FACULTY OF ELECTRICAL ENGINEERING
UNIVERSITY OF BANJA LUKA

X International Symposium
INDUSTRIAL ELECTRONICS

INDEL – 2014

SYMPOSIUM PROCEEDINGS

ISBN 978-99955-46-22-9

Banja Luka
Republika Srpska
November 6–8, 2014

X International Symposium
INDUSTRIAL ELECTRONICS
INDEL – 2014

Banja Luka

November 6–8, 2014

Organized by

FACULTY OF ELECTRICAL ENGINEERING
UNIVERSITY OF BANJA LUKA

Patre 5
78000 Banja Luka
Republika Srpska
Bosnia and Herzegovina

Phone: +387 (0)51 - 221 - 820
Dean Office: +387 (0)51 - 221 - 824
Fax: +387 (0)51 - 211 - 408
Web: www.etfbl.net
E-mail: office@etfbl.net

Under the auspices of

MINISTRY OF SCIENCE AND TECHNOLOGY
REPUBLIC OF SRPSKA GOVERNMENT

Technical Co-Sponsor

Symposium partners and sponsors

ECONOMICS INSTITUTE BANJA LUKA
NOVA BANKA
NLB RAZVOJNA BANKA

Symposium Chairman

Prof. dr Branko Dokić, University of Banja Luka, Bosnia and Herzegovina

Program Committee

Dimitar Alexiev, Technical University of Sofia
Goce Arsov, University St. Cyril and Methodius, Macedonia
Zdenka Babić, University of Banja Luka, Bosnia and Herzegovina
Narcis Behlilović, University of Sarajevo, Bosnia and Herzegovina
Petar Biljanović, University of Zagreb, Croatia
Branko Blanuša, University of Banja Luka, Bosnia and Herzegovina
Milorad Božić, University of Banja Luka, Bosnia and Herzegovina
Dražen Brđanin, University of Banja Luka, Bosnia and Herzegovina
Zlatko Bundalo, University of Banja Luka, Bosnia and Herzegovina
Branko Dokić, University of Banja Luka, Bosnia and Herzegovina
Gordana Gardašević, University of Banja Luka, Bosnia and Herzegovina
Cvetan Gavrovski, University St. Cyril and Methodius, Macedonia
Vladimir Katić, University of Novi Sad, Serbia
Tom Kazmierski, University of Southampton, United Kingdom
Branko Kovačević, University of Belgrade, Serbia
Igor Krčmar, University of Banja Luka, Bosnia and Herzegovina
Božidar Krstajić, University of East Sarajevo, Bosnia and Herzegovina
Mirza Kušljugić, University of Tuzla, Bosnia and Herzegovina
Vančo Litovski, Niš Cluster of Advanced Technologies, Serbia
Duško Lukač, University of Applied Sciences, Germany
Danilo Mandić, Imperial College, United Kingdom
Dragan Mančić, University of Niš, Serbia
Darko Marčetić, University of Novi Sad, Serbia
Petar Marić, University of Banja Luka, Bosnia and Herzegovina
Slavko Marić, University of Banja Luka, Bosnia and Herzegovina
Petar Matić, University of Banja Luka, Bosnia and Herzegovina
Octavio Nieto-Taladriz Garcia, Polytechnic University of Madrid, Spain
Vojin Oklobdžija, New Mexico State University, USA
Predrag Pejović, University of Belgrade, Serbia
Tatjana Pešić-Brđanin, University of Banja Luka, Bosnia and Herzegovina
Predrag Petković, University of Niš, Serbia
Nikola Rajaković, University of Belgrade, Serbia
Predrag Rapajić, University of Greenwich, United Kingdom
Miran Rodič, University of Maribor, Slovenia
Ferid Softić, University of Banja Luka, Bosnia and Herzegovina
Goran Stojanović, University of Novi Sad, Serbia
Milić Stojić, University of Belgrade, Serbia
Nikola Teslić, University of Novi Sad, Serbia
Slobodan Vukosavić, University of Belgrade, Serbia
Čedomir Zeljković, University of Banja Luka, Bosnia and Herzegovina
Volker Zerbe, University of Applied Sciences of Erfurt, Germany
Miloš Živanov, University of Novi Sad, Serbia

Organizing Committee

Branko Dokić, University of Banja Luka, Bosnia and Herzegovina
Petar Marić, University of Banja Luka, Bosnia and Herzegovina
Zoran Đurić, University of Banja Luka, Bosnia and Herzegovina
Tatjana Pešić-Brđanin, University of Banja Luka, Bosnia and Herzegovina
Igor Radojičić, National Assembly of the Republic of Srpska, Bosnia and Herzegovina
Duško Torbica, ELNOS, Bosnia and Herzegovina
Dragan Praštalo, Mikroelektronika, Bosnia and Herzegovina
Dragoljub Davidović, The union of energy professionals of Republic of Srpska, Bosnia and Herzegovina
Željko Jungić, Adviser to the Director of strategy department at Telekom Srpske a.d., Bosna and Herzegovina
Branko Blanuša, University of Banja Luka, Bosnia and Herzegovina
Petar Matić, University of Banja Luka, Bosnia and Herzegovina
Igor Krčmar, University of Banja Luka, Bosnia and Herzegovina
Čedomir Zeljković, University of Banja Luka, Bosnia and Herzegovina
Željko Ivanović, University of Banja Luka, Bosnia and Herzegovina
Mladen Knežić, University of Banja Luka, Bosnia and Herzegovina

Technical Secretaries

Đorđe Lekić, University of Banja Luka, Bosnia and Herzegovina
Predrag Mršić, University of Banja Luka, Bosnia and Herzegovina
Vanja Todorović, University of Banja Luka, Bosnia and Herzegovina
Milosava Radonjić, University of Banja Luka, Bosnia and Herzegovina
Bojana Trbić, University of Banja Luka, Bosnia and Herzegovina

www.indel.etfbl.net
indel@etfbl.net

PREFACE

The first symposium INDUSTRIAL ELECTRONICS – INDEL was held in 1997, and since 1998 it is organized every other year in the first half of November. So far, 461 papers were presented, which in average means 51 presented papers per Symposium. All presented papers were printed in the Symposium Proceedings. The following topics, in the area of application of electrical engineering and information and communication technologies in industry, are covered by the Symposium: materials and components, power electronics, circuits and systems, electrical machines and drives, measurement methods and systems, signal processing, modelling, identification and process control, renewable energy, signal processing in electric power, and information technologies. Since 2012, the Symposium also covers the topic of energy efficiency. The first Symposium (INDEL 1997) was held on national level, whereas since 2010 the Symposium is under IEEE technical co-sponsorship. Highly respected scientists from all around the world (all ex-Yugoslavia states, USA, England, Germany, Austria, Bulgaria) are in the Symposium Program Committee. At the 9th INDEL – 2012 authors from Austria, B&H, France, Germany, Iran, Italy, Macedonia, Serbia, Slovenia, Spain, Turkey, Great Britain, and the USA, presented their papers. All papers were reviewed, and since 2010 every accepted paper has to have two positive reviews.

At plenary sessions, respected scientists and entrepreneurs presented introductory lectures from then current thematic fields as: "40 Years of the Faculty of Electrical Engineering Banja Luka", "TEMPUS Projects and Cooperation between Higher Education Institutions in the Region", "150 Birth Anniversary of Nikola Tesla", "Renewable Energy – Wind Power Use", "Power Converters in the Context of Environmental Equipment", "Stimulus for Preferential Producers of Electrical Power in Serbia", "Electronic Power Measuring Devices", "Principles of High Speed Computational Processes", "Technological and Logical Aspects of CMOS VLSI Circuits with Low Consumption", "Equal Ripple Optimization of Generalized Chebyshew Low-Pass Filters", "Power Quality – Measurements and Analysis of Voltage Drop", "Electric Cars". While celebrating 50 years of foundation of the Faculty of Electrical Engineering in Banja Luka, three plenary lectures were presented: "50 Years of Electronics in Banja Luka", "Cooperation between the Faculty of Electronics Niš and Faculty of Electrical Engineering Banja Luka", and "Cooperation with European (UK) Universities in Research and Education". Plenary session authors were, among others: Vojin Oklobdžija (USA), Branko Dokić (FEE Banja Luka), Milić Stojić (FEE Belgrade), Vladimir Katić (FTS Novi Sad), Vančo Litovski (FE Niš), Slobodan Vukosavić (FEE Belgrade), Volker Zerbe (Germany), Tom Kazmierski (England), Nedeljko Perić (FEEC Zagreb), Nikola Rajaković (FEE Belgrade), Predrag Rapajić (England), Predrag Pejović (FEE Belgrade), Duško Lukač (Germany), Octavio Nieto-Taladriz (Spain), Goce Arsov (FEE Skopje), Dejan Raca (Austria) and others.

Briefly, I will summarize the contents of the plenary and keynote sessions of the 9th Symposium INDEL 2012. Plenary session was dedicated to the 50th anniversary of the Faculty of Electrical Engineering in Banja Luka. On that occasion, three opening speeches were given. Prof. dr Vančo Litovski spoke about the cooperation between the Faculty of Electronic Engineering in Niš and the Faculty of Electrical Engineering in Banja Luka. Prof. dr Predrag Rapajić from the University of Greenwich spoke about the cooperation with European universities. The speech about 50 years of electronics in Banja Luka was given by prof. dr Branko Dokić. Notes from that speech are comprised in the following:

"Essentially, development of electronics in Banja Luka has started a little more than fifty years ago, with engineers which were brought to Banja Luka from bigger centres of the former Yugoslavia. Finest years of electronics in Banja Luka were the eighties of the last century, when "Cajavec" and "Kosmos", together had about 12000 employees. In Banja Luka, the first Yugoslav microelectronics circuits and microprocessor were produced. Further, Banja Luka was the place where radar systems, specialized computer and communications devices, electronic devices for households, automobile equipment, medical equipment, measuring devices, detectors, etc. were designed and produced. "Cajavec" had a share of approximately 45% of military electronics market in former Yugoslavia. Beginning of dissolution of SFRY is, at the same time, the beginning of decline of electronic devices production. Four year long war left numerous consequences. There was a significant loss of domestic and especially foreign markets. Also, the pace of technological development was lost. Inappropriate and irresponsible concept of economy transition just speeded up the fall. In what where production facilities and scientific research laboratories of "Cajavec" and "Kosmos", today there are private faculties, universities, various schools, wedding and funeral saloons, etc. Out of three institutions that carried the development of electronics in Banja Luka, just the Faculty of Electrical Engineering has remained. The first beginning was in the overall poverty and without educated professionals, but those were the times rich with enthusiasm of the youth, fully aware that it had to work on itself. Today we have professionals, but we live in the time without enthusiasm, when we wait for somebody to some form outside and to invest in the development of modern technologies. The sooner we get rid of this misconception and unite individual enthusiasms and turn to ourselves, the sooner we will have a new beginning and see a new beauty of modern technologies. Is this enough for a new beginning and a long expected rise?"

Co-chair of the keynote session, at which two keynote speeches were given, was prof. dr Predrag Pejović from the School of Electrical Engineering, University of Belgrade. Prof. dr Tom J. Kazmierski from the University of Southampton

had a notable lecture on the subject of "Energy Harvesting in Electronics: From Low Power to Zero Power". He gave a review of the academic and industrial state of the art in the energy harvesting technology with specific examples of small scale energy harvester systems which were recently developed and put into commercial practice. Prof. dr Octavio Nieto-Taladriz from the Polytechnic University of Madrid had an interesting lecture about "Cognitive Wireless Sensor Networks – New challenges". He gave an introduction to the concept of Cognitive Wireless Sensor Network (CWSN) and discussed important characteristics of sensor node and sensor network architectures.

I would like to emphasize a very interesting lecture "150 Birth Anniversary of Nikola Tesla", which was presented by prof. dr Milić Stojić at the plenary session of the 6th INDEL 2006. I have chosen this lecture as for the importance of impressive work of Nikola Tesla, and the way prof. Stojić presented it. Here is the abstract of that lecture:

"Nikola Tesla (1856-1943) significantly influenced technological development with his polyphase system inventions. The system is in cornerstone of modern electrical energy production, long-distance transmission, and use of electrical currents. Beside inventing the induction motor, he invented the Tesla coil — a high frequency transformer, which is an essential part of all contemporary high frequency devices. Tesla also pioneered research into other effects produced by his currents, such as the possibility of induction heating, ozone production, and the effects on the human organism. His inventions have been crucial to the development of many of today's technologies including the radio, radar, television, motors of all kinds, and computers. He is also credited with predicting the emerging energy problem as early as 1900. After death of Nikola Tesla in 1943, all his belongings have been inherited by his nephew and transferred to Belgrade where in 1955 the Nikola Tesla Museum has been opened. His ashes are also in the Museum. After his death, the name "Tesla" was given to the unit of magnetic induction.

The Nikola Tesla Archive in Belgrade (Serbia) constitutes a unique collection of over 160000 pages of the patents documentations, scientific correspondence, scientific papers, manuscripts, technical drawings, scientific measuring data, personal documents, and legal papers as well as over 1000 original photographs of Tesla's experiments and inventions, all of which are indispensable to the study of the history of electrification. Nikola Tesla's Archive in Belgrade joins Memory of the World register."

Sponsors of the Symposium were: City of Banja Luka, Telekom Srpske, Power utility of the Republic of Srpska, Siemens d.o.o., Elektrokrajina Banja Luka, Elektroprenos Banja Luka, Nokia Siemens Networks Banja Luka, Government of the Republic of Srpska, Ministry of Science and Technology, Elnos, Caldera Company, Integra, Inova.

Since 2006 accompanying activity of the symposium is the International Students Competition Hardware&Software.

The 10th International Symposium Industrial Electronics – INDEL 2014 was held in Banja Luka on November, 6-8, 2014. The Symposium was organized by Faculty of Electrical Engineering, University of Banja Luka, and under the auspices of Ministry of Science and Technology, Republic of Srpska Government. The Symposium was organized under the technical co-sponsorship of IEEE Bosnia and Herzegovina Section.

At the Opening Ceremony, awards were presented to scientists and organizations that have significantly contributed to the organization of the symposium since its inception:

- Milić Stojić, Faculty of Electrical Engineering, Belgrade, Serbia
- Nikola Rajaković, Faculty of Electrical Engineering, Belgrade, Serbia
- Slobodan Vukosavić, Faculty of Electrical Engineering, Belgrade, Serbia
- Predrag Pejović, Faculty of Electrical Engineering, Belgrade, Serbia
- Vladimir Katić, Faculty of Technical Sciences, Novi Sad, Serbia
- Miloš Živanov, Faculty of Technical Sciences, Novi Sad, Serbia
- Vančo Litovski, Faculty of Electronic Engineering, Niš, Serbia
- Predrag Petković, Faculty of Electronic Engineering, Niš, Serbia
- Ministry of Science and Technology, Republic of Srpska Government
- Power Utility of Republic of Srpska
- Society of ETRAN, Serbia

Symposium was opened by the President of National Assembly of the Republic of Srpska, mr. Igor Radojičić.

Plenary session was dedicated to the the electric vehicles. On that occasion, three opening lectures were given. Tamás Ruzsányi, Executive Board member of IEEE Industry Application Society, spoked about new generation of public transport vehicles for replacing diesel buses. The speech about plug-in electric vehicles was given by prof. Vladimir Katić from Faculty of Technical Sciences, Novi Sad. Miloš Pilipović, from RT-RK Institute from Novi Sad spoked about

intelligent driver-assist technologies and piloted driving. Co-chairs of the plenary session were prof. Predrag Pejović, Faculty of Electrical Engineering, Belgrade, and prof. Goce Arsov, University St. Cyril and Methodius, Skopje.

We would like to highlight that two round tables, *Digital Television* and *Is the Sun able to cool?*, were organized in addition to the regular sessions. The second one was organized in cooperation with the Economics Institute Banja Luka.

Regular papers were presented in the following sessions: *Materials and components*, *Power electronics*, *Circuits and systems*, *Electrical machines and drives*, *Measurement methods and systems*, *Modeling, identification and process control*, *Renewable energy and energy efficiency* and *Telecommunication technologies*, and merged sessions *Signal processing* and *Information technologies*. For the first time, along with regular oral sessions, the poster session was organized. In regular session, the following invited papers were presented: Zoran Jakšić, "Optical Chemical Sensors for Industrial Application", Miroslav Lazić, "The Use of Microcontrollers in Modern Solutions of Power Electronics", Darko Brodić, "Script Recognition by Statistical Analysis of the Image Texture", and Milan Bjelica "The Personalization of a Cloud Ecosystem: Adding Dimensions to Situational Awareness".

Symposium Chairman
Prof. dr Branko L. Dokić

Symposium Program Chair
Prof. dr Tatjana Pešić-Brđanin

SYMPOSIUM TOPICS AND SESSIONS

Plenary session
Co-chair(s): **Goce Arsov and Predrag Pejović**

T-01: Materials and components
Co-chair(s): **Goce Arsov**

T-02: Power electronics
Co-chair(s): **Branko Blanuša**

T-03: Circuits and systems
Co-chair(s): **Vančo Litovski**

T-04: Electrical machines and drives
Co-chair(s): **Petar Matić**

T-05: Measurement methods and systems
Co-chair(s): **Igor Krčmar**

T-06 and T-09: Signal processing and Information technologies
Co-chair(s): **Vladimir Risojević**

T-07: Modeling, identification and process control
Co-chair(s): **Milorad Božić**

T-08: Renewable energy and energy efficiency
Co-chair(s): **Vladimir Katić**

T-10: Telecommunication technologies
Co-chair(s): **Gordana Gardašević**

Poster session
Co-chair(s): **Čedomir Zeljković**

CONTENT

PLENARY SESSION

Plug-in Electric Vehicles – State of Technology and Market Perspectives

Invited Paper

Vladimir A. Katić
University of Novi Sad, Faculty of Technical Sciences
Novi Sad, Serbia
katav@uns.ac.rs

***Abstract* — Transportation sector is contributing in great percentage (23%) to GHG emission in EU countries. In order to reach EU 202020 targets significant efforts to reduce vehicles´ exhausting gases have been made. Plug-in vehicles are foreseen as a part of solution, as it is expected that their share in GHG emission will be decreased below 10% in 2050. In the paper, an overview of plug-in vehicles technologies is presented. All systems are classified and details of electric power processing units are given. Special attention is placed to market perspectives of presented solutions. A review of present state in plug-in passenger cars production and prediction for coming years are given.**

Keywords - Plug-in Vehicles, Power train technologies, Market perspectives

I. INTRODUCTION

Although EU 2020 policy target to decrease GHG emission by 20% in 2020, the transportation sector will actually increase the emission and their share up to 26% in 2020 from 16.5% in 1990 [1]. This alarming trend, which is presented with dotted line in Fig.1, needs swift and effective counter-measures and additional policies applied in a short time. Effects of these measures will hold the GHG emission rise at the level of 22%. Significant reduction is possible if low and zero emission vehicles are implemented. In this case it is expected that by 2050 the transportation sector will reduce GHG emission down to 7% - 9% (Fig.1 [1]).

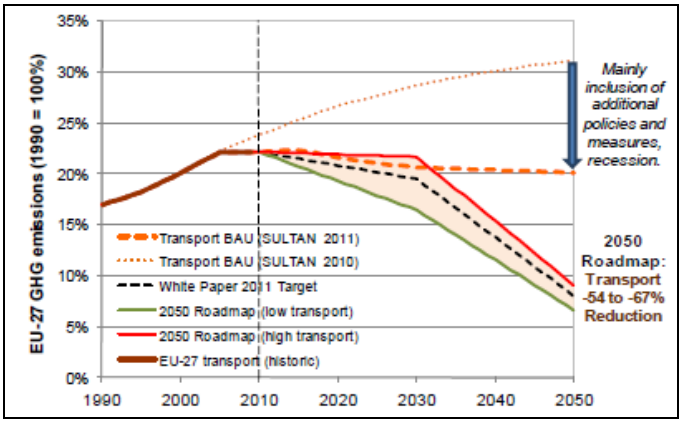


Figure 1. GHG emission of transportation sector [1]

However, there are several options of technological solution, which may ensure such future mobility. This can be achieved with the development of mobility technologies and by focusing on the sustainable business concepts. The solution should include social aspects, environmental issues and economic views, also [2]. Besides electric mobility, management strategies must also consider other alternatives like hydrogen, biogas or fuel cells. These strategies actually are considering different energy sources and engine types, although the vehicle's exhaust gasses depend on applied engine and its efficiency. In that sense (if the propulsion motor is considered) possible vehicles may be classified into three categories:

1. IC engine vehicles - gasoline, diesel, biodiesel, natural gas (CNG, LNG, LPG), hydrogen (H_2)
2. Electric motor vehicles – Battery, Fuel-cells
3. Hybrid ICE/electric motor vehicles – series, parallel, series/parallel

Regarding reduction of GHG emission, future development is oriented toward low emission vehicles (LEV). Above classification indicates that there are several options for such development. Many research studies predict that there will be no single solution, i.e. that different vehicle types will be competing on the market [3, 4, 5].

The study prepared for the European Commission predicts significant improvements in transportation power train efficiency [3]. This will be achieved mainly by fuel substitution; in particular switching from diesel to electricity in areas where electrification is an economically viable option Fig. 2 shows two projections for future development (up to 2050): shares by vehicles´ structure in car stocks (in %, left bars) and by fuel (energy) consumption by cars (in %, right bars). Left bars show significant decrease of classical IC engines (gasoline conversion and diesel conversion) share - from 96% in 2010 down to 55% in 2050. These vehicles will be replaced mainly by electric and hybrid ones and in smaller percentages by LPG and CNG powered cars. Plug-in hybrid vehicles (PHEVs) will hold the largest share among this new group due to their ability to use both power-trains alternatively (ICE or electric motor). However, it will not mean that the oil consumption (gasoline and diesel) will be decreased in the same percentage – from 92% it will drop to 81%, only.

Another study shows that annual sales of passenger light-duty vehicles LDV sales will dramatically change structure in next decades (Fig.3). Conventional gasoline and diesel cars share will drop from 98% in 2010 down to only 8.5% in 2050, while

hybrid, plug-in hybrid and electrical vehicles will grow up to their domination in the market (71% in 2050) [4].

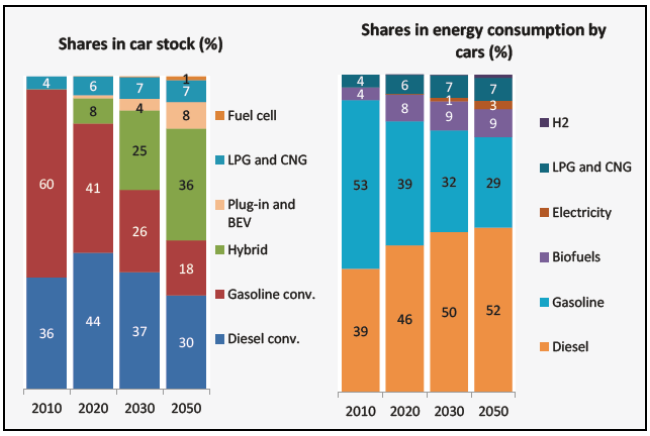


Figure 2. Projections of vehicles´ structures and fuels consumptions [3].

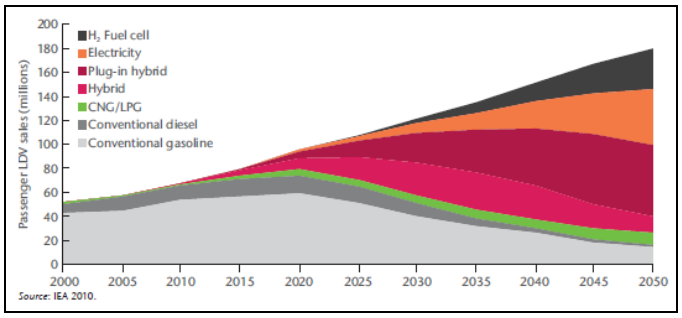


Figure 3. Annual (LDV) sales by technology type (Blue map scenario) [4]

Still, some authors are more conservative and show that different models for analyzing future developments do not give (predict) single vehicle or fuel technology for sustainable mobility to dominate in next decades [5]. They concluded that both regarding the lowest cost solutions and carbon constrains, only a diversity of technology combinations may give acceptable effects, so they highlights the need for research and development of multiple technologies. Still, the main technologies are based on electrical and hybrid sources.

The success of electric mobility products depends on the fulfilment of users' expectations. But in the past, electric vehicles (EV) were not able to meet all the needs of users and to survive on the market. The first ones appeared way back in the past, at the end of 19th century. In early 20th century they were relatively popular (38% of U.S. Market) until the internal combustion engine displaced them, due to better driving range and open-road energy supply infrastructure [6, 7]. They re-appeared in 1970s, when oil prices increased rapidly and fears about possible depletion of fossil fuels started to upset the public. Although environmental awareness contributed also to significant investments to R&D in this field, the resources soon became very limited, as driving range limitation and gasoline prices decrease discourage further investments. Fig. 4 shows U.S. market data – regular gasoline ($/gallon) vs. residential electricity ($/gallon equivalent). It can be seen that in that period these prices were very close, so there is no economic interest for the users to switch from ICE to electric motor vehicles [8].

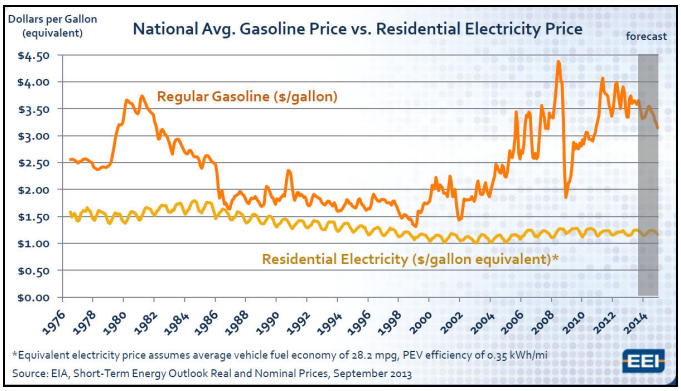


Figure 4. US average gasoline prices vs. Residential electricity prices – a long term overview.

In the last decade the EVs are entering the third age as several factors are working in their favour, like enhanced technology, extended range, environmental awareness, very good performances, significantly lower energy prices, much better energy efficiency, etc. Today, all major car manufacturer companies are offering such products in a form of electric or hybrid ICE/electric motor vehicles (dual fuel vehicles). They are known with common name plug-in electric vehicles (PEVs).

The paper gives an overview of present technology of PEVs and some future trends, with special look on market perspectives. The technical aspects are considered regarding driving train, while market trends are stated according to applied technology and present research studies.

II. PLUG-IN ELECTRIC VECHICLES CLASSIFICATION

Usually the PEVs are classified into plug-in hybrid electric vehicles (PHEVs), extended-range hybrid electric vehicles (E-RHEVs), or battery electric vehicles (BEVs) [7, 9]. Such classification is presented in Fig. 5.

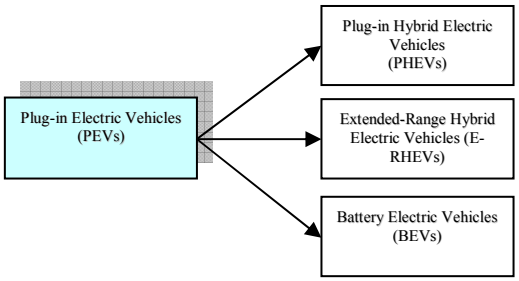


Figure 5. Standard PEVs classification

The PHEV is an evolutionary extension (successor) of the hybrid electric vehicle (HEV), which originated in the nineties in the form of the famous Toyota's Prius model [7]. The HEVs may be classified into serial HEVs, parallel HEVs and serial/parallel HEVs [7, 9, 10, 11]. They have different engines – ICE and electrical ones, so they are often called dual fuel vehicles. In that sense, they represent the first step in ICE vehicles electrification,

which is further extended with PHEVs. The main advantage of the PHEVs, beside enlarged capacity of battery is ability to charge the battery from external power supply in contrary to HEVs where the battery is charged only internally (by ICE-Electric generator train or by energy recuperation). The PHEVs may be divided in the same manner into serial PHEVs, parallel PHEVs, and serial/parallel PHEVs [11].

The extended-range hybrid electric vehicles (E-RHEVs) are developed to overcome the main disadvantage of electric vehicles – the limited range. The batteries are enlarged and hybrid structure has been applied. With appearance of PHEVs this feature of E-RHEVs lost its importance. However, additional options to achieve range extensions have been proposed using renewable energy or different fuel-based units [12]. The most feasible solution for renewable energy range extender is application of solar arrays. It could be mounted on the vehicle top (internal source) to supply energy to battery during vehicle operation and parking or used as an external source for battery charging during vehicle parking time. The fuel-based range extenders use fuel-cells and ICE/microturbine generators. The fuel-cells are delivering electrical energy for battery charging, which has been transform from hydrogen (H_2). They could be used as the only source of electricity, but in that case, they are classified as fuel-cells electric vehicle (FCEV) and represent a special type of EVs. The ICE/microturbine generator is not popular solution for passenger cars as it requires additional space and its operation is connected with audio noise.

The battery electric vehicles (BEVs) are also known as full-electric or simply electric vehicles (EVs). Their main source of power is electric battery, which converts chemical energy into electrical one and vice-versa. The energy capacity of lead-acid batteries used to be the main obstacle for wider application of EVs. Today, more advanced battery technologies in a form of NiMH and lithium-ion batteries are giving promising features for further spreading of this type of vehicle [13].

Beside batteries, other electrical storage units may be also implemented - ultra-capacitors, making BEV´s energy storage system a hybrid one. For efficient application of these sources power converters (DC/DC) with specialized control strategies are implemented. Depending on the level of power converters application passive hybrids (parallel connection of battery and ultra-capacitor without power conversion units), semi-active hybrids (power converters are applied for control either battery or ultra-capacitor source) and active hybrids (power converters are applied for control of both battery and ultra-capacitor sources) are distinguished [12].

To enable power transfer, conversion and processing in power flow of electrical energy a complex internal electrical power system is needed [9, 11, 13, 14, 15, 16, 17]. Both DC and AC voltage lines are may be used, so BEVs with single or dual voltage systems can be distinguished. In single voltage systems a low DC voltage level of 12V or 42V are applied. For a wider driving range, such system has limiting transfer capabilities, so dual one is implemented. Dual power system combines DC and AC voltage lines. The DC grid consists of low voltage DC (12V) and high voltage DC (200-400-600V), while AC voltage of 600V is used for AC motor [15]. A schematic of a BEV´s electric power dual voltage system is presented in Fig. 6 [16]. In both single and dual power systems, AC voltage exists at the battery charger inputs (230V/400V), also.

Although different electric motor types may be used for propulsion purposes in PHEVS, E-RHEVs or BEVs, classification regarding this item is not preset in literature or practice. Still, it should be noted that serial DC motor and a number of AC motor types (induction, synchronous, permanent magnet (PM), switched reluctance) are used [9, 13].

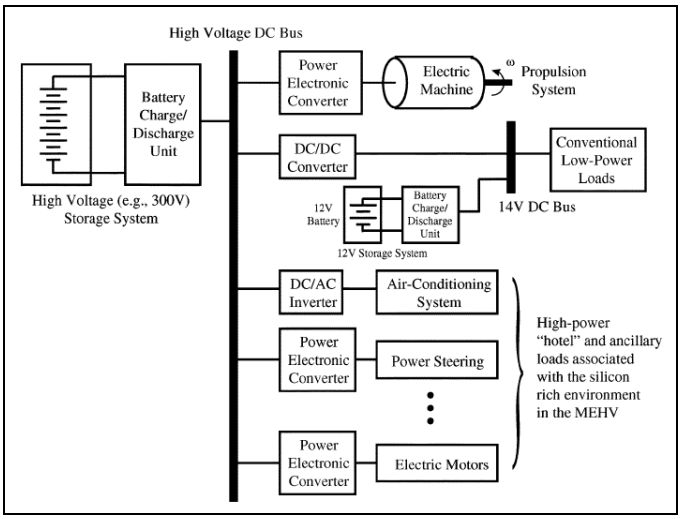


Figure 6. BEV´s electric power system configuration [16].

III. PEV´S POWER FLOW

The conventional ICE vehicle power train consists of energy source (tank), hydraulic link (fuel delivery and preparation), engine (ICE), transmission, a drive shaft and wheels (Fig. 7.a). The plug-in hybrid vehicles bring in power-train the electric system and electric motors/generators units.

In series connection hybrid power train is supplied from an ICE and from batteries (Fig. 7.b). The ICE is driving an electrical generator producing electricity for propulsion done by an electric motor. In excess of electrical energy or in motor recuperating mode the electrical energy is stored in batteries. The batteries are also supplying the electric motor, but their role is restricted, due to limited stored energy. The batteries can be charged from outside, using on-board or off-board level 1 or mode 1 chargers [18, 19].

Parallel plug-in hybrid vehicles have both ICE and electrical motor for propulsion, connected in parallel (Fig. 7.c). To achieve low emission, electric power train is used for start and acceleration, while ICE is used in steady driving conditions and for range extension.

A combination of the series and parallel power-train is called series/parallel plug-in hybrid (Fig. 7.d). Again, the batteries may be charged from outside, using on-board or off-board chargers. In this case ICE may be used directly for propulsion, but also for driving an electricity generator for battery charging.

The BEVs power flow is presented in Fig. 7.e. Huge batteries are now the main (and only) energy source for propulsion. Advanced energy management is applied, so energy is stored in batteries during braking and downhill driving. For start fast

acceleration a supercapacitor module may be added. The batteries are charged using on-board or off-board chargers, but the charging time can be shortened down to 20 min. (level 3 or mode 3, 4 chargers) [18, 19].

IV. PEV´S TECHNOLOGY

For transferring energy from the battery to the electric motor a complex power system with different power electronics converters is needed. In such a system four main areas can be distinguished:

1. High voltage energy source/storage system with charging-discharging unit
2. Electric Propulsion system
3. Low voltage energy source/storage system with conventional low power loads.
4. High power grid with ancillary loads

Many books, papers, expertises, studies and other documents have been devoted to these topics, so in this paper they will be only slightly mentioned [7, 9, 11, 12, 14, 20, 21, 22, 23].

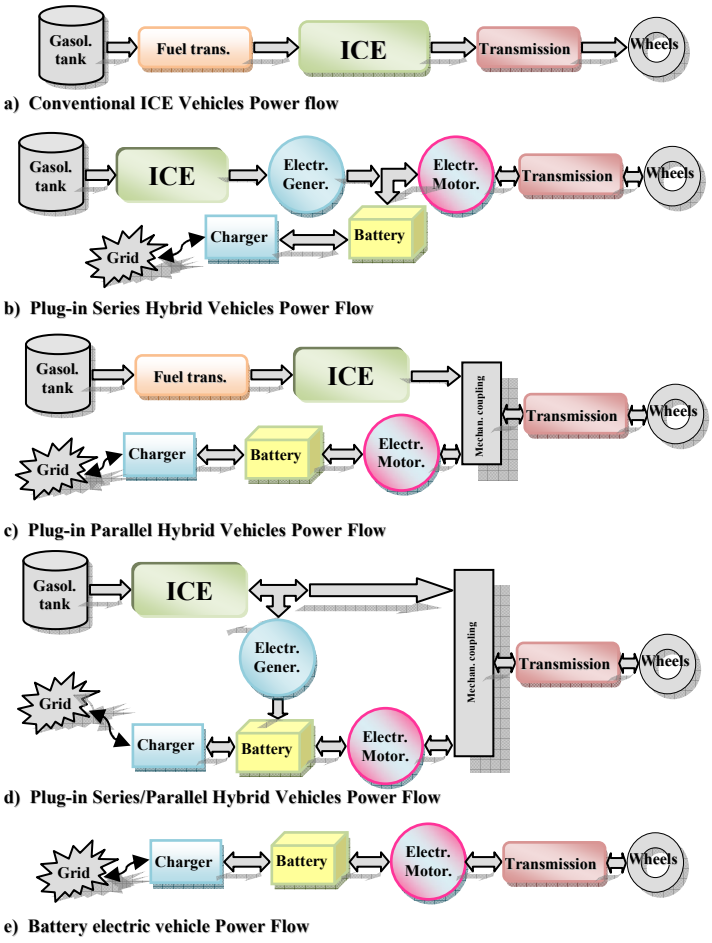


Figure 7. The structure of an ICE, different plug-in hybrid and electric vehicles

A. *HV energy source/storage and charging/discharging systems*

The main sources of electrical energy for propulsion are batteries. They form a high voltage source (200-600V) using stored electrical energy [13]. But, the energy supply is obtained from battery charging stations, where the energy, coming from the public grid is converted into electrical one in a form suitable for transferring it to the batteries. Different types of electric chargers are available [14].

Batteries have dual role in EVs, as electrical energy storage unit in charging mode and energy source in powering mode. The main features are energy density (in Wh/kg), which define the driving range, and power density (in W/kg), which define start and acceleration capabilities. Important characteristics are weight, cycle life, volume and costs. Four types of batteries were considered for the BEVs - Lead Acid, Nickel Cadmium (NI-Cd), Nickel-Metal hydride (NiMH) and Lithium Ion (Li Ion), but nowadays Li Ions are becoming dominant. Comparison of their main characteristics is given in Fig. 8 [24]. In future some improvements in lithium based batteries (Lithium-Metal, Lithium-Polymer, Lithium Phosphate....) are expected, although concerns on available Li reserves are present.

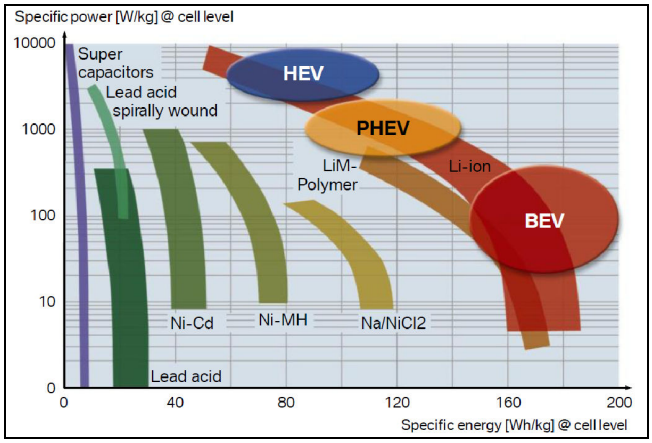


Figure 8. Power and energy densities of selected energy storage technologies [24]

For charging process a special power converter is needed. The charger could be placed off-board or on-board (lower size, long time charging). On-board charging systems can be conductive or inductive. Conductive chargers have wired (direct) contact with the grid using special socket/plug unit, while inductive have wireless (magnetic) transfer of energy [11, 14, 18, 19, 25].

In conductive system a grid-isolated AC/DC converter system is needed. Such a converter in a grid-to-vehicle (G2V) power flow should enable AC voltage rectification (AC/DC), voltage level adaptation (DC/DC) and galvanic isolation of vehicle from the grid (DC/DC). Fig. 9 presents a typical schematic of this charger supplied from a three-phase public grid.

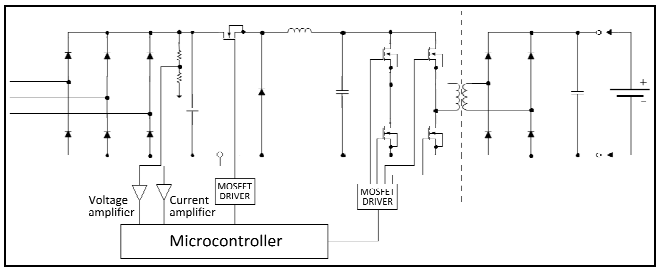
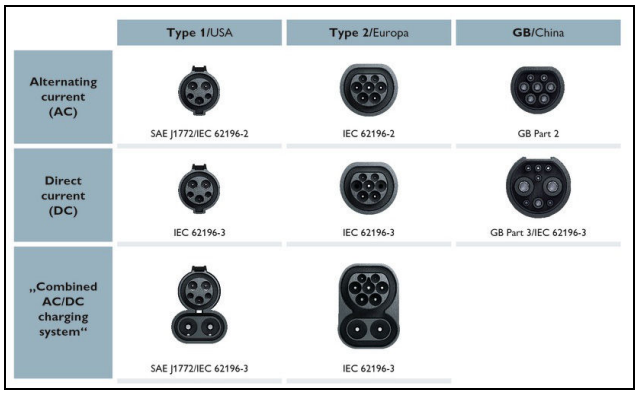


Figure 9. Topology of battery charging station.

Depending on its power size and complexity three levels are distinguished in USA [14, 18], and four charging modes in EU [14, 19, 25]. Usually a charger is on-board (Level 1, Mode 1 or 2), giving EV flexibility and autonomy. Due to weight and size limitation for on-board chargers, recently chargers are placed in public, outside the vehicle, off-board, forming a wide network of battery charging station and completing energy supply infrastructure for EVs. They enable fast charging in 15-20 min or moderate speed charging in 2-3 hours (Level 3, Mode 3 or 4). The main obstacle is different plugs and sockets, which have been developed Worldwide. Three types can be distinguished and they are presented in Fig.10.a. A plug and inlet of the combined AC/DC charging system developed in Europe is presented in Fig. 10.b.

a)



b)

Figure 10. a) Charging plug-in systems Worldwide; b) DC and AC Charging plugs Type 2/Europe (left) and inlet of the combined AC/DC plug-in system (right)

Future solutions predict a case of vehicle-to-grid (V2G) operation, so above mentioned power converters should be constructed in a way to provide bilateral power flow [26].

Some of the proposals gives unified role of traction hardware (electric motor and inverter) and charging one making their functions integrated - integrated charger [14, 27]. In that case problems of charger weight, space, and cost are avoided. The integration may also allow galvanic isolation, voltage level adaption, better efficiency, low current harmonic content, and mandatory unity power factor operation. Application of multi-phase electric motor enables easier charger integration and improved reliability of the whole driving train [28].

Inductive chargers are attractive as there is no wired contact (cables and cords are eliminated) [14, 22]. The energy is transferred via electromagnetic coupling, which makes them very convenient for the user. There are several solutions for the coupling itself, but all are based on high-frequency off-board AC source and on-board pickup inductance (charge port) which should be brought together on close distance. Disadvantage includes relatively low efficiency and power density, manufacturing complexity, size, and costs.

B. Electric propulsion system

Electric propulsion system consists of high-power DC/AC converter (traction inverter), which is connected to the HV DC grid, and traction motor with mechanical coupling (transmission) to the wheels. In some cases (4 wheel drive) the motors may be built into the wheels, so separate inverter for driving each motor is needed. There is a huge experience and practical solutions from industrial electric drives, including very sophisticated control algorithms [9, 20, 23, 29, 30, 31, 32].

The traction inverter, which function is to provide AC power to the main, traction AC motor, has input voltage range between 190 V and 400 V. The inverter is H bridge topology, composed of IGBT switches with free-wheeling diodes and controlled with space vector modulated PWM or similar advanced control algorithms. Different other solutions, like multi-level inverters, multi-phase configurations and others, have been proposed in order to provide easier and more precise control, increased reliability, higher efficiency and lower costs [13, 29, 33]. Although these converters are operating in switch-mode, with high efficiency and low losses, a sophisticated energy management is needed to coordinate energy flow and provide rational energy usage. Further improvements in these directions are expected in the future.

A number of electric motor types are considered for a role of traction motor. Traditionally the DC motor is used in different electrical vehicles. It offers good driving characteristics and very simple and cost-effective speed control. However, high maintenance costs, due to sparkling and a need for frequent brush replacement, size and weight made DC motor unattractive in comparison to AC ones.

Most of today's solutions are based on different AC motors. Induction or synchronous AC motors are used as traction motors, due to their lower weight and costs, higher reliability and lower maintenance requirements.

In HEVs and EVs interior permanent magnet synchronous machine (IPMSM) is employed in most of current available models because of its high efficiency, high torque, high power density, and relatively ease of field weakening operation. For example Toyota Prius is using 60 kW motor and 42 kW generator, Chevy Volt model has 110 kW motor and 55 kW generator built-in and Nissan Leaf a 80 kW machine. For four wheel drive vehicles, IPMSM are mounted inside the wheel,

eliminating mechanical gears and differential. This gives higher efficiency, less weight, and improved reliability, but has usual size and weight restriction. The problem with all PM machines is that their price strongly depends on current market conditions and availability of rare earth-based magnets.

For high power propulsion, induction motor is used. For example, Tesla Roadster is using a 3-phase 4-pole induction motor of 185 kW power and with maximum speed of 6,000 rpm. They have reasonably good performance, simple structure and people have long experience in their industrial drives applications. Still, some improvements in increasing operating voltage, using copper cage rotors and custom design for automotive applications are possible [13, 30].

Significant research work is going on to improve operational characteristics of switched reluctance machine (SRM), which is a type of synchronous machine and has the simplest mechanical design. At the moment they are extremely noisy, have torque pulsation, and larger size and weight in comparison to PM machines. They also have lower efficiency, but the lowest costs and most robust structure [13, 23, 30]. They are seen as possible alternative solution EVs propulsion engine, if costs of rare earth magnets become too high.

C. *Low voltage energy source/storage system with conventional low power loads.*

This part of the system is dedicated to the conventional low-power loads, which operate on 14V (or 12V) bus and which can be seen at the ICE vehicles, as well. As number of such loads is growing, there were concerns about ability of standard battery to supply them. In the past decade, very strong ideas to introduce 42V bus in traditional ICE vehicles, instead of 14V existed, in order to meet increased electricity demand [16]. However, it was decided not to make such a change, so this part of the system remains in current shape.

D. *High power grid with ancillary loads*

Other electrical loads, which require higher power is planned to be attached to the HV grid. It will enable installation of more powerful electrical and electronic equipment, and widening the scope of their application. Still, the concerns rise over safety issues, as HV on board may be hazardous for the vehicle and dangerous for passengers. This is especially a problem in case of traffic accidents, when HV wires may be hanging unprotected [34].

V. MARKET PERSPECTIVES

Plug-in electric vehicles are young at the market, coming in a new wave in automotive industry, with a goal to establish electrical transportation in a domain of passenger vehicles. A number of leading car manufacturers is turning their product lines in this direction. They are driven by imposed restrictions to emission levels of exhausting gasses introduced in USA legislation, new market regulation in some states of USA, huge subsidies by European and far-east Asian governments and increased awareness of population for environmental protection and climate change fighting. These products are entering to a huge, lucrative marked, which has been growing for more than hundred years. Therefore, the challenge for companies is to find a sustainable and economically competitive solution. The PEVs offer such a feature. Their high efficiency brings the energy costs to very low level. A comparison on energy costs per mile of two ICE vehicles (gasoline costs), a HEV and two EVs is presented in Fig. 11. It can be seen that ICE gasoline costs are almost five times higher than the EV´s and twice the HEV´s, which is very attractive feature of the PEVs [35].

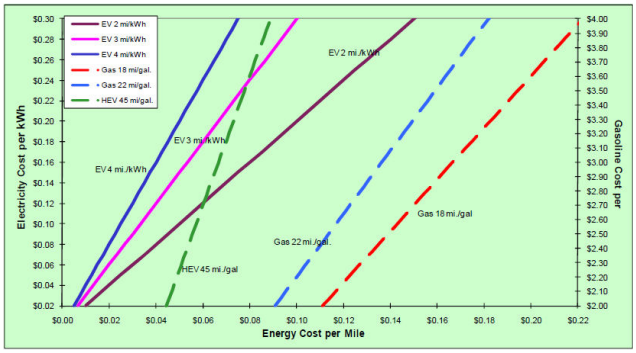


Figure 11. Comparison of energy costs per mile [35]

However, limited range and poor energy supply infrastructure is main obstacle for the EVs. In that sense, for current conditions on the road and for average driving daily schedule the PHEVs are more attractive, although it is not zero emission solution. Figs. 2 and 3 are actually showing these perspectives. Similar conclusion can be obtained from a Deutsche Bank study report, which shows that PEVs´ penetration in 2015 and 2020 will be 11% (2015) and 23% (2020) [36]. These results are presented in Fig. 12. It can be seen that EVs share in 2015 will be only 1% of total sale or 10% of PEVs. In 2020 it is predicted that EVs´ share will be increase to 4% (or 17% of PEVs).

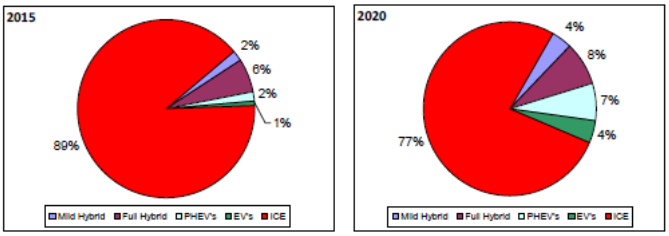


Figure 12. U.S. PEVs penetration by type (2015 and 2020) [36].

Current overview of data for PEVs cumulative sale in the World, given in Fig. 13 shows that the most of PEVs (43%) are sold in U.S. [37]. Next are far-eastern markets of Japan and China (29% in total) and then EU sales with 25% of share.

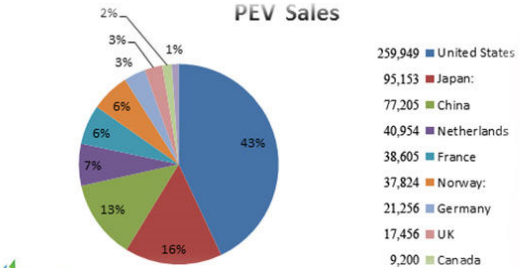


Figure 13. PEVs sales share by countries [37.]

Data for U.S. market are available for each month and by vehicle models. Fig. 14 shows an extract for the period from December 2010 up to August 2014. It can be seen that the sale is constantly increasing and that the most popular models are Chevrolet Volt, Nissan Leaf and Toyota Prius PHEV.

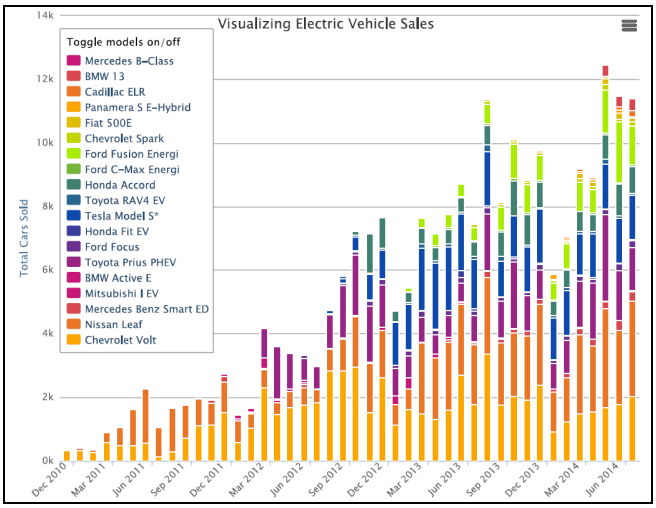


Figure 14. EV sales in US (2010-2014) by models – updated Aug.2014.
(Source: US Dept. of Energy, http://energy.gov/articles/visualizing-electric-vehicle-sales)

VI. CONCLUSIONS

Plug-in vehicles are gaining momentum and more and more models are offered to the market. For the present state of technology, especially battery characteristics, the ICE and Electric motor hybrid solutions are more popular. All forecasts for future market sales predict that PHEVs will be dominant. However, improvement in technology and development of electricity charging infrastructure in cities and on open road ion future may bring full electric vehicles in buyers' attention. The main driving force will be their low energy consumption (high efficiency) and zero emission of gasses.

ACKNOWLEDGMENT

This research was partially co-funded by the by the Provincial Secretariat for Science and Technological Development of AP Vojvodina under contract No. 114-451-3508/2013-04 „Research and Development of Energy Efficient Power Supply and Propulsion Systems of Electric Vehicles".

REFERENCES

[1] N. Hill, C. Brannigan, R. Smokers, A. Schroten, H. van Essen, I. Skinner, "Developing a better understanding of the secondary impacts and key sensitivities for the decarbonisation of the EU´s transport sector by 2050", Final project report produced as part of a contract between EU Commission Directorate – General Climate Action and AET Technology plc., 2012, www.eutransportghg2050.eu

[2] C. Hanke, M. Hülsmann, D. Fornahl, "Socio-Economic Aspects of Electric Vehicles: A Literature Review", In: M. Hülsmann and D. Fornahl (Eds.), "Evolutionary Paths Towards the Mobility Patterns of the Future", Springer-Verlag, Berlin-Heidelberg, 2014.

[3] ***, "EU Energy, Transport and GHG Emissions Trends to 2050 - Reference scenario 2013", European Commission, Luxembourg, 2014, Available On-line: http://ec.europa.eu/energy/observatory/trends_2030/doc/trends_to_2050_update_2013.pdf

[4] ***, "Technology Roadmap – Electric and Plug-in Electric Vehicles", Paris, International Energy Agency, June 2011, Available On-line: http://www.iea.org/publications/freepublications/publication/EV_PHEV_Roadmap.pdf

[5] M. Grahn, E. Klampfl, M. Whalen, T.J. Wallington, "Sustainable Mobility: Using a Global Energy Model to Inform Vehicle Technology Choices in a Decarbonized Economy", Sustainability, 2013, No.5, pp.1845-1862, DOI: 10.3390/su5051845

[6] ****, "Overview of Early Electric Cars (1895-1925)", Low-Tech Magazine, Available On-line: http://www.lowtechmagazine.com/overview-of-early-electric-cars.html

[7] V.A. Katić, B. Dumnić, Z. Čorba, D. Milićević, "Electrification of the Vehicle Propulsion System - An Overview", Facta Universitatis, Series: Electronics and Energetics, Vol.27, No.2, June 2014, pp. 299-316, DOI: 10.2298/FUEE1402299K

[8] Edison Electric Institute, "Utility Perspective on Plug-In Electric Vehicles", Dec.2013, Available On-line: http://www.ncsl.org/documents/energy/KSchefter1213.pdf

[9] M. Eshani, Y. Gao, A. Emadi, "Modern electric, hybrid electric and fuel cell vehicles-Fundamentals, Theory and Design", 2nd Edition, CRC Press, Taylor & Francis Group, Boca Raton (USA), 2010.

[10] C. C. Chan, A. Bouscayrol, and K. Chen, "Electric, hybrid, and fuel-cell vehicles: architectures and modeling," IEEE Transactions on Vehicular Technology, vol. 59, no. 2, Article ID5276874, pp. 589–598, 2010.

[11] A. Ayob, W.M.F.W. Mahmood, A. Mohamed, M.Z.C. Wanik, M.F.M. Siam, S. Sulaiman, A.H. Azit, M.A.M. Ali, "Review on Electric Vehicle, Battery Charger, Charging Station and Standards", Research Jour. of Applied Scien., Eng. and Tech., Vol.7, No.2, 2014, pp.364-373.

[12] I. Aharon, A. Kuperman, "Topological Overview of Powertrains for Battery-Powered Vehicles With Range Extenders", IEEE Trans. on Power Electronics, Vol.26, No.3, May 2011, pp.868-876.

[13] K. Rajashekara, "Present Status and Future Trends in Electric Vehicle Propulsion Technologies", IEEE Journal of Emerging and Selected Topics in Power Electronics, Vol.1, No.1, March 2013, pp.3-10.

[14] M. Yilmaz, P. Krain, "Review of Battery Charger Topologies, Charging Power levels, and Infrastructure for Plug-In Electric and Hybrid Vehicles", IEEE Trans. on Power Electronics, Vol.28, No.5., May 2013, pp. 2151-2169.

[15] C.C. Chan, K.T. Chau, "An overview of power electronics in electric vehicles", IEEE Trans. on Industrial Electronics, Vol.44, No.1, Feb.1997, pp.3-13.

[16] A. Emadi, S.S. Williamson, A. Khaligh, "Power Electronics Intensive Solutions for Advanced Electric, Hybrid Electric, and Fuel Cell Vehicular Power Systems", IEEE Trans. on Power Electronics, Vol.21, No.3, May 2006, pp.567-577.

[17] A. Emadi, Y.J. Lee, K. Rajashekara, "Power Electronics and Motor Drives in Electric, Hybrid Electric, and Plug-In Hybrid Electric Vehicles", IEEE Trans. on Industrial Electronics, Vol.55, No.6, June 2008, pp.2237-2245.

[18] SAE International, "SAE Electric Vehicle and Plug in Hybrid Electric Vehicle Conductive Charge Coupler", J1772-201210, 2012.

[19] IEC, "Electric Vehicle Conductive Charging System - Part 1: General Requirements", 2.0 Edn., IEC 61851-1, Geneva, 2010.

[20] R. Garcia-Valle and J.A. Pecas Lopes (Eds.), "Electric Vehicle Integration into Modern Power Networks", Springer Science + Business Media, New York, 2013

[21] K.T. Chau, "An overview of energy sources for electric vehicles", Energy Conversion and Management, Vol.40, Issue 10, July 1999, pp. 1021–1039.

[22] G.A. Covic, J.T. Boys, "Modern Trends in Inductive Power Transfer for Transportation Applications", IEEE Journal of Emerging and Selected Topics in Power Electronics, Vol.1, No.1, March 2013, pp.28-41.

[23] K. Kiyota, H. Sugimoto, A. Chiba, "Comparing Electric Motors – An analysis using four standard driving schedules", IEEE Industry Application Magazine, Vol.20, No.4, July/Aug.2014, pp.12-20.

[24] http://www.iwe.kit.edu/

[25] IEC 62196-1 standard: "Plugs, socket-outlets, vehicle couplers and vehicle inlets–Conductive charging of electric vehicles", Geneva, 2003.

[26] W. Su, H. Rahimi-Eichi, W. Zeng, M.Y. Chow, "A Survey on the Electrification of Transportation in a Smart Grid Environment", IEEE Trans. on Industrial Informatics, Vol.8, No.1, Feb.2012, pp.1-10.

[27] S. Haghbin, S. Lundmark, M. Alaküla, O. Carlson, "Grid-Connected Integrated Battery Chargers in Vehicle Applications: Review and New Solution", IEEE Trans. on Industrial Electronics, Vol. 60, No. 2, Feb. 2013, pp.459-473.

[28] I. Subotic, E. Levi, "An Integrated Battery Charger for EVs Based on a Symmetrical Six-Phase Machine", IEEE Intern. Symposium on Industrial Electronics, ISIE 2014, Istanbul, Turkey, 1-4 June 2014.

[29] A. Emadi (Ed.), "Handbook of Automotive Power Electronics and Motor Drives", CRC Press, Boca Raton (USA), 2005

[30] B. Bilgin, A. Emadi, "Electric Motors in Electrified Transportation", IEEE Power Electronics Magazine, Vol.1, No.2, June 2014, pp.10-17.

[31] S. Soylu (Ed.), "Electric Vehicles - Modelling and Simulations", InTech, Rijeka, 2011.

[32] Q. Huang, J. Li, Y. Chen "Control of Electric Vehicle", Chapter in a book S. Soylu (Ed.) "Urban Transport and Hybrid Vehicles", InTech, Rijeka, 2010, http://cdn.intechweb.org/pdfs/12061.pdf

[33] H. van Hoek, M. Boesing, D. van Treek, T. Schoenen, R.W. De Doncker, "Power Electronic Architectures for Electric Vehicles", Int. Conf. on Emobility - Electrical Power Train, Leipzig, 8-9 Nov. 2010, pp.1-6.

[34] H. Uwai, A. Isoda, H. Ichikawa, N. Takahashi, "Development of Body Structure For Crash Safety of The Newly Developed Electric Vehicle", 22nd Enhanced Safety of Vehicles Conf., Washington DC (USA), June 13-16, 2011, http://www-esv.nhtsa.dot.gov/Proceedings/22/files/22ESV-000199.pdf

[35] US Department of Energy, "Comparing Energy Costs per Mile for Electric and Gasoline-Fueled Vehicles", Available On-line: http://avt.inel.gov/pdf/fsev/costs.pdf

[36] R. Lache, D. Galves, P. Nolan, "Electric Cars: Plug in 2", FITT Research, Deutsche Bank, 2009, Available On-line: http://www.libralato.co.uk/docs/Electric%20Cars%20Plugged%20In%202%20Deutsche%20Bank%202009.pdf

[37] ***, "Tesla Motors - Summative And Algorithmic Evaluation", Nov. 2014, Available On-line: http://seekingalpha.com/article/2629525-tesla-motors-summative-and-algorithmic-evaluation

Toward Intelligent Driver-Assist Technologies and Piloted Driving: Overview, Motivation and Challenges

Plenary paper

Milos Pilipovic
Faculty of Technical Sciences, University of Novi Sad
Department of Computer Engineering and Communications
Novi Sad, Serbia
milos.pilipovic@rt-rk.com

Danijel Spasojevic, Ivan Velikic, Nikola Teslic
RT-RK Institute
Narodnog fronta 23a, 21000 Novi Sad, Serbia
danijel.spasojevic, ivan.velikic, nikola.teslic @rt-rk.com

Abstract— **In this paper we present an extensive literature review on the research progresses in Intelligent Driver-Assistance Systems (IDAS). IDAS monitor the car's environment and driving behaviour to identify and avoid a potentially dangerous situations at an early stage without human input. Based on intelligent sensor fusion technology with full or partial context-aware autonomy in decision-making IDAS aim to combat obstacles in a traffic scene using various advanced control systems such as Adaptive Cruise Control, Collision Avoidance System, Driver Drowsiness Detection System, Parking Assistance System, Lane Departure Warning System. Several adaptive safety control systems have been proposed and discussed as well as their interoperability issues to address different aspects in road situation analysis. Any autonomous advanced control system manufacturers launches should be selectable by the driver. In addition, the implementation complexity is analyzed. At the end of the paper we envision some research directions. To the best of our knowledge series-built vehicles with a piloted driving function will be technically feasible this decade over the next two to three years period.**

Keywords-autonomous; driver-assist; piloted driving

I. INTRODUCTION

The automotive industry has been moving toward more autonomous vehicles (AVs) over the past few years. Autonomous technology is a technology which is installed on a motor vehicle and which has the capability to drive the motor vehicle without the active intervention or monitoring of a human operator. The term does not include an active safety system or a system for driver assistance IDAS, including, without limitation, a system to provide Adaptive Cruise Control (ACC), Collision Avoidance (CA), Electronic Blind Spot Detection (EBSD), Emergency Braking (EB), Lane Keeping Assistance (LKA), Lane Departure Warning (LDW), Parking Assistance (PA), Traffic Jam Assistance (TJA), or other similar systems that enhance safety or provide driver assistance, unless any such system, alone or in combination with any other system, enables the vehicle on which the system is installed (also called autonomous, self-driving, driverless or robotic) to be driven without the active control or monitoring of a human operator over roads that have not been adapted for its use [1][2]. Many different competing teams by major commercial manufacturers or universities are engaged in research in this area, serving as a backbone for improving the AV technology further. With each subsequent year, the winner of the race raises the bar toward truly AVs one step higher. Companies that currently developing self-driving cars include Audi, BMW, Ford, Google, General Motors, Volkswagen and Volvo. Those AVs are still in development stage but the technology is starting to show up in consumer market as well as receiving widespread media attention. As of beginning of 2013, we have AVs that have driven autonomously for hundreds of thousands of miles. By April 2014, Google's AVs collectively have logged over 700.000 miles on city's streets and highways without a single accident attributable to the automation [3].

A. *Safety*

There are numerous reasons for bringing AV technology into people's everyday-life. The first and most important reason is safety. Proponents of electronic driving aids in AVs claim they would eliminate accidents caused by driver error, which is the primary cause of almost all traffic accidents. According to the World Health Organization (WHO), worldwide there are 1.3 million road traffic fatalities and over 20 million serious injuries per year [4]. Based on a literature survey, Smiley and Brookhuis (1987) concluded that more than 90 percent of all traffic accidents are caused by driver failure via inattention, fatigue or drowsiness. As an example, a retrospective research recently showed that applying brakes half a second earlier in a car traveling at 50 km/h can reduce the crash energy by 50 percent, but also showed that 40 percent of drivers didn't activate or apply effectively brakes before the collision. Research clearly shows that conventional fully human-driven vehicles pose a considerable safety risk. According to researchers, with use of IDAS, 50 percent of serious road traffic accidents could be avoided [5].

B. *Economic and Social Impacts*

The introduction of AVs could significantly ameliorate economic losses and negative social consequences from traffic accidents, as well as economic losses and environmental damage from traffic congestion. AVs certainly hold the potential to disrupt a number of Industries with high stakes in this rapidly evolving market: Automotive, Insurance, and Healthcare to name a few. Simultaneously, an opportunity arises for innovative firms to introduce a scale of creative new business models. Companies engaged in automotive software development for in-car applications will boom. The semiconductor market for IDAS will grow significantly in the foreseeable future, according to IMS Research's automotive semiconductor study. For instance, the microcontroller (MCU)/DSP market in IDAS applications was $70 million in 2011 and is estimated to rise to $150 million by 2018. To achieve economies of scale, IDAS must be targeted also at mid and low-end vehicles segments. IDAS would allow drivers and passengers now with new time on their hands to do other things while traveling such as working, relaxing or sleeping. Speed limits are likely to be raised significantly. Furthermore, use of IDAS could dramatically improve traffic flow, eliminate or reduce traffic collisions, cut insurance costs, while fuel economy from congestion avoidance and more efficient long-distance cruising are big potential bonuses. According to a report from investment bank Morgan Stanley, AVs can contribute $1.3 trillion in annual savings to the U.S. economy alone, with global savings estimated at over $5.6 trillion.

From a marketing point of view, driver comfort also poses a strong impetus for the development of IDAS. Commercial car manufacturers invest substantial effort in the development and improvement of comfort enhancing electronic features such as navigation, ACC and PA systems. Studies show that drivers identify with the contributions of immersive experience and enjoyment from increasing interaction while using these new technologies.

This paper is organized as follows. Below, in section II, we will give a description of the IDAS and the functional blocks that comprise it. Section III concludes the paper and presents future development directions.

II. INTELLIGENT DRIVER-ASSISTANCE SYSTEMS (IDAS)

IDAS are considered a key technology when it comes to moving toward accident-free driving. IDAS monitor the car's environment and driving behaviour to identify and avoid a potentially dangerous situations at an early stage without human input. Based on intelligent sensor fusion technology with full or partial context-aware autonomy in decision-making IDAS aim to combat obstacles in a traffic scene in a non-intrusive and naturalistic manner using various advanced control systems. IDAS features are either built into cars or available as add-on packages. Bolstered by increasingly powerful and affordable sensors, vision/camera systems, car data networks, sophisticated algorithms, these invisible multi-tasking co-pilots will redefine the nature of driving. Next-generation IDAS will increasingly utilize wireless network connectivity improving Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication. Before introducing any IDAS feature, every component of the same must be certified and the consequences of such system use should be identified.

A. *Current Trends in IDAS Design*

In order to process multiple algorithms in parallel and develop a scalable architectures, IDAS designers are increasingly turning to FPGAs and PSoCs (Programmable SoC) [6]. FPGAs present a compelling viable alternative to IDAS design compared to fixed-function devices. FPGAs provide IDAS designers capability to quickly customize the I/O structures, hardware and data pipeline to be optimized for a specific algorithm as well as faster time to market, lower risk and cost of ownership than standard ASIC (Application-Specific Integrated Circuit) and ASSP (Application Specific Standard Product) approaches. The industry lacks interoperability specifications for laser, radar and video data in the car network. Also lacks standards for embedded vision-processing algorithms. For instance, commercial manufacturers use multiple data communication standards for audio-video data, such as MOST (Media Oriented System Transport), Ethernet AVB, and LVDS. IDAS must support a mass of interfaces to ensure adoption across a spectrum of possible interfaces. Most important requirements of IDAS includes:

- Higher levels bandwidth and performance for processing video streams from multiple cameras.
- Real-time processing due to multitude of sensor inputs.
- Transmit, receive, and translate between multiple communication standards.

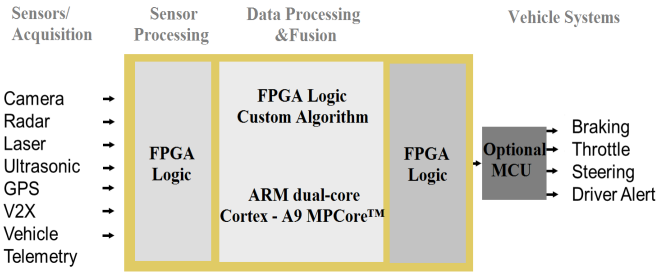


Figure 1. An example of PSoCs based IDAS architecture

B. *Digital Sensors*

Intelligent data fusion technology also known as the car's Artificial Intelligence (AI) use a probabilistic approach aggregating and analyzing many sensor readouts into a combined assumption. IDAS can respond to dangerous situations using actuators roughly 1000 times faster than humans on average, and with potentially far greater precision. By comparison, even the professional drivers can only react within about ½ second to new stimuli, while average human reaction time is over second under most driving circumstances [7]. State-of-the-art AVs use a variety of sensors such as:

1. Multi Purpose Camera – Offering megapixels resolution, monitor driver eye/face inside of car,

lane markings by spotting the contrast between the road surface and the boundary lines.

2. Stereo Vision Camera – Offer three dimensional (3D) environment detection, spotting potential hazards like pedestrians and predicting where they are headed.
3. Infrared Camera – For instance, Mercedes' Night View Assist system uses two headlamps to beam invisible, infrared light onto the road ahead. A computer converts the detected IR signature information from the windshield-mounted camera and shows the illuminated image on the driver's display.
4. Radar - High-end cars are already equipped with radars, which can detect and track nearby objects. For instance, Audi's ACC with stop & go function, uses radar system which monitors the area in front of the car in a 35-degree field of view and at a distance of up to 250 meters. Two radar sensors in the rear monitor events behind the car at a distance of up to 70 meters.
5. LIDAR – Google's AVs uses rooftop mounted Light Detection and Ranging system, which uses 64 lasers, spinning at upwards of 900 rpm, to generate 360° real-time map of the vehicle's surroundings.
6. GPS - Provide vehicle's geo-location information and navigation capabilities.
7. Wheel Encoder – Google's AVs uses wheel-mounted sensors which can measure the velocity of the car as it maneuvers through traffic.

Below is an example of how Google's cars work fusing multiple sensor inputs [8]**:**

- The driver sets a destination. The car's software calculates a route and starts the car on its way.
- A rotating LIDAR sensor monitors a 60-meter range around the car and creates a dynamic 3D map of the car's surroundings.
- A sensor on the left rear wheel monitors sideways movement to detect the car's position relative to the 3D map.
- Radar systems in the front and rear bumpers calculate distances to obstacles.
- Artificial Intelligence (AI) software in the car is connected to all the sensors and has input from Google Street View and video cameras inside the car.
- The AI simulates human perceptual and decision-making processes and controls actions in driver-control systems such as steering and brakes.
- The car's software consults Google Maps for advance notice of things like landmarks and traffic signs and lights.
- An override function is available to allow a human to take control of the vehicle.

Due to the diverse characteristics of driving sensors may fail. IDAS must function under a variety of weather and lighting conditions. For instance, if the vision based system detects that the camera lens is occluded it can warn the driver that it is currently in non-operational mode. Fusing multiple sensor inputs can often provide a more effective solution than using a single in isolation. It will be critical for IDAS to have reliable and cost-effective internal feedback algorithms that can detect electrical or physical damages of these sensing components. A sensor that fails to provide new updates is easily detected, but a sensor that occasionally sends bogus data may be more difficult to detect.

C. IDAS Technologies – Several Examples

1) Adaptive Cruise Control (ACC)

Today's ACC relies on radar or laser technology. The system regulates the vehicle's speed and distance to the vehicle ahead [9][10]. ACC enables car's speed synchronization, adjust smooth lane change, reduces the number of sudden accelerations and decelerations and reduce accident possibility [11]. For instance, Audi's ACC with stop & go function uses one or two radar sensors at the front of the vehicle to detect the reflections of objects as far as 250 meters away from the car. The driver can adjust the distance to the car ahead and the control dynamics in multiple levels. After stopping at a traffic light, it automatically drives off and follows the vehicle ahead and conversely after a longer stop, the driver must tap the accelerator pedal or activate the control stalk to resume driving. ACC with stop & go function interacts with other IDAS to continuously analyze all of the car's surroundings. The system uses knowledge base to recognize complex scenarios and offer early stage support to the driver. Because it also works together with the optional Multi Media Interface (MMI) navigation plus, it knows the course of the selected route and can use this information to regulate the car's speed. Car manufacturers typically offer a number of versions of ACC for different model series.

2) Lane Departure Warning (LDW)

LDW alerts the driver when the car begins to unintentionally stray from its lane without obvious input from the driver (for instance, due to driver's distraction or momentary lapse in concentration) [12]. Lane Keeping Support (LKS) help to counteract unintentional drift out of the marked lane by the driver, typically via audio alert or haptic warning. Both systems make use of a video camera to detect lane markings ahead of the vehicle by analyzing differences in contrast between the road surface and the lane markings. The car's position in the lane is then compared with additional information taken from the steering angle, brake and accelerator position sensors and whether or not the indicators are in use [12][13]. LDW and LKS rely on visible lane

markings, hence faded, missing, incorrect, markings covered in snow or ice can hinder the ability of those systems.

3) Parking Assistance (PA)

The system finds parking spaces using ultrasound sensors and/or video cameras that scan the side of the road in two dimensions while driving at moderate speed. When system detects a space of sufficient size, it executes a near-perfect reversing maneuver while the driver operates the accelerator or brake pedal. The PA will maneuver forward and back multiple times and also help driver when leaving parallel parking spots. Visual (via the central display) and acoustic signals that feature in the PA warn the driver of any obstacles.

4) Collision Avoidance (CA)

The system typically uses radar, laser and/or video sensors to detect an imminent crash. CA allows the driver to take appropriate corrective actions in order to mitigate, or completely avoid the collision event, or furthermore, takes action autonomously without any driver input by braking, steering or both [14]. CA detects proximity of other vehicles or pedestrians in surroundings, how much vehicle's speed needs to be reduced and how close the vehicle is going off the road. Some of the precaution measures include: tension of the seat belts, moving optional memory seats forward to protect the car's occupants, closing of the car's sunroof, closing of any open windows, partial or full braking, pedestrian detection, warning occupants through a buzzer.

5) Driver Drowsiness Detection

The system prevents accidents when the driver is getting drowsy. Driver fatigue ("falling asleep at the wheel") is a major cause of road accidents, accounting for up to 50% of serious accidents on motorways and monotonous roads. The Driver Warning System (DWS) uses different acoustic, visual and haptic signals to alert the driver against his drowsiness which prevents from accidents [15].

6) Night Vision Asisstant (NVA)

NVA increases a driver's visual perception in darkness or poor weather conditions beyond the reach of the vehicle's headlights. For example, Audi's NVA uses thermal imaging camera with a 24-degree angle of aperture at the front of the automobile. It works in the Far Infrared Range (FIR) which can look up to 300 meters ahead, and reacts to the objects radiated heat. A computer converts the information from the camera into black and white images and shows them on the driver information system's central display, if desired. If the system predicts a potentially dangerous situation because a person or animal is crossing the road in front of the car, for example, the person is marked in red and a warning tone sounds. The brake system is conditioned at the same time [16].

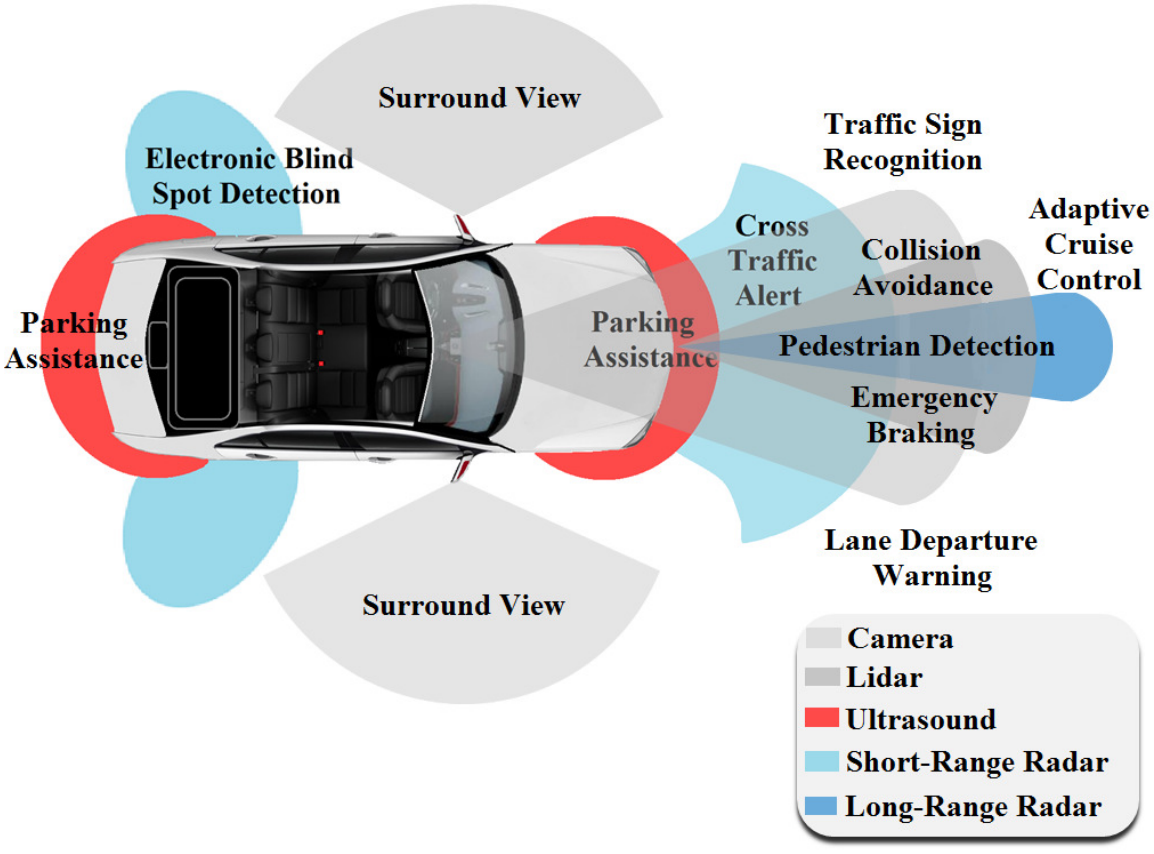


Figure 2. IDAS & sensors

III. CONCLUSION

IDAS is one of the fastest growing segments in automotive industry. In IDAS, sensors and sophisticated algorithms are combined to identify the vehicle environment so that the driver can receive assistance or be warned of potentially dangerous situations at an early stage. The research is based on numerous interviews with commercial manufacturers and a through literature review. After conducting a survey we determined that the benefits of the technology outweigh the disadvantages so IDAS should be encouraged when it is superior to an average human driver.

The social and economic implications of the IDAS are enormous. Increased safety awareness as well as government's legislation presents a driving force contributing to the IDAS market growth. IDAS features have already become established in luxury, mid-range and compact cars. Automotive OEMs need to deploy advanced IDAS technologies in their next-generation models in order to achieve five-star safety ratings as mandated by car's safety organizations such as New Car Assessment Program (NCAP). Since these systems are very similar with the systems that are used in AVs prototypes, they are regarded as the transition elements on the way to the implementation of fully AVs.

IDAS are currently used to increase safety in speed zones where driver error is most common: at lower speeds, traffic congestion, and at higher speeds as cruising on a highway for instance. There are some improvements required to be done in the future. It is expected that next-generation IDAS will likely offer greater vehicle autonomy at lower speeds and may reduce the incidence of low-impact crashes. Recent reports show that Audi's new autonomous driving system, dubbed "Piloted Driving", will debut on its luxury A8 sedan around the year 2016. The Mercedes system, dubbed "Steer Assist," is able to autonomously drive in traffic and on highways at speeds of up to 124 mph. Automotive supplier Bosch has also announced that at least one of its customers will be using its autonomous driving technology in 2014. Bosch's system is designed for heavy traffic and works at speeds of up to 30 mph.

"AVs are inevitable. It's only a matter of time," says Andrew Chatham, senior staff engineer and off-board software lead of Google's self-driving car program.

ACKNOWLEDGMENT

This work was partially supported by the Ministry of Education and Science of the Republic of Serbia, under the project No. 32034, year 2014.

REFERENCES

[1] S. K. Gehrig and F. J. Stein, "Dead reckoning and cartography using stereo vision for an autonomous car," in International Conference on Intelligent Robots and Systems, Kyongju, Korea, pp. 1507-1512, 1999.

[2] "NRS: Chapter 482A-Autonomus Vehicles," Nevada Legislature. N.p., 2012. Web.11 <http://www.leg.state.nv.us/NRS/NRS-482A.html>.

[3] U. Chris, "The Latest Chapter for the Self-Driving Car: Mastering City Street Driving," Retrieved April 28, 2014, from Google's Official Blog Web site: <http://googleblog.blogspot.com/2014/04/the-latest-chapter-for-self-driving-car.html>.

[4] World Health Organization, Global Status Report on Road Safety 2013: Supporting a Decade of Action, Luxembourg: World Health Organization, <http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/index.html>.

[5] F. Guido, "Driver Assistance Systems (DAS) Intelligent Accident Prevention in Road Traffic," ICT and Road Safety, WISS Forum, Geneva, 2011.

[6] M.R Nithin, B. Raisa and M. Sreela, "Advanced Driver Assistance System using FPGA," White Paper, March 2014.

[7] Human Factors Group, "Reaction Time of Drivers to Road Stimuli," Monash University, June 1982, <http://www.monash.edu.au/miri/research/reports/other/hfr12.pdf>

[8] G. Erico, "How Google's Self-Driving Car Works," Retrieved September 10, 2014, from IEEE Spectrum Magazine: <http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works>

[9] Hella KGaA, Hueck & Co, "Electronics – Driver Assistance Systems," Technical Information, 2005.

[10] P. Huei, "Evaluation of Driver Assistance Systems - A Human Centered Approach," supported by the U.S. Army TARDEC, NSF and the TRW Automotive, 2006.

[11] W.S. Lee, D.H. Sung, J.Y. Lee, Y.S. Kim and J.H. Cho,"Driving Simulation for Evaluation of Driver Assistance Systems and Driving Management Systems," sponsored by the Korea Transportation Institute under the national project, 'Development of National Traffic Core Technology,' 2007.

[12] K.Henricksen and J.Indulska, "Software Engineering Framework for Context-Aware Pervasive Computing," 2nd IEEE Conference on Pervasive Computing and Communications (PerCom), 2004.

[13] A. Iihoshi, "Driver Assistance System (Lane Keep Assist System)," Presentation to WP-29 ITS Round Table Geneva, 2004.

[14] S. Peter, B. Song and J.K. Hedrick, "Development of a Collision Avoidance System," Society of Automotive Engineers, 1998.

[15] E.Bekiaris, S.Nikolaou and A.Mousadakou, "System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation," National Center for Research and Technology, Hellas (CERTH) AWAKE Consortium, August 2004.

[16] Audi AG, "Next Generation," Audi AG at the 2014 International CES, Technical Information, January 2014.

Session T-01
MATERIALS AND COMPONENTS

Optical Chemical Sensors for Industrial Applications

Keynote Paper

Zoran Jakšić
Center of Microelectronic Technologies
Institute of Chemistry, Technology and Metallurgy, University of Belgrade
Belgrade, Serbia
jaksa@nanosys.ihtm.bg.ac.rs

Abstract— **The market for accurate and inexpensive sensors of chemical compounds in gases and liquids is increasing at a steady pace and it is expected that in 2015 the global demands should reach a level of more than US$ 17 billion. Industrial applications of optical chemical sensors include process and quality control, as well as environmental pollution measurements in different fields that include process industry, automotive, biomedical industry, power engineering, oil and gas industry and food industry. Optical sensing methods offer some distinctive advantage over other approaches, including the applicability in dangerous, flammable and explosive environments, high sensitivities, compact design and insensitivity to external interference. They can be used to recognize trace amounts of a wide range of different chemicals, are compatible with the existing control systems and can be used for multianalyte detection. This work first shortly reviews different schemes utilized for optical chemical sensing in gases and liquids. These include various planar and fiber-optic waveguide systems that utilize absorption, transmission or reflection spectroscopy, fluorescence measurements that may be applied to propagating or evanescent waves. Affinity sensing is considered, where the optical sensor surface or a special intermediate receptor layer binds the targeted analyte, which in turn causes physicochemical transformations that modulate the optical properties of the system. The most part of the presentation is dedicated to sensors utilizing surface plasmon polaritons in metal-dielectric nanocomposites. These devices offer label-free, real-time and multianalyte operation with sensitivities even reaching single molecule level. Different schemes are considered including the use of optical metamaterial structures for adsorption-based sensing. The advantages and disadvantages of these sensing schemes are considered.**

Keywords-chemical sensors; adsorption; plasmonic sensing; nanophotonics; metamaterial sensors

I. INTRODUCTION

The continual development of industry and the increasing sophistication of the applied methods and protocols pose a general requirement for more complex sensing and control of different processes. The measurement results must be accurate, there should not be a lag between the appearance of a signal and its reading which calls for sensor speed and its online availability, the sensing devices must often be applied under difficult conditions that include aggressive media, high or low temperatures, vibrations, electric or magnetic disturbances, etc. In addition to that, a sensor must not perturb the industrial process itself – for instance, electronic devices may ignite flammable or explosive media or cause electromagnetic interference.

All-optical sensors meet many of these conditions, since they do not convey electric signals at all and are thus neither sensitive to external interference, nor they cause any; they are often made of robust materials and are very fast. Because of that they are applicable in dangerous, flammable and explosive environments, high sensitivities, compact design. Their design may be very compact, especially if nanophotonic devices with evanescent waves are used. Their sensitivities are extremely high, ensuring even the detection of single molecules [1]. They can be used to recognize trace amounts of a wide range of different chemicals, are compatible with the existing control systems and can be used for multianalyte detection.

Obviously, the qualitative and quantitative assessment of chemical substances is not limited to process and quality control in processing industry and chemical plants. Chemical sensing finds intensive use in automotive, biomedical industry, power engineering, oil and gas industry and food industry, among other fields. For instance, a thermal power station will need to control its greenhouse gases emission, as well as the levels of different other pollutants. Generally, one will need to observe the levels of different industrial by-products in the environment and monitor the possible appearance of dangerous substance to prevent, avoid or mitigate chemical accidents.

It is expected that at the present pace the global sensor market will reach US$ 154.4 billion in 2020 [2]. The second fastest growing segment in it are chemical and biological sensors, which should reach US$ 32.8 billion in 2020, while in 2015 the global demands should exceed a level of US$ 17 billion. The photonic sensors should reach a level of about US$ 8 billion in 2015.

A large body of literature is dedicated to chemical sensors that base their operation on optical mechanisms. Numerous research articles are being published in this field (e.g. [1, 3-11]. A number of review paper has been published dedicated to

The paper is a part of the research funded by the Serbian Ministry of Education, Science and Technological Development within the project TR32008.

optical chemical sensors [12-19], as well as a number of books [20-32]. These are written from different viewpoints and illuminate different facets of optical chemical – but also biological – sensor technologies and applications.

Contrary to the majority of the previous publications, this paper gives a broad overview of different optical sensors of chemical or biological analytes through an attempt to give a systematic "bigger" picture. The accent is given to nanophotonic affinity-based sensors. A possible classification of various methods an mechanisms is proposed. The goal of such a concept is to create a broad framework where not only all of the existing devices can fit, but it may also point out to novel structures and methods.

II. General Considerations

A chemical sensor can be described as an analytical device that converts the chemical information about a chemical substance (analyte) into a readable output signal, usually electric or optical one. The chemical information may range from the simple presence, over the concentration of the analyte to a total analysis of multiple analytes. The output signal may be readable by an instrument or directly by an observer, and is proportional to the concentration of the analyte. The sensor is self-contained, i.e. all its units are located within the same package.

Analyte itself is defined as a chemical substance present as a constituent in a complex sample and is of interest for an analytical procedure. It may denote chemical element, component or ion. From this point of view, a simple litmus paper strip may be defined as a non-reusable pH sensor with optical readout. Analyte may also be a biological substance, since the sensing mechanisms applied are identical in both cases, i.e. we assume there is no fundamental difference between a chemical and a biological sensor.

Some requirements posed to a chemical sensor include:

- Maximum sensitivity (defined as the ratio between the amount of the target analyte in a complex sample and the output signal.)
- Maximum selectivity (the ability of the device to discerns between different analytes, including the similar ones.)
- High signal-to-noise ratio (much larger value of the useful signal than that of the extrinsic and intrinsic noise).
- High dynamic range (large ratio of the highest to lowest measurable concentration)
- Fast response (response time being defined as the time between the onset of a concentration change at the sensor input and the moment the sensor measures a certain percent of the final value. Ideally it is a real-time operation).
- Maximum resolution (the smallest distinguishable concentration variation).

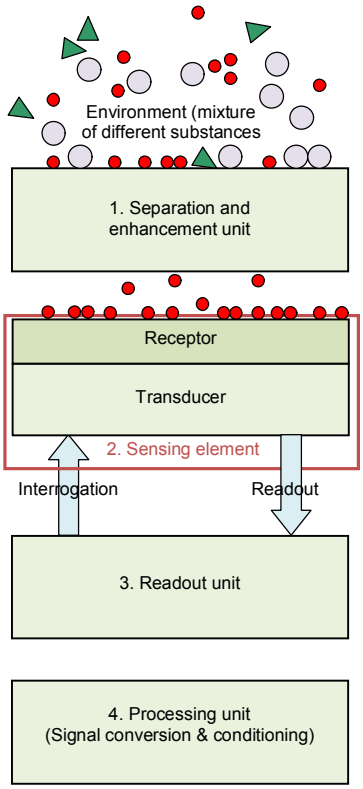


Figure 1. Layout of a general chemical sensor consisting of (1) separation and enhancement unit; (2) sensing element consisting of a receptor and a transducing element, (3) readout unit and (4) processing unit for signal conversion and conditioning. Red circles denote targeted analyte particles

- Compact dimensions.
- Simplicity (as simple design and use as possible; for instance, label-free operation).
- Reusability (no need to discard the sensor after use or to regenerate it).
- Possibility of parallel sensing of several analytes (multianalyte operation).
- Nondestructive operation.
- Possibility of online use.
- Minimal hysteresis (defined as the difference between the sensor readout with measurand increasing and decreasing)
- Minimum drift (signal variation with time for the same value of analyte concentration)
- Maximum operating life (defined either for continuous operation or for the number of repeated measurements)
- Low cost.

A general affinity-based chemical sensing system is schematically presented in Fig. 1. It consists of four main blocks, (1) the unit for separation and possibly reaction

enhancement, (2) the sensing element – the structure where the analyte concentration causes a proportional change of the optical properties of the material, for instance refractive index change, (3) readout unit, the part of the device where the signal is generated and (4) the processing unit, where signals are conditioned and communicated further.

The only parts that each sensor must have are the sensing element and the readout unit, while separation/enhancement unit and processing unit are optional. Also, two or more presented blocks may be integrated into a single unit, so that Fig. 1 represents a functional, rather than a structural division of the device.

According to IUPAC, the sensing element itself consists of a receptor that transforms chemical information into energy measurable by transducer and a transducing unit that converts this energy carrying chemical information into a useful analytical signal [33]. Different operating mechanisms of transduction can be met in chemical sensors, including optical, electrochemical, chemical, mass sensitive (piezoelectric and surface acoustic), magnetic, thermometric [33]. Other mechanisms may include different types of radiation like ionizing (X-ray, beta, gamma), microwave, terahertz, etc. In our case the transducing mechanism must include optical processes.

Further we consider the sensing process itself. The liquid or gaseous environment around the sensor represents a complex mixture of different substances that includes one or more analytes to be detected and quantified. In Fig. 1 these substances are represented by different symbols (triangles, small or larger circles).

In order to improve sensor selectivity, i.e. to avoid false positives, we are interested to separate the target analyte from the mixture of other substances even before it triggers the sensor response. This may be done by different schemes for analyte separation, that include the use of receptors (ligands), the application of filtering structures, etc. The same structure may be utilized to enhance the reaction, thus amplifying the sensor signal. Upon that analyte arrives to the active surface of the sensing element.

In this text we limit ourselves to all-optical sensors of chemical substances. This means that some optical property of the material of the sensing element is modified by the presence of analyte. Such sensors are sometimes denoted as optodes or optrodes. Lübbers and Opitz utilized the term optode in 1975 [34], while the term optrode was first utilized in 1984 [35] to denote a chemically functionalized fiber tip.

The change of the chosen optical property modulated by the presence of analyte is further read out by a propagating or evanescent electromagnetic wave with a frequency in the optical range. This readout may be passive (there is no external interrogating beam, the structure itself emits the output signal) or active (an interrogating optical beam impinges on the sensing element, the presence of analyte changes its properties, and the modulated beam is further used as the output signal.

The processing unit may include light source for interrogating beam (e.g. LED diode or laser), photodetector for the conversion of the optical output signal to electric signal, amplifiers, multiplexers/demultiplexers, etc.

III. Sensor Exposure Modes

Chemical sensors generally may function in four possible modes of exposure to the analyte:

- Remote mode
- Immersion
- Adsorption
- Absorption

Remote mode is used when the sensor is removed far from the analyte-containing medium and the sensor has either to act actively and use electromagnetic beam to probe the medium at a distance or act passive and detect the emissions from the tested medium. An example of the remote mode of operation would be spectral measurement, used even to assess the chemical properties of deep space objects. This mode of operation that generally belongs to remote sensing is sometimes excluded from the definitions of the chemical sensor because of the lack of physical contact with analyte.

Immersion is the situation when the sensor is wholly or partially surrounded by the medium, but no further binding processes occur at its surface or within its bulk. This would be the case when a sensing probe is dipped into fluid and the optical properties of fluid are directly measured. Alternatively, the measured fluid may fill a large cavity within the sensing probe. In both cases the bulk properties of the measured fluid are those that trigger the sensor response.

Adsorption is a process in which an atom or molecule of a liquid or gaseous substance is bound to a solid surface. The adsorbed layer may be extremely thin, even monatomic or monomolecular, and it may not cover the whole surface, and still be sufficient to trigger a response of the chemical sensor. The bound substance is called adsorbate, and the solid surface to which the adsorbate adheres is called adsorbent. Adsorption occurs due to the presence of surface forces. The result of adsorption is an increase of the concentration of adsorbate at the adsorbent surface. A typical example of adsorption is condensation at the surface. Adsorbate may be physisorbed (bound by van der Waals forces or by electrostatic forces) or chemisorbed (bound by covalent bonding). The process opposite to adsorption is desorption – the removal of adsorbate particles from the adsorbent.

If the binding of substance occurs within the volume of a solid object, the effect is called absorption, the substance dissolved within the volume is the absorbate and the volume permeated by the absorbate is the absorbent. Absorption is not a consequence of surface forces. A typical example of absorption is incorporation of hydrogen atoms within palladium crystal lattice.

If the binding of particles from liquid or gas occurs within pores, like in nanoporous materials, the process will occur within the volume of the object and still it will be denoted as adsorption, because binding is a consequence of surface forces.

An example of such volume adsorption is binding of gases in zeolites or in activated carbon.

Adsorption and absorption both belong to the group of processes called sorption. Chemical sensors generally may any or both of the two sorption mechanisms for their function. In the case of optical devices, the increased concentration of the analyte (sorbate) changes the optical properties of sorbent material. These changes are detected by the readout beam.

The adsorption kinetics is usually described by isotherms that define the amount of the adsorbed substance at a constant temperature. For gas adsorption one typically uses the simple Langmuir isotherm that assumes a monolayer adsorption. If multilayers are formed by the adsorption of particles onto already adsorbed particles, one utilizes Brunauer-Emmett-Teller (BET) isotherm. An illustration can be seen in Fig. 3. There are other isotherms that are used to analyze adsorption at the surface of chemical sensors [34, 35], but we will not consider them in this text.

More than one process may occur concurrently within the same structure. For instance, an analyte may adsorb to the surface of a structure and at the same time diffuse and permeate its bulk.

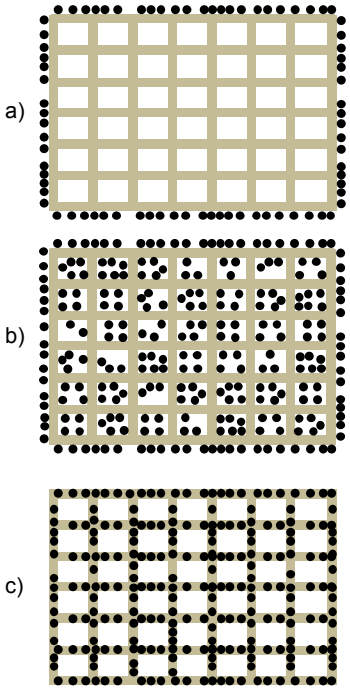


Figure 2. Sensor exposure modes: a) surface adsorption; b) bulk adsorption in nanopores; c) absorption (permeation/dissolution). In all three cases the active part of the sensor is a porous ordered structure.

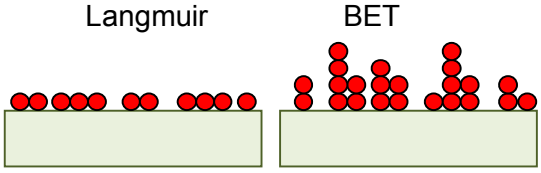


Figure 3. Examples of Langmuir (left) and BET (right) adsorption.

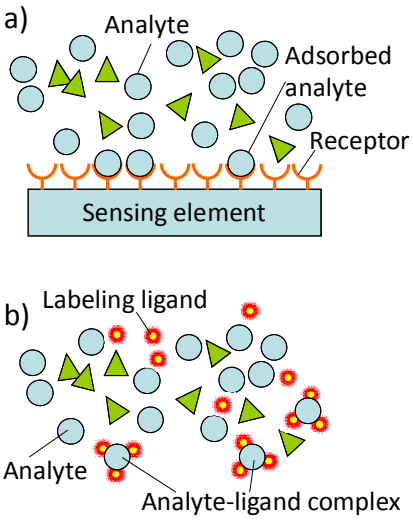


Figure 4. Two mechanisms of operation of affinity-based sensors. a) receptor layer at the sensor surface binds targeted analyte and discards undesired similar species; b) ligand (e.g. fluorescent molecule) labels the analyte particles.

IV. Affinity Sensors

In our further consideration we analyze adsorption-based sensors. The medium around a chemical sensor typically contains several competing species that may simultaneously adsorb on the same substrate. It may happen that two or more of these adsorbates cause the same modulation of sensor parameters and thus an identical output. This makes a great problem with sensor sensitivity. A way to avoid it is to utilize affinity sensors, i.e. the devices that specifically bind to a target analyte.

The term "affinity sensor" was used by Schultz et al in 1970ties [36, 37] to describe a glucose sensing device. In affinity sensors one utilizes material that preferentially binds the targeted analyte. This material is denoted as receptor and can be deposited to the active surface of the sensor to promote target-specific binding. This material is sometimes also denoted as ligand. An illustration of this case can be seen in Fig. 4a.

One also utilizes the term ligand to denote signal-triggering substance (for instance, fluorescent labeling molecule) that binds to the target analyte in order to ensure its detection, Fig. 4b. Thus affinity sensor may be a device utilizing a layer of target-analyte binding receptor material deposited onto its surface, but also a device utilizing ligand molecules for labeling.

V. A Possible Systematization

In the further text we introduce a classification of optical sensors of chemical (or possibly biological) analytes. To this purpose we observe the main properties of interest. We can consider *what* parameter is read out (sensor output), *how* is it modulated, i.e. which optical phenomenon is utilized to generate the output signal and *where* it all occurs, i.e. what type of structure is utilized as the active part of the sensor. These

what, how and where are interdependent and define the behavior of the sensor.

A. What is modulated (macroscopic readout)

The external world can only see changes of the macroscopic properties of a chemical sensor caused by the presence of an analyte. These may be any electric, thermal, electric, chemical, etc. properties observable by our instruments. One could imagine devices that change temperature due to some kind of interaction between the active part and the analyte, modulate electric potential, change optical transmission, even, in some more exotic situations, change shape (like in bimaterial devices). The sensor may be called optical as long as the basic interaction that causes modulation is optical.

In a strict sense we can assume that the macroscopic output is also optical. This is the simplest situation, and one can always assume that a device converting optical interaction to non-optical signal basically consists of two or more units, one generating optical output, the other converting the optical output to non-optical. Thus for the sake of simplicity, in this text we consider only the optical output.

Two most obvious optical parameters to modulate are the scattering parameters of the sensor – its optical transmittance T and optical reflectance R. These are connected with the optical absorptance as $A = 1 - T - R$. Of course, the interaction can also cause optical emission or modify the already existing emission. One can also utilize devices that change the polarization of the utilized optical waves (chirality or handedness of the light). Thus the following macroscopic optical outputs can be utilized for sensor readout:

1. Reflection
2. Transmission
3. Absorption
4. Emission
5. Polarization
6. Phase change

One should mention that the use of the above may not be obvious at a first glance. For instance, in conventional surface plasmon polariton-based devices one varies the measurement spatial angle in order to obtain the sensor. Yet, the underlying principle of signal readout is the determination of reflectance minimum related with plasmon resonance.

The following parameters may be varied when observing the quoted outputs:

1. Intensity (amplitude)
2. Wavelength (spectral dependence)
3. Temporal dependence
4. Spatial distribution.

Different outputs and varying parameters can be combined when obtaining the macroscopic output from the sensor. For instance, when measuring reflection, one can observe its intensity at a given moment and for a given wavelength, or alternatively observe its distribution across wavelengths (reflection spectroscopy), temporal dependence (changes of reflection coefficient over time), spatial dependence (reflection may vary across the analyzed surface). This does not relate with the intrinsic transduction mechanism used to obtain a particular change. Different mechanisms that can be used to modulate macroscopic output are given in the next subsection.

B. How is it modulated (optical transduction phenomena)

The optical phenomena that are being modulated are the heart of an optical chemical sensor. They are the transduction mechanisms itself that define the sensor behavior. These mechanisms further modulate transmission, reflection, absorption or emission of optical radiation.

The IUPAC definition of the main modes of transduction of optical sensors include absorbance, reflectance, fluorescence, refractive index, optothermal effect and light scattering [33]. However, such a definition simultaneously uses several criteria instead of a single one, since for instance absorbance modulation actually represents a consequence of refractive index change (being proportional to the imaginary part of the complex refractive index that describes losses), reflectance changes are the consequence of the changes of the real part of the refractive index, etc. On the other hand, it does not take into account some other phenomena, like e.g. photoacoustic, electrochromic, etc. In this text we opted to utilize instead different microscopic mechanisms that modulate the properties listed as macroscopic outputs in subsection A. Some of them include:

1) Wave superposition

- Interferometry
- Optical resonators
- Diffractive methods

2) Evanescent wave mechanisms

- Total internal reflection
- Surface plasmon polaritons
- Extraordinary optical transmission
- Surface enhanced Raman scattering (SERS)
- Surface enhanced infrared absorption (SEIRA)
- Dyakonov waves

3) Photoluminescence phenomena

- Fluorescence
- Fosforescence

4) Polarimetry/ellipsometry

- Optical rotation (optical activity)
- Optical rotatory dispersion
- Circular dichroism

5) *Scattering*
   - Rayleigh scattering
   - Mie scattering
   - Tyndall scattering
   - Brillouin scattering
   - Raman scattering
   - Fano resonance

6) *Nonlinear optical phenomena*
   - Harmonic generation by frequency mixing (second, third, high harmonic generation)
   - Optical rectification
   - Optical parametric generation, amplification, oscillation
   - Spontaneous parametric down conversion
   - Light-plasma interaction
   - Stimulated Brillouin scattering
   - Cross-polarized wave generation
   - Optical soliton interactions
   - Kerr effect
   - Four-wave mixing

7) *Quantum optical phenomena*
   - Upconversion
   - Electromagnetically induced transparency
   - optically active quantum structures (e.g. quantum dots)

8) *Photothermal effects*

9) *Photoacoustic and acousto-optic effects*

10) *Photoelasticity*

11) *Electro-optical mechanisms*
   - Electroabsorption
   - Stark effect
   - Franz-Keldysh effect
   - Electrochromic effect
   - Pockels effect
   - Electron-refractive effect
   - Electro-gyration

12) *Transformation optics*
   - Superabsorption and "optical black holes"
   - Superlensing
   - Hyperlensing
   - Sensor cloaking

Many of the above described sensing schemes reduce to refractometric ones since they are based on modification of the complex refractive index. Also, two or more may be combined in the same sensor.

C. Where is it modulated (device structures)

We assume that an optical chemical sensor is basically an optical waveguide in a most general sense, i.e. a structure that supports electromagnetic (EM) waves. These waves may be propagating or localized, sinusoidal or exponentially decaying (evanescent); they may be transmitted, reflected, absorbed or emitted. The structure itself may have an arbitrary geometry and its electromagnetic properties can be described by a given spatial distribution of its complex refractive index

$$n(\omega,\vec{r}) = n_{\mathrm{Re}}(\omega,\vec{r}) + i \cdot n_{\mathrm{Im}}(\omega,\vec{r}) \qquad (1)$$

where ω is the angular frequency, r is the spatial position vector, "Re" stands for real and "Im" for imaginary part of the refractive index. Thus in a general case n is frequency dispersive, inhomogeneous (spatially varying), anisotropic (with properties in a given point varying in dependence on the EM wave direction) and lossy.

The complex refractive index is the basic material property in optical systems generally. In

A division of waveguiding structures for optical sensors can be done based on the dimensions of their characteristic features compared to the operating wavelength λ:

1) Superwavelength (characteristic dimensions much larger than λ),
   - Spectroscopic cells

2) Mesoscopic (characteristic dimensions comparable with λ),
   - Planar waveguides
   - Fiber optics
   - Photonic crystals

3) Subwavelength (characteristic dimensions much smaller than λ)
   - Surface plasmon polariton devices
   - Nanoplasmonics
   - Metamaterials

In the further text we will shortly describe the properties of some mesoscopic and subwavelength structures. Our stress will be on the subwavelength structures for chemical/biological sensors.

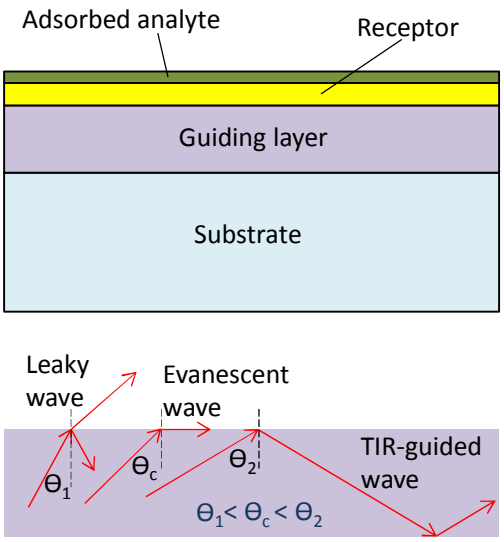


Figure 5. Generic structure of a chemical sensor based on optical waveguide (top) and the scheme of optical wave propagation in the guiding layer (bottom).

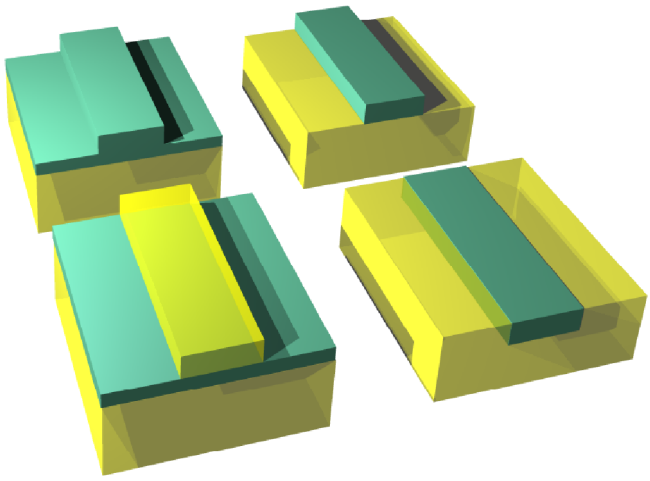


Figure 6. Some types of rectangular channel waveguides. Top left: ridge waveguide; top right: rib waveguide; bottom left: strip-loaded waveguide; bottom right: embedded waveguide. Waveguiding layer is denoted by darker color.

VI. Planar Waveguides

Many chemical and biological sensors base their operation on all-dielectric optical waveguides that utilize total internal reflection (TIR) for the wave propagation. In order to ensure waveguiding, the dimensions of such an optical guide along at least one axis should be comparable to the operating wavelength. A generic structure of a chemical sensor based on planar optical waveguide is shown in Fig. 5.

The bottom part of figure 5 illustrates how TIR supports a propagating wave within the waveguide. If the refractive index within the guide is larger than that of its environment, the propagation of a light beam within it will depend on its angle of incidence on the boundary between the guide and the ambient. If this angle is small enough, a part of the beam will reflect from the boundary, and a part will refract. When increasing the angle, the output beam will refract at a increasingly larger angle. For the incidence under the critical angle θ_c no light will leave the guide, and no light will be reflected too. All of it will instead propagate as an evanescent wave strictly along the boundary. For even larger angles of incidence the beam will be fully reflected from the boundary, it will travel to the boundary at the opposite side of the guide, bounce from it, and thus propagate through the guide after repeated reflections from its boundaries – this is the TIR effect.

If analyte is adsorbed at a boundary, usually via a receptor layer that specifically binds it, it will modify the local value of the refractive index. Thus the propagation constant of the guided mode will be changed and the ensuing change in propagation will be detected by some of the previously quoted external means. Thus the waveguide-based sensors are typical refractometric sensors. However, any of the previously listed transducing mechanism could be in principle used for their operation. Also, besides the described affinity method, some other mechanism could be used as well, for instance bulk absorption that causes physical swelling and the resulting change of the waveguide dimensions.

Fig. 6 illustrates some typical planar waveguide structures utilized in chemical sensing. These are rectangular waveguides denoted as channel waveguides. Top left picture represents a ridge waveguide; top right is a rib waveguide. Bottom left is a strip-loaded waveguide where a lower index material is deposited over a strip waveguide. Finally, bottom right picture represents an embedded waveguide. Besides these geometries, different other refractive index profiles and geometries can be also used.

VII. Optical Fibers

An optical fiber may be regarded as a special case of general optical waveguide where full cylindrical symmetry exists and where the length of the structure is much larger than its radial dimensions. Optical fibers are widely used in optical communications [38], but one of their important applications is in fiber optical sensing [31]. They are used for sensing of a wide range of physical quantities, and one of their applications is chemical and biological sensing.

In conventional optical fibers the propagation is again based on the principle of total internal reflection, the same as in planar waveguides. A typical optical fiber consists of a core with a higher refractive index and a cladding with a lower refractive index. If the refractive index abruptly changes between the core and the cladding, such structure is denoted as step-index fiber. If there is a gradient of refractive index, it is denoted as graded index structure. Fig. 7 shows some typical situations. The case (a) represents a multimode fiber with step index profile. The core has a larger refractive index and thus a TIR-guided propagation is enabled, the same as in the case of planar guides. The case (b) represents a single-mode fiber (smaller diameter core), (c) is a fiber with a gradient of refractive index and (d) is the situation with multiple steps.

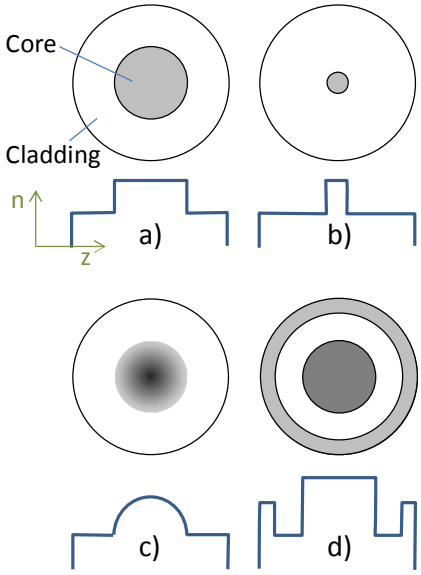


Figure 7. Some profiles of optical fibers: cross section and refractive index values. Darker shading corresponds to higher refractive indices. a) multimode fiber with step profile; b) single-mode fiber, step profile; c) graded index fiber; multiple-step fiber.

The usual way to utilize an optical fiber is to remove a part of cladding and let the evanescent field at the core boundary interact directly with the analyte. Another way is to utilize inverted profile fibers, for instance inverted graded index (IGI), where larger value of index are near the fiber surface [39].

As far as the readout is concerned, the most often used mechanisms are absorbance (as defined by Beer-Lambert-Bouguer extinction law), fluorescence measurement, diffraction on grating inscribed directly in the fiber and surface plasmon polaritons (to be described in more detail further in this text). Various phenomena can be utilized here, as defined in the description of transduction mechanisms.

One can use one fiber for illumination and another for detection, or alternatively the same fiber is used for both (Fig. 8).

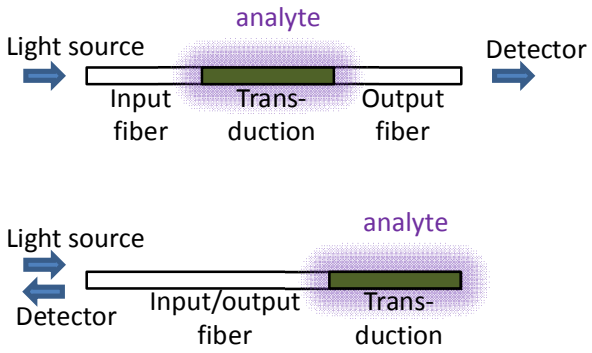


Figure 8. Readout schemes for chemical fiber optic sensors. Top: separate input and output (two-fiber scheme); bottom: single fiber scheme for simultaneous input and output.

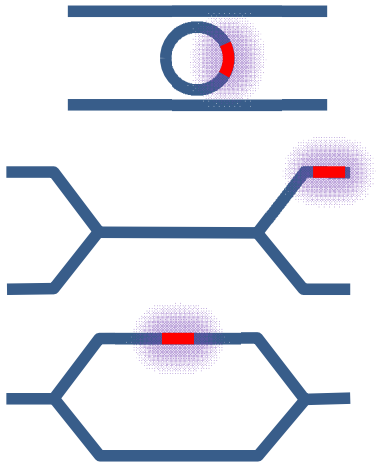


Figure 9. Top: Fiber ring resonator for chemical sensing; middle: fiber-based Mach-Zehnder interferometer as sensor; bottom: Michelson interferometer for chemical sensing.

One can fabricate resonators from fibers by performing coupling between a straight fiber section and a fiber ring, or between two fibers divided by a ring, as illustrated in Fig. 9 top. This ensures the appearance of sharp peaks in spectral characteristics, while the index change introduced by analyte results in a shift of the center frequency of the resonator.

Another possibility is to fabricate different interferometers using two branches of fibers, one of which is then exposed to analyte. Examples of such fiber structures for chemical sensing are Mach-Zehnder and Michelson interferometers (Fig. 9 middle and bottom).

VIII. Photonic Crystals

Photonic crystals are artificial structures with periodically changing refractive index in 1D, 2D or 3D, where the contrast between the low and high values of refractive index is sufficient to ensure the existence of the wavelength range in which no electromagnetic waves can propagate through the structure. The range of forbidden wavelengths is the photonic bandgap (PBG) and in that range a photonic crystal behaves as a near-perfect mirror. It could be said that photonic crystals represent a generalization of Bragg interference mirrors and also of Bragg gratings. Light in photonic crystals behaves similarly to charge carriers in semiconductors – besides the appearance of a bandgap, energy bands exist that correspond to the valence and the conductive band, it is possible to introduce defect and surface modes, etc.

The behavior of a photonic crystal is defined by the dimension of their unit cell and by the value of the refractive index (or dielectric permittivity) of each of its constituent parts.

The features of a unit cell of a photonic crystal are comparable to the wavelength and their behavior is based on the diffraction effects. In order to ensure the appearance of a full photonic bandgap, a photonic crystal structure must have a sufficiently large contrast of refractive index between its "high" and "low" parts.

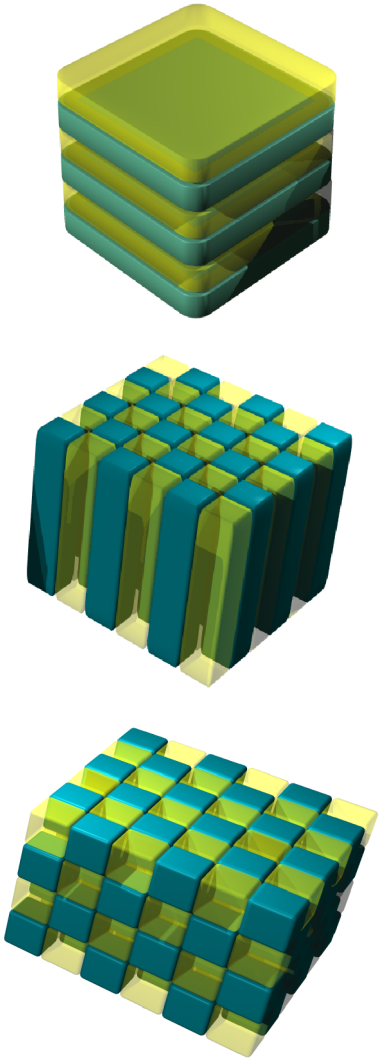


Figure 10. 1D, 2D and 3D Cartesian geometries (from top to bottom) for PBG-based chemical sensors

Photonic crystals were introduced in 1987 [40]. Until now they found a plethora of applications [41-43], among which an important place belongs to their use for chemical and biological sensing [19].

Structurally, there are three types of photonic crystals for chemical sensors, Fig. 10: high-contrast Bragg layers, basically functioning as interference-based dielectric mirrors (1D structures); planar waveguides or resonators utilizing 2D geometry and full 3D crystal.

The properties of photonic crystals can be tuned by the presence of analyte by two mechanisms:

a) changing refractive index due to the presence of analyte (for instance by adsorption or absorption) and

b) changing the unit cell dimensions, the simplest case being the shrinkage or swelling of the structure due to the presence of analyte, for instance by utilizing hydrogel that interacts with the targeted substance.

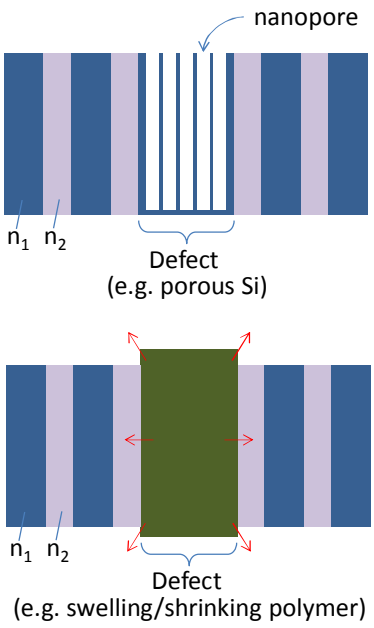


Figure 11. USe of 1D photonic crystals as chemical sensors. Top: absorption of analyte by porous defect (refractive index change); bottom: absorption of analyte in polymer that causes swelling (photonic crystal dimensions change)

The two basic methods of sensing of chemical substances are illustrated in Fig. 11 for the case of 1D photonic crystal. A layer with different thickness and/or refractive index is introduced in the structure. Depending on its properties, it functions as an acceptor or donor defect, as per equivalence with semiconductors. In the transmission spectrum it shows as a sharp peak (a dip in reflection). In the case shown in top the defect is made of a porous material with subwavelength holes, for instance nanoporous silicon. Upon exposure to chemicals, analyte fills the pores and modifies the effective value of refractive index. Instead of the scheme shown in Fig. 11, one can alternatively use structures where each layer is porous, but with different densities in different parts (higher hole-to-solid ratio means a lower effective refractive index).

The scheme shown in bottom illustrates the possibility to change the dimensions of the photonic crystal. To this purpose, a polymer layer is introduced as a defect. The exposure to analyte initiates specific binding between the polymer part and the analyte and the result is swelling or shrinking of the defect, which causes a shift of the defect peak. Again, as in the previous case, one can utilize the variable volume material not only for the defect, but for the whole photonic crystal or only for its high or low refractive index parts. In this case the whole photonic crystal is infused with analyte and the resulting thickness change modifies the spectral characteristics.

Obviously, the same kind of tunability is applicable in the case of 2D and 3D photonic crystals as well.

A special kind of PBG structures often used for chemical sensing are the photonic crystal fibers [44-46], also called the microstructured optical fibers. These are a special kind of 2D photonic crystals that are produced in the form of fibers. A high refractive index contrast in these structures is obtained by

introducing hollow (air filled) channels within the fiber, or using some other type of microstructuring to ensure the contrast. There are several basic types of photonic crystal fibers, as shown in Fig. 12. The first type are the so-called holey fibers, structures with a hollow core and hollow channels around it (top left). The core may also be solid (top right). There are combination of hollow channels with conventional step index optical fibers (bottom left) where the channels ensure better index contrast and thus better guiding. Finally, there is a Bragg-type fiber (bottom right) which basically represents a cylindrical version of 1D photonic crystal and where concentric layer with alternating low and high refractive index are used to confine light.

Of the listed structures, those with hollow core are the most convenient for sensing, since the core can be filled with analyte, the interrogating light is simultaneously confined within the same space and the optical paths can be very long. However, structures with solid core were also proposed for chemical sensing [42].

IX. Surface Plasmon Polariton Devices

Surface plasmons polaritons (SPP) may be defined as oscillations of free charge carriers in conductive material near an interface with dielectric which are coupled with electromagnetic radiation propagating across the interface. The relative dielectric permittivity ε in the conductive part is negative in the range of interest, and positive in the dielectric part. Typically the conductive material is metal (usually gold or silver. Other alternative materials can be sometimes used, for instance transparent conductive oxides (indium tin oxide, doped zinc oxide, tin oxide, etc.), semiconductor materials, intermetallics, graphene, etc. These conductors are all denoted as plasmonic materials [47-49].

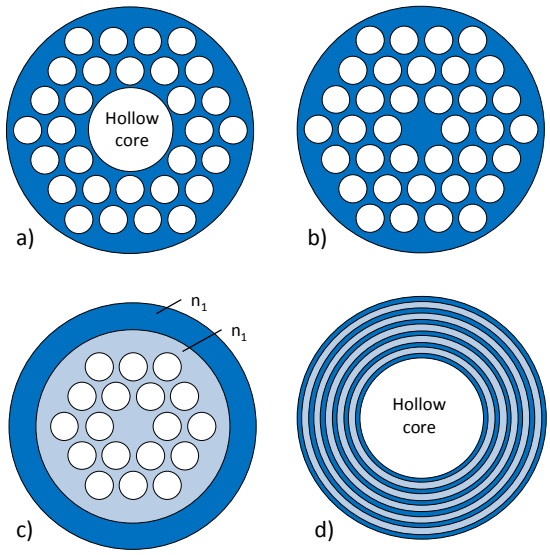


Figure 12. Different types of photonic crystal fibers for chemical sensors. a) hollow core fiber (holey fiber); b) solid-core fiber; c) mixed-type (hole-assisted fiber); d) Bragg fiber ("OmniGuide")

The electromagnetic wave in SPP is confined to the interface and evanescent (exponentially decaying) in the perpendicular direction. In chemical sensors one also encounters the spatially confined SPP that exist on a conductive nanoparticle (localized surface plasmons polaritons, LSPP). The science dedicated to the research and application of SPP phenomena and devices is called plasmonics [50-53].

In the simplest case, SPPs at the interface between two semi-infinite media with opposite signs of dielectric permittivity are not plane electromagnetic waves. Their energy is concentrated in the narrow region near the boundary plane. SPPs are TM (transverse-magnetic) polarized, and because of that they are called *polaritons*. In other words, magnetic field and wavevector of the SPP lay in the plane of interface, while electric field of the wave is both perpendicular and parallel to the wavevector components. TE polarized component of electromagnetic field cannot form a surface wave at a simple metal-dielectric interface.

A SPP wave has its largest value in the interface between the ε-positive and ε-negative part and decays exponentially in the both normal directions. If an analyte adsorbs to the sensor surface in an ultrathin layer, this will mean that the regions where the SPP is concentrated and the region where the analyte is present will overlap. Thus an SPP-based sensor will be able to "see" even minute amounts of analyte. Actually the conventional SPP sensors have long been able to distinguish even a few percent change of the composition of a monatomic or monomolecular layer [54].

The wavevector of a SPP is much larger than that of a propagating optical beam in the free space for the same frequency. Thus it is necessary to match the wavevector between the two. The most often used approach for the conventional SPP chemical sensors is the Kretchmann method [55]. An illustration of this method is shown in Fig. 5, with a dielectric hemicylinder presented (often a prism is used instead), with a goal to show that the interrogating beam is incident perpendicularly to the dielectric surface. The interrogating beam arrives to the metalized flat surface at exactly the critical angle and thus the created evanescent wave is matched to the SPP.

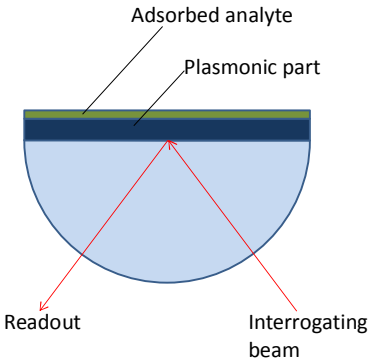


Figure 13. Kretchmann configuration for SPP sensor readout

At the surface plasmon resonance there will be a strong dip in the reflection spectrum that corresponds to a given angle of incidence and output. When an ultrathin layer of analyte adsorbs to the surface, its refractive index changes the propagating conditions of the evanescent SPP waves and thus the angle is changed. The angular shift is detected in the readout signal, and its amount is determined by the refractive index of analyte. Thus a SPP sensor is actually a refractometric sensor, basing its operation on the refractive index change within the near-surface layer with a subwavelength thickness.

The resonant nature of the SPPs actually represents an advantage for sensing purposes, since it causes steep gradients in their frequency dispersion, which results in large changes of scattering parameters for very small changes of the surface refractive index. Absorption losses in SPP structures are very large due to collisions of free carriers in the ε-negative material, which leads to shorter propagation paths, but this is not an issue for chemical sensing purposes.

The relative dielectric permittivity ε of plasmonic materials is negative below plasma frequency, and its dispersion is well-described by the electron resonance model of Drude [56], also denoted as Drude-Sommerfeld model

$$\varepsilon = \varepsilon_\infty - \frac{\omega_p^2}{\omega(\omega + i\Gamma)}, \tag{5}$$

where ω_p is the plasma frequency, Γ is the damping factor (i.e. the imaginary part of the complex dielectric permittivity), and ε_∞ is the asymptotic dielectric permittivity. The plasma frequency is

$$\omega_p^2 = \frac{n_e e^2}{m^* \varepsilon_0}, \tag{6}$$

where n_e is the free electron concentration in epsilon-negative part, e is the free electron charge ($1.6 \cdot 10^{-19}$ C), ε_0 is the vacuum permittivity ($8.854 \cdot 10^{-12}$ F/m), and m^* is the electron effective mass.

The damping factor is

$$\Gamma = \frac{e}{\mu m^*} \tag{7}$$

where μ is the mobility of free electrons.

In the case of interband transitions between the valence and the conduction bands, dielectric permittivity is given by Lorentz model [57]

$$\varepsilon = \varepsilon_\infty + \frac{\omega'^2_p}{(\omega_0^2 - \omega^2) - i\Gamma'\omega}, \tag{8}$$

where ω_o is the resonant frequency of the electron oscillator, while the apostrophe denotes values related with the concentration of bound electrons that participate in the interband transitions.

The propagation constant of a TM polarized wave at the metal (*m*)-dielectric (*d*) interface is given as

$$\beta = \frac{\omega}{c}\sqrt{\frac{\varepsilon_m \varepsilon_d}{\varepsilon_m + \varepsilon_d}}, \tag{8}$$

Thus a change of the dielectric permittivity (and therefore the refractive index) of the dielectric part at the interface will modify the propagation constant, thus changing the propagation of the SPP. This is the basis of the use of SPP structures for sensing.

The applicability of SPP resonance for chemical sensing enhancement has been recognized very early [15], and indeed some of the first proposed applications of the SPP were in refractometric sensing in thin layers at the ultrathin boundary layer between metal and dielectric. Currently this is probably the most often used optical method for chemical sensing [4, 53, 58-60].

X. Optical Metamaterial Sensors

The possibility to structure material at nanometer level has brought to a new paradigm, the electromagnetic metamaterials [61-67]. These can be defined as artificial structures with subwavelength features and with effective optical properties that exceed those of natural media. Theoretically envisioned in 1967 [68], practically described in 1999 [69] and experimentally proven in 2001 [70], metamaterials showed themselves applicable in a vast number of areas. The first described metamaterials were those with negative refractive index [69], but subsequently the term has spread to other structures with unusual values of refractive index, for instance zero refractive index . Metamaterials for the optical range were described in [71, 72], and those utilizing surface plasmons polaritons in [61].

Since one is able to tailor a metamaterial structure almost at will, this means that its dispersion relation could be fully tailored. It is thus possible to design modes with superluminal group velocities ("fast light"), near-zero ("slow light") and, as mentioned above, negative ("left-handed light," propagating in the direction opposite to that of the phase velocity) [73].

The possibility to design a frequency dispersion at will ensures tailoring of spatially strongly localized near-field modes (due to shortening of the wavelength for a given frequency), thus obtaining a very high density of states compared to the free space. The same energy that a propagating wave would carry and at the same frequency is here compacted into a much smaller space, thus ensuring much higher wave densities and much shorter wavelengths for that frequency. This ensures highly enhanced interaction of optical radiation with the surrounding material. This kind of engineering of optical interaction with matter ensures its maximization in the active area of a sensor, with a consequence

that its response is not only vastly increased, but also localized to a specific nano-volume. Thus a chemical sensor based on optical metamaterials will offer a significantly improved sensitivity. At the same time, a plethora of novel modes of operation will exist. Contrary to e.g. SPP sensors, no special readout schemes like Kretschmann prisms are necessary and the sensor designer freedom is much larger.

Among the first descriptions of metamaterial chemical sensors are [74, 75]. There are many different schemes for this kind of devices [7, 13, 76-78]. Among the most convenient ones are the so-called metasurfaces, i.e. planar metamaterials that are compatible with the conventional planar technologies [13, 79, 80].

Probably the best known negative index metamaterial for the optical range is the fishnet, [81-84] the structure that consists of a metal-dielectric-metal sandwich perforated by an array of nanoholes, Fig. 14. The use of fishnet metamaterials for chemical sensing was described in [85, 86].

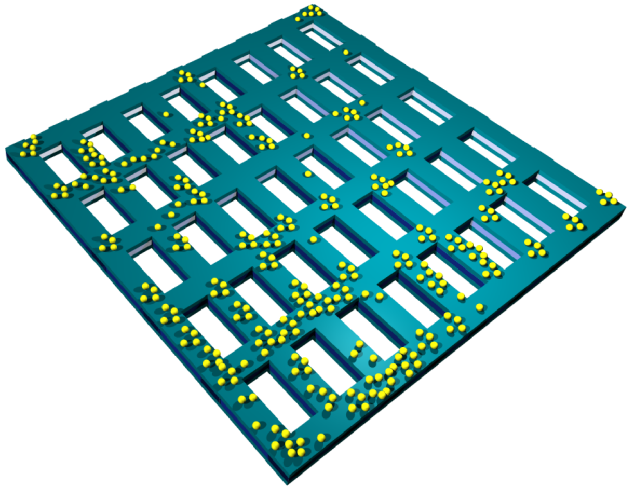


Figure 14. Fishnet metamaterial with rectangular holes. Spheres denote analyte atoms or molecules.

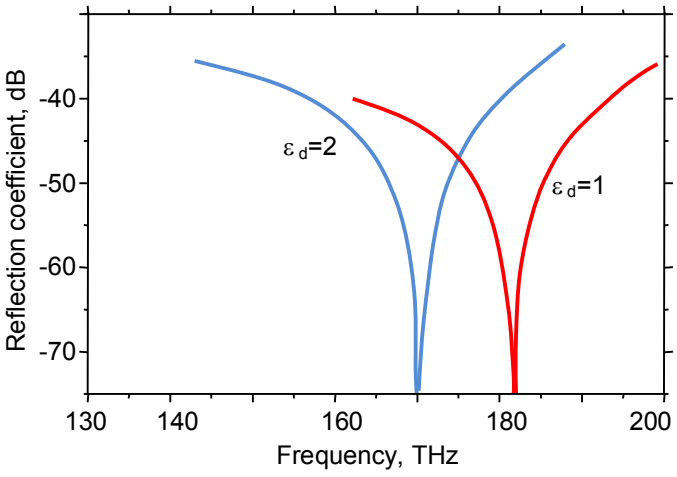


Figure 15. Spectral response of a fishnet metamaterial. Aperture size 320 nm x 280 nm, metal thickness 20 nm, dielectric thickness 200 nm, TM polarized in-plane response observed.

The whole structure is dipped into analyte-containing fluid, i.e. the absorption scheme is used to obtain the response shown in Fig. 15. This causes the shift of the dip in spectral reflection. This is a setup obtained by applying microfludics where analyte flows through nanoapertures, as described in [87] for extraordinary optical transmission case. Thinner fishnets and smaller apertures are more strongly modified by analyte presence.

Due to their wider range of available values of refractive index, metamaterial structures ensure much better control over electromagnetic fields. Thus a whole new field of electromagnetic optics emerged, the transformation optics [88-92]. This ensured different applications, like hyperlenses for conversion between near-field and propagating far-field modes, cloaking devices, superconcentrators and superabsorbers [93-98]. All of these structures can be utilized in chemical or biological sensing [99], ensuring a large number of novel modes of operation of these devices.

XI. Nanoplasmonic Structures

Nanoplasmonic structures [25, 100-102] are nanocomposites consisting of metals and dielectrics – or, in a more general case, of materials with positive and with negative relative dielectric permittivity). They represent a generalization of surface plasmon polariton structures – instead of having a simple flat boundary between semi-infinite metal and semi-infinite dielectric, the boundary here may have virtually any shape. A plethora of different field modes is possible in such structures. They ensure full control over both frequency dispersion and spatial dispersion, resulting in very high field enhancements and unprecedented resolutions. The behaviour of these structures is covered by the emerging science of plasmonics [50, 52]. Probably the best contemporary sensors of chemical and biological analyte belong to the class of nanoplasmonic devices [4, 12, 17, 25, 103].

Nanoplasmonic structures can be one-dimensional – planar superlattices with alternating metal and dielectric layers, 2D – metallic nanowires, nanorods, etc. embedded in dielectric, like nanometer-sized analogue of artificial dielectrics[104], or 3D – arrays of metallic nanoparticles embedded in dielectric). If the structure of nanoplasmonic composites is ordered and periodic, they are called plasmonic crystals [105]. Nanoplasmonic structures can be also quasiperiodic (plasmonic quasicrystals) [106], aperiodic [107] or even random [108].

The basic building block of a nanoplasmonic structure, i.e. the artificial "atom" of such a plasmonic metamaterial, can have any form. This is in contrast to the natural materials, where there is no choice regarding the shape of the building blocks. The shapes of this metallic "atom" within the dielectric host may vary from simple sphere to very complex form from regular to irregular [77].

A fact of large importance for sensing is the possibility to produce additional field localizations by changing the shape of the plasmonic inclusions and their geometry. Similar to the edge phenomena in macroscopic world, the plasmonic inclusions with sharp corners will localize electromagnetic fields, the amount of localization being larger for sharper corners. Also, plasmonic nanoparticles with small interparticle

distance (compared to the skin depth in metal) will cause strong coupling of surface plasmons polaritons between these particles, creating large field localizations in the space between them. This will cause high nonlocal effects and extreme localizations of electromagnetic field. A plethora of new modes and novel effects of interest for chemical sensing appears in such structures [105, 109].

One of the typical situations is encountered if randomly distributed plasmonic nanoparticles are utilized. The may be placed on a surface (Fig. 16) or mixed with a fluid that contains analyte.

The field intensity around the nanoparticles is high due to the establishment of localized surface plasmons polaritons (LSPP) [110-113]. When analyte is present, its refractive index will change the electromagnetic conditions and the result will be for instance color shift. Fluorescence schemes can be also used in this case.

Another simple structure used for sensing are plasmonic nanomembranes [6, 114-117]. These are freestanding structures with thickness below 100 nm and extremely high aspect ratios, making their lateral dimensions reach millimeter, even centimeter range. When a surface plasmon polariton travels along one side of such a structure, it couples electromagnetically with the SPP on the other side. The result is the establishment of a long range surface plasmon polariton, which decays more slowly away from the surface and propagates at much larger distances than the usual SPP. This means that its "information distance" from the surface at which it can sense analytes is much larger than in conventional SPP structures. Because of that it is able to detect much larger analytes, including macromolecules and even different living cells.

A nanoplasmonic structure of a large importance for chemical sensing are extraordinary transmission arrays. If one opens nanoholes with dimensions much smaller than the operating wavelength in an optically opaque metal layer, according to the classical Bethe theory no light should be transmitted through the nanoholes because of the polarization of light. However, in 1998 Ebessen experimentally proved that a near 100% transmission can be obtained through such aperture array and denoted the effect extraordinary optical transmission (EOT) [118]. It was proven later that the effect is a consequence of the establishment of surface plasmons polaritons at the front and the back surface and the coupling of the SPP between these surfaces.

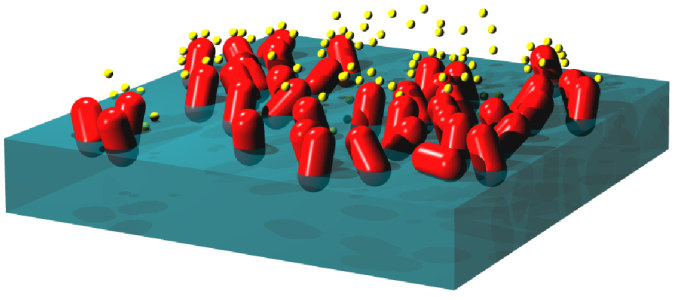


Figure 16. Nanoparticles placed on a substrate (capsule shapes) surrounded by analyte (small spheres).

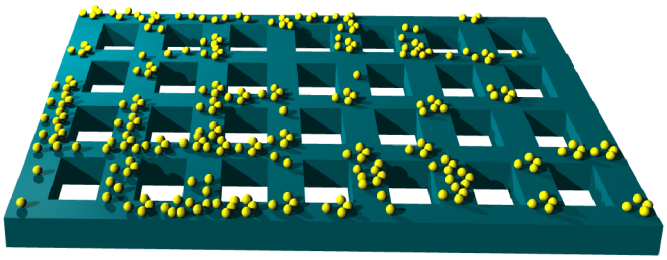


Figure 17. Extraordinary optical transmission structure.

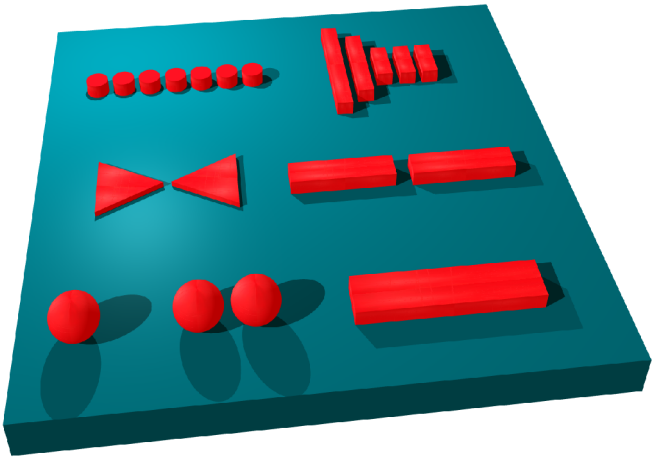


Figure 18. Various types of nanoantennas. Top left: linear nanoparticle array; top right: Udo-Yagi nanoantenna; middle left: bow tie antenna; middle right: two-wire nanoantenna. Bottom left: nanosphere, the simplest nanoantenna; bottom middle: nanodimer, two coupled spherical nanoantennas; bottom right: nanorod (single wire antenna).

The field intensity within the nanoapertures becomes several orders of magnitude larger than that impinging at the surface. Thus such structures ensure strong enhancement of nonlinear effects which is beneficial for selectivity of chemical sensing since different effects will occur in different analytes. At the same time, field concentration ensures high sensitivity chemical sensing at low analyte concentrations. Fig. 17 shows an EOT array.

Similarly to fishnet metamaterials, EOT structures are very convenient for microfluidic chemical sensing schemes, where fluid that contains analyte is pumped through nanoapertures [87].

Probably the most sensitive nanoplasmonic sensors are those based on nanoantennas [3, 119-122]. Nanoantenna or optical antenna is a plasmonic device that converts propagating waves into evanescent ones and back with a high efficiency. Field localization depends on the nanoantenna design and can be made deeply subwavelength. Thus interaction with photodetector active region can be vastly enhanced.

Nanoantennas are isolated structures, i.e. they function without a connection to any circuitry. Thus a simple spherical nanoparticle can be basically regarded as a dipole nanoantenna,

Fig. 18 bottom left. Other nanoantenna types met in literature include nanodimer, consisting of two nanosphere at a small (subwavelength) distance so that they are electromagnetically coupled. The distance between the two spheres in a nanodimer is denoted as feed gap and electromagnetic field localizations in it can be extremely strong. Other types are nanorod antennas (bottom right), bow tie antennas (middle left) that combine very small feed gap with the effect of sharp edges and thus ensures very high field localizations. Besides that it furnishes a broader bandwidth of optical frequencies. Two nanorods coupled together make a two-wire antenna (middle right). An array of nanodisks makes a complex linear nanoparticle array antenna. Stacking nanorods can make an equivalent of the classical Udo-Yagi antenna (top right). A vast number of different other nanoantennas has been reported until now.

In 2012 a golden nanorod antenna has been used to ensure sensing of single non-absorbing molecules [1]. The same type of nanoantenna was utilized for in-situ, real time interactions with biomolecules in infrared spectrum (the previously quoted SEIRA method) [3]. Different types of plasmonic antennas are being considered for various applications in chemical and biological sensing.

XII. Conclusion

A review of different optical sensing schemes for the determination of chemical, biochemical or biological properties has been given. From the point of view of industrial measurements one of the most convenient families of sensors are those utilizing optical effects, because of their inherent insensitivity to electromagnetic interference, robustness under difficult operating conditions (vibrations, temperature stress, corrosive or aggressive media, etc.) and built-in fire protection and Ex-proofing since typically no electric signals are present. We proposed a new classification that is more general than that given by IUPAC and thus points out to some possible novel measurement mechanisms. Our accent is given to the more recently proposed subwavelength structures that utilize near field effects, including those based on optical metamaterials and plasmonics. Such structure ensure an increased degree of design freedom and introduce a number of new modes of operation. Combined with possible wireless transfer of sensor signal and thus an increased density of measurement locations and even larger robustness against difficult measurement conditions, this ensures a bright future for optical chemical sensor applications in industry. Obviously, the described sensors can be utilized in other areas as well, including but not limited to homeland defense (detection of toxins, explosives, etc), biomedical applications and environmental protection.

References

[1] P. Zijlstra, P. M. R. Paulo, and M. Orrit, "Optical detection of single non-absorbing molecules using the surface plasmon resonance of a gold nanorod," *Nature Nanotech.,* vol. 7, no. 6, pp. 379-382, 2012.

[2] "Global Markets and Technologies for Sensors, Report Code: IAS006F, July 2014," http://www.bccresearch.com/market-research/instrumentation-and-sensors/sensors-ias006f.html.

[3] R. Adato, and H. Altug, "In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas," *Nature Comm.,* vol. 4, pp. 2154, 2013.

[4] J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, "Biosensing with plasmonic nanosensors," *Nature Mater.,* vol. 7, no. 6, pp. 442-453, 2008.

[5] Y. Cui, Q. Wei, H. Park, and C. M. Lieber, "Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species," *Science,* vol. 293, no. 5533, pp. 1289-1292, 2001.

[6] C. Jiang, S. Markutsya, Y. Pikus, and V. V. Tsukruk, "Freely suspended nanocomposite membranes as highly sensitive sensors," *Nature Mater.,* vol. 3, no. 10, pp. 721-728, 2004.

[7] A. V. Kabashin, P. Evans, S. Pastkovsky, W. Hendren, G. A. Wurtz, R. Atkinson, R. Pollard, V. A. Podolskiy, and A. V. Zayats, "Plasmonic nanorod metamaterials for biosensing," *Nature Mater.,* vol. 8, no. 11, pp. 867-871, 2009.

[8] S. Lal, S. Link, and N. J. Halas, "Nano-optics from sensing to waveguiding," *Nature Photonics,* vol. 1, no. 11, pp. 641-648, 2007.

[9] D. K. Lim, K. S. Jeon, H. M. Kim, J. M. Nam, and Y. D. Suh, "Nanogap-engineerable raman-active nanodumbbells for single-molecule detection," *Nature Mater.,* vol. 9, no. 1, pp. 60-67, 2010.

[10] T. Rindzevicius, Y. Alaverdyan, A. Dahlin, F. Höök, D. S. Sutherland, and M. Käll, "Plasmonic Sensing Characteristics of Single Nanometric Holes," *Nano Lett.,* vol. 5, no. 11, pp. 2335-2339, 2005.

[11] Z. S. Siwy, and M. Davenport, "Biosensors: Making nanopores from nanotubes: Single-walled carbon nanotubes can be used to detect single DNA molecules as they pass through the nanotubes under the influence of an applied electric field," *Nature Nanotech.,* vol. 5, no. 3, pp. 174-175, 2010.

[12] I. Choi, and Y. Choi, "Plasmonic nanosensors: Review and prospect," *IEEE J. Sel. Top. Quant. Electr.,* vol. 18, no. 3, pp. 1110-1121, 2012.

[13] Z. Jakšić, S. M. Vuković, J. Matovic, and D. Tanasković, "Negative Refractive Index Metasurfaces for Enhanced Biosensing," *Materials,* vol. 4, no. 1, pp. 1-36, 2011.

[14] V. K. Khanna, "Nanoparticle-based sensors," *Defence Science Journal,* vol. 58, no. 5, pp. 608-616, 2008.

[15] B. Liedberg, C. Nylander, and I. Lunström, "Surface plasmon resonance for gas detection and biosensing," *Sens. Act.,* vol. 4, pp. 299-304, 1983.

[16] L. Senesac, and T. G. Thundat, "Nanosensors for trace explosive detection," *Materials Today,* vol. 11, no. 3, pp. 28-36, 2008.

[17] M. E. Stewart, C. R. Anderton, L. B. Thompson, J. Maria, S. K. Gray, J. A. Rogers, and R. G. Nuzzo, "Nanostructured plasmonic sensors," *Chem. Rev.,* vol. 108, no. 2, pp. 494-521, 2008.

[18] R. Bogue, "Optical chemical sensors for industrial applications," *Sensor Review,* vol. 27, no. 2, pp. 86-90, 2007.

[19] C. Fenzl, T. Hirsch, and O. S. Wolfbeis, "Photonic crystals for chemical sensing and biosensing," *Angewandte Chemie - International Edition,* vol. 53, no. 13, pp. 3318-3335, 2014.

[20] F. Baldini, A. N. Chester, J. Homola, and S. Martellucci, eds., *Optical Chemical Sensors*, Dordrecht: Springer, 2006.

[21] J. Becker, *Plasmons as Sensors*, Springer, Berlin, Heidelberg, 2012.

[22] W. J. Bock, I. Gannot, and S. Tanev, eds., *Optical Waveguide Sensing and Imaging*, Dordrecht: Springer, 2006.

[23] A. Cusano, F. J. Arregui, M. Giordano, and A. Cutolo, eds., *Optochemical Nanosensors*, Boca Raton: CRC Press, Taylor & Francis Group, 2013.

[24] A. P. Demchenko, *Introduction to Fluorescence Sensing*, Springer, New York, 2009.

[25] A. Dmitriev, eds., *Nanoplasmonic Sensors*, Berlin, Heidelberg: Springer, 2012.

[26] X. Fan, eds., *Advanced Photonic Structures for Biological and Chemical Detection*, Dordrecht: Springer, 2009.

[27] W. Fritzsche, and J. Popp, eds., *Optical Nano- and Microsystems for Bioanalytics*, Berlin Heidelberg: Springer-Verlag, 2012.

[28] J. Haus, *Optical Sensors: Basics and Applications*, Wiley-VCH, Weinheim, 2010.

[29] K. Kneipp , M. Moskovits, and H. Kneipp, eds., *Surface-Enhanced Raman Scattering: Physics and Applications*, Berlin Heidelberg: Springer 2006.

[30] F. S. Ligler, and C. R. Taitt, eds., *Optical biosensors: today and tomorrow*, Amsterdam: Elsevier, 2011.

[31] S. Yin, P. B. Ruffin, and F. T. S. Yu, eds., *Fiber Optic Sensors*, Bocca Raton: CRC Taylor & Francis Group, 2008.

[32] M. Zourob, and A. Lakhtakia, eds., *Optical Guided-wave Chemical and Biosensors I-II*, Berlin Heidelberg: Springer-Verlag, 2010.
[33] A. Hulanicki, S. Glab, and F. Ingman, "Chemical sensors definitions and classification," *Pure & Appl. Chem,* vol. 63, no. 9, pp. 1247-1250, 1991.
[34] D. W. Lübbers, and N. Opitz, "The pCO2-/pO2-optode: a new probe for measurement of pCO2 or pO in fluids and gases (authors transl)," *Zeitschrift fur Naturforschung. Section C: Biosciences,* vol. 30, no. 4, pp. 532-533, 1975.
[35] F. P. Milanovich, T. B. Hirschfeld, F. T. Wang, S. M. Klainer, and D. Walt, "Clinical measurements using fiber optics and optrodes," *Proc. SPIE*, 494, pp. 18-24.
[36] J. S. Schultz, S. Mansouri, and I. J. Goldstein, "Affinity sensor: A new technique for developing implantable sensors for glucose and other metabolites," *Diabetes Care,* vol. 5, no. 3, pp. 245-253, 1982.
[37] J. S. Schultz, and G. Sims, "Affinity sensors for individual metabolites," *Biotechnology Bioengineering Symposium,* vol. No. 9, pp. 65-71, 1979.
[38] G. P. Agrawal, *Fiber-optic communication systems*, Wiley, New York, NY, 1997.
[39] V. Matějec, M. Chomát, I. Kašík, J. Čtyroký, D. Berková, and M. Hayer, "Inverted-graded index fiber structures for evanescent-wave chemical sensing," *Sensors and Actuators B: Chemical,* vol. 51, no. 1–3, pp. 340-347, 1998.
[40] E. Yablonovitch, "Inhibited spontaneous emission in solid-state physics and electronics," *Phys. Rev. Lett.,* vol. 58, no. 20, pp. 2059-2062, 1987.
[41] J. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals: Molding the Flow of Light 2nd ed*, Princeton University Press, Princeton, 2008.
[42] J. M. Lourtioz, H. Benisty, V. Berger, J.-M. Gerard, D. Maystre, and A. Tchelnokov, *Photonic Crystals: Towards Nanoscale Photonic Devices 2nd ed*, Springer-Verlag Berlin Heidelberg, 2008
[43] F. J. P. Schuurmans, "Light in complex dielectrics," PhD. dissertation, University of Amsterdam, Amsterdam, 1999.
[44] P. Russell, "Photonic crystal fibers," *Science,* vol. 299, no. 5605, pp. 358-362, 2003.
[45] J. Broeng, D. Mogilevstev, S. E. Barkou, and A. Bjarklev, "Photonic crystal fibers: A new class of optical waveguides," *Optical fiber technology,* vol. 5, no. 3, pp. 305-330, 1999.
[46] F. Poli, A. Cucinotta, and S. Selleri, *Photonic Crystal Fibers: Properties and Applications*, Springer, Dordrecht, 2007.
[47] A. Boltasseva, and H. A. Atwater, "Low-Loss Plasmonic Metamaterials," *Science,* vol. 331, no. 6015, pp. 290-291, 2011.
[48] P. Avouris, and M. Freitag, "Graphene photonics, plasmonics, and optoelectronics," *IEEE J. Sel. Top. Quant. Electr.,* vol. 20, no. 1, 2014.
[49] Z. Jakšić, S. M. Vuković, J. Buha, and J. Matovic, "Nanomembrane-Based Plasmonics," *J. Nanophotonics,* vol. 5, pp. 051818.1-20, 2011.
[50] S. A. Maier, *Plasmonics: Fundamentals and Applications*, Springer Science+Business Media, New York, NY, 2007.
[51] S. A. Maier, and H. A. Atwater, "Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures," *J. Appl. Phys.,* vol. 98, no. 1, pp. 1-10, 2005.
[52] E. Ozbay, "Plasmonics: Merging Photonics and Electronics at Nanoscale Dimensions," *Science,* vol. 311, no. 5758, pp. 189-193, 2006.
[53] R. B. M. Schasfoort, and A. J. Tudos, eds., *Handbook of Surface Plasmon Resonance*, Cambridge, UK: Royal Society of Chemistry 2008.
[54] L. S. Jung, C. T. Campbell, T. M. Chinowsky, M. N. Mar, and S. S. Yee, "Quantitative interpretation of the response of surface plasmon resonance sensors to adsorbed films," *Langmuir,* vol. 14, no. 19, pp. 5636-5648, 1998.
[55] E. Kretschmann, "Die Bestimmung optischer Konstanten von Metallen durch Anregung von Oberflächenplasmaschwingungen," *Z. Phys. A Hadr. Nucl.,* vol. 241, no. 4, pp. 313-324, 1971.
[56] P. Drude, *The Theory of Optics*, Dover Publications, Mineola, New York, 2005.
[57] H. A. Lorentz, *The Theory of Electrons*, Dover Publications, Mineola, New York, 1952.
[58] I. Abdulhalim, M. Zourob, and A. Lakhtakia, "Surface plasmon resonance for biosensing: A mini-review," *Electromagnetics,* vol. 28, no. 3, pp. 214-242, 2008.
[59] N. J. de Mol, and M. J. E. Fischer, eds., *Surface Plasmon Resonance Methods and Protocols*, New York: Humana Press – Springer Science, 2010.
[60] J. Homola, and M. Piliarik, "Surface Plasmon Resonance (SPR) Sensors," *Surface Plasmon Resonance Based Sensors*, J. Homola, ed., pp. 45-67, Berlin Heidelberg: Springer, 2006.
[61] W. Cai, and V. Shalaev, *Optical Metamaterials: Fundamentals and Applications*, Springer, Dordrecht , Germany, 2009.
[62] F. Capolino, eds., *Metamaterials Handbook: Theory and Phenomena of Metamaterials*, Boca Raton: CRC, 2009.
[63] T. J. Cui, D. R. Smith, and R. Liu, eds., *Metamaterials: Theory, Design, and Applications*, New York Springer 2009.
[64] N. Engheta, and R. W. Ziolkowski, eds., *Metamaterials: Physics and Engineering Explorations*, Hoboken, New Jersey: Wiley-IEEE Press, 2006.
[65] A. Sarychev, and V. Shalaev, *Electrodynamics of metamaterials*, World Scientific, Singapore, 2007.
[66] L. Solymar, and E. Shamonina, *Waves in Metamaterials* Oxford University Press, Oxford 2009.
[67] E. J. Tremblay, eds., *Metamaterials: Classes, Properties and Applications*, Hauppauge, New York: Nova Science Publishers, 2010.
[68] V. G. Veselago, "The electrodynamics of substances with simultaneously negative values of mu and epsilon," *Sov. Phys. Uspekhi,* vol. 10, no. 4, pp. 509-514, 1968.
[69] J. B. Pendry, A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE T. Microw. Theory,* vol. 47, no. 11, pp. 2075-2084, 1999.
[70] R. A. Shelby, D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science,* vol. 292, no. 5514, pp. 77-79, 2001.
[71] V. M. Shalaev, "Optical negative-index metamaterials," *Nature Photonics,* vol. 1, no. 1, pp. 41-48, 2007.
[72] V. M. Shalaev, W. Cai, U. K. Chettiar, H. K. Yuan, A. K. Sarychev, V. P. Drachev, and A. V. Kildishev, "Negative index of refraction in optical metamaterials," *Opt. Lett.,* vol. 30, no. 24, pp. 3356-3358, 2005.
[73] P. W. Milonni, *Fast Light, Slow Light and Left-Handed Light*, Taylor & Francis, Abingdon, Oxford, 2004.
[74] A. Ishimaru, S. Jaruwatanadilok, and Y. Kuga, "Generalized surface plasmon resonance sensors using metamaterials and negative index materials," *Prog. Electromagn. Res.,* vol. 51, pp. 139-152, 2005.
[75] Z. Jakšić, O. Jakšić, Z. Djurić, and C. Kment, "A consideration of the use of metamaterials for sensing applications: Field fluctuations and ultimate performance," *J. Opt. A-Pure Appl. Opt.,* vol. 9, no. 9, pp. S377-S384, 2007.
[76] C. M. Bingham, H. Tao, X. Liu, R. D. Averitt, X. Zhang, and W. J. Padilla, "Planar wallpaper group metamaterials for novel terahertz applications," *Opt. Express,* vol. 16, no. 23, pp. 18565-18575, 2008.
[77] Z. Jakšić, "Optical metamaterials as the platform for a novel generation of ultrasensitive chemical or biological sensors," *Metamaterials: Classes, Properties and Applications*, E. J. Tremblay, ed., pp. 1-42, Hauppauge, New York: Nova Science Publishers, 2010.
[78] N. Liu, T. Weiss, M. Mesch, L. Langguth, U. Eigenthaler, M. Hirscher, C. Sönnichsen, and H. Giessen, "Planar Metamaterial Analogue of Electromagnetically Induced Transparency for Plasmonic Sensing," *Nano Lett.*, pp. ASAP, 2009.
[79] A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Planar photonics with metasurfaces," *Science,* vol. 339, no. 6125, pp. 12320091-12320096, 2013.
[80] N. Yu, and F. Capasso, "Flat optics with designer metasurfaces," *Nature Mater.,* vol. 13, no. 2, pp. 139-150, 2014.
[81] S. Zhang, W. Fan, N. C. Panoiu, K. J. Malloy, R. M. Osgood, and S. R. J. Brueck, "Experimental demonstration of near-infrared negative-index metamaterials," *Phys. Rev. Lett.,* vol. 95, no. 13, pp. 1-4, 2005.
[82] K. B. Alici, and E. Ozbay, "A planar metamaterial: Polarization independent fishnet structure," *Photonic. Nanostruct.,* vol. 6, no. 1, pp. 102-107, 2008.
[83] M. Kafesaki, I. Tsiapa, N. Katsarakis, T. Koschny, C. M. Soukoulis, and E. N. Economou, "Left-handed metamaterials: The fishnet structure and its variations," *Phys. Rev. B,* vol. 75, no. 23, pp. 235114, 2007.

[84] S. Xiao, U. K. Chettiar, A. V. Kildishev, V. P. Drachev, and V. M. Shalaev, "Yellow-light negative-index metamaterials," *Opt. Lett.,* vol. 34, no. 22, pp. 3478-80, 2009.
[85] Z. Jakšić, D. Tanasković, and J. Matović, "Fishnet-based metamaterials: spectral tuning through adsorption mechanism," *Acta Phys. Pol. A,* vol. 116, no. 4, pp. 333-335, 2009.
[86] D. Tanasković, M. Obradov, O. Jakšić, and Z. Jakšić, "A low-loss double-fishnet metamaterial based on transparent conductive oxide," *Physica Scripta,* vol. 2014, no. T162, pp. 014048, 2014.
[87] R. Gordon, D. Sinton, K. L. Kavanagh, and A. G. Brolo, "A new generation of sensors based on extraordinary optical transmission," *Acc. Chem. Res.,* vol. 41, no. 8, pp. 1049-1057, 2008.
[88] H. Chen, C. T. Chan, and P. Sheng, "Transformation optics and metamaterials," *Nature Mater.,* vol. 9, no. 5, pp. 387-396, 2010.
[89] U. Leonhardt, "Optical Conformal Mapping," *Science,* vol. 312, no. 5781, pp. 1777-1780, 2006.
[90] U. Leonhardt, and T. G. Philbin, "Transformation Optics and the Geometry of Light," *Progress in Optics*, E. Wolf, ed., pp. 69-152, Amsterdam, The Netherlands: Elsevier Science & Technology 2009.
[91] J. B. Pendry, "Negative refraction makes a perfect lens," *Phys. Rev. Lett.,* vol. 85, no. 18, pp. 3966-3969, 2000.
[92] J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling Electromagnetic Fields," *Science,* vol. 312, no. 5781, pp. 1780-1782, 2006.
[93] A. Alù, and N. Engheta, "Achieving transparency with plasmonic and metamaterial coatings," *Phys. Rev. E,* vol. 72, no. 1, pp. 016623, 2005.
[94] W. Cai, U. K. Chettiar, A. V. Kildishev, and V. M. Shalaev, "Optical cloaking with metamaterials," *Nature Photonics,* vol. 1, no. 4, pp. 224-227, 2007.
[95] T. Ergin, N. Stenger, P. Brenner, J. B. Pendry, and M. Wegener, "Three-dimensional invisibility cloak at optical wavelengths," *Science,* vol. 328, no. 5976, pp. 337-339, 2010.
[96] Z. Jacob, L. V. Alekseyev, and E. Narimanov, "Optical hyperlens: Far-field imaging beyond the diffraction limit," *Opt. Express,* vol. 14, no. 18, pp. 8247-8256, 2006.
[97] J. Ng, H. Chen, and C. T. Chan, "Metamaterial frequency-selective superabsorber," *Opt. Lett.,* vol. 34, no. 5, pp. 644-646, 2009.
[98] T. Yang, H. Chen, X. Luo, and H. Ma, "Superscatterer: Enhancement of scattering with complementary media," *Opt. Express,* vol. 16, no. 22, pp. 18545-18550, 2008.
[99] A. Alù, and N. Engheta, "Cloaking a sensor," *Phys. Rev. Lett.,* vol. 102, no. 23, 2009.
[100] W. L. Barnes, A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature,* vol. 424, no. 6950, pp. 824-830, 2003.
[101] S. V. Boriskina, H. Ghasemi, and G. Chen, "Plasmonic materials for energy: From physics to applications," *Materials Today,* vol. 16, no. 10, pp. 375-386, 2013.
[102] J. A. Schuller, E. S. Barnard, W. Cai, Y. C. Jun, J. S. White, and M. L. Brongersma, "Plasmonics for extreme light concentration and manipulation," *Nature Mater.,* vol. 9, no. 3, pp. 193-204, 2010.
[103] Z. Jakšić, O. Jakšić, and J. Matovic, "Performance limits to the operation of nanoplasmonic chemical sensors – noise equivalent refractive index and detectivity," *J. Nanophotonics,* vol. 3, pp. 031770, 2009.
[104] J. Brown, "Artificial dielectrics," *Progress in Dielectrics*, J. B. Birks, ed., pp. 193–225, Hoboken, New Jersey: Wiley, 1960.
[105] S. M. Vuković, Z. Jakšić, I. V. Shadrivov, and Y. S. Kivshar, "Plasmonic crystal waveguides " *Appl. Phys. A,* vol. 103, no. 3, pp. 615-617, 2011.
[106] C. Bauer, G. Kobiela, and H. Giessen, "2D quasiperiodic plasmonic crystals," *Scientific reports,* vol. 2, pp. 0681.1-6, 2012.
[107] M. Maksimović, and Z. Jakšić, "Emittance and absorptance tailoring by negative refractive index metamaterial-based Cantor multilayers," *J. Opt. A-Pure Appl. Opt.,* vol. 8, no. 3, pp. 355-362, 2006.
[108] K. Vynck, M. Burresi, F. Riboli, and D. S. Wiersma, "Photon management in two-dimensional disordered media," *Nature Mater.,* vol. 11, no. 12, pp. 1017-1022, 2012.
[109] S. M. Vuković, Z. Jakšić, and J. Matovic, "Plasmon modes on laminated nanomembrane-based waveguides," *J. Nanophotonics,* vol. 4, pp. 041770, 2010.
[110] P. Ginzburg, N. Berkovitch, A. Nevet, I. Shor, and M. Orenstein, "Resonances on-demand for plasmonic nano-particles," *Nano Lett.,* vol. 11, no. 6, pp. 2329-2333, 2011.
[111] J. Zhang, T. Atay, and A. V. Nurmikko, "Optical detection of brain cell activity using plasmonic gold nanoparticles," *Nano Lett.,* vol. 9, no. 2, pp. 519-524, 2009.
[112] M. Quinten, *Optical Properties of Nanoparticle Systems: Mie and Beyond*, Wiley-VCH, Weinheim, Germany, 2011.
[113] T. K. Sau, and A. L. Rogach, eds., *Complex-shaped Metal Nanoparticles: Bottom-Up Syntheses and Applications*, Weinheim, Germany: Wiley-VCH, 2012.
[114] P. Berini, "Long-range surface plasmon polaritons," *Adv. Opt. Photon.,* vol. 1, no. 3, pp. 484-588, 2009.
[115] P. Berini, R. Charbonneau, and N. Lahoud, "Long-range surface plasmons along membrane-supported metal stripes," *IEEE J. Sel. Top. Quant. Electr.,* vol. 14, no. 6, pp. 1479-1495, 2008.
[116] Z. Jakšić, and J. Matovic, "Functionalization of Artificial Freestanding Composite Nanomembranes," *Materials,* vol. 3, no. 1, pp. 165-200, 2010.
[117] Z. Jakšić, and J. Matović, "Nanomembrane-Enabled MEMS Sensors: Case of Plasmonic Devices for Chemical and Biological Sensing," *Micro Electronic and Mechanical Systems*, K. Takahata, ed., pp. 85-107, Vienna: In-Tech, 2009.
[118] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays," *Nature,* vol. 391, no. 6668, pp. 667-669, 1998.
[119] A. Alu, and N. Engheta, "Theory, modeling and features of optical nanoantennas," *IEEE T. Antenn. Propag.,* vol. 61, no. 4, pp. 1508-1517, 2013.
[120] P. Biagioni, J.-S. Huang, and B. Hecht, "Nanoantennas for visible and infrared radiation," *Reports on Progress in Physics,* vol. 75, no. 2, pp. 024402, 2012.
[121] N. Liu, M. L. Tang, M. Hentschel, H. Giessen, and A. P. Alivisatos, "Nanoantenna-enhanced gas sensing in a single tailored nanofocus," *Nature Mater.,* vol. 10, no. 8, pp. 631-636, 2011.
[122] X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Broadband light bending with plasmonic nanoantennas," *Science,* vol. 335, no. 6067, pp. 427, 2012.

Parametric SPICE Model for Static Induction Transistor (SIT) in triode mode of operation

Goce L. Arsov
Faculty of Electrical Engineering and Information Technologies
SS Cyril and Methodius University - Skopje
Skopje, Macedonia
g.arsov@ieee.org

***Abstract*—The static induction transistor (SIT) is a device which belongs to the multichannel power FET family. Depending on its internal source resistance it can achieve triode or pentode like output characteristics. The purpose of this work is to present a simple model suitable for computer analysis and simulation of circuits containing SITs in triode-like mode of operation. The model is based on the SITs static and dynamic behaviour, and not on the physical structure and characteristics of the device. A possible method for extraction of the model parameters is proposed. The simulated static and frequency characteristics correspond well to the experimental results available from the references. The model has passed all performed tests. Although the proposed model is not yet an ideal solution it will help to start simulating circuits containing the static induction transistors for different applications.**

Keywords—Static induction transistor (SIT); static characteristics; SPICE; modeling; parameter extraction

I. INTRODUCTION

The basic principles of the field-effect transistor (FET) were described by J. E. Lilienfeld [1] - [3], while the theoretical analysis was performed by W. Shockley in 1952 [4]. In early 1950's there were two main goals which occupied the scientists dealing with further development of the FET. The first one was a creation of semiconductor device with triode-like *I-V* characteristics and the other was the development of high power device based on field effect principles – power FET. In 1950 Y. Watanabe and J. I. Nishizawa, investigating the possibilities of creating semiconductor device with triode like characteristics have developed the concept of vertical "analog" transistor [5], while in 1952 W. Shockley proposed the, so-called, "analog transistor" [6]. It was shown that the saturation region in the FET's *I-V* characteristics is a result of the internal negative feedback caused by the series channel resistance (r_s). As a result the FET's overall transconductance (G_m') changes to value [7]

$$G_m' = G_m/(1+ r_s \cdot G_m). \quad (1)$$

For $r_s \cdot G_m >> 1$ (1) reduces to:

$$G_m' \approx 1/r_s \quad (2)$$

On the other hand, if $r_s \cdot G_m << 1$ in the entire region, (1) becomes

$$G_m' = G_m, \quad (3)$$

and the device will not exhibit saturation. Instead, its characteristics will be triode-like.

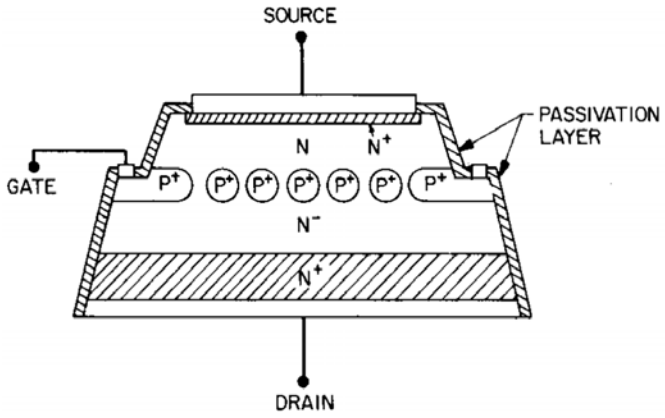
The research considering the development of power FETs resulted in similar conclusions. Namely, any power semiconductor device should have large current carrying capabilities, low power losses in conducting mode of operation, and high blocking voltage in non conducting mode. These imply very short and as wide as possible device channel. The first attempt in creating power FET was the "cylindrical field-effect transistor", proposed by H. A. R. Wegener [8] in 1959. It was clear, from the beginning, that the width of the channel cannot be increased, without losing part of the wanted device characteristics, beyond certain critical value. The main problem was the control of the channel width for very short channel device. Possible solutions of this problem were proposed by R. Zuleeg [9] – [12] and his "Multichannel field–effect transistor", as well as by S. Teszner and R. Gicquel [13] and their "Gridistor". Both components are based on the vertical multichannel structure where many vertical channels are connected in parallel forming "parallel multichannel FET". Due to the very short channel these devices could operate, depending on their gate-to-source voltage, in both: triode-like and pentode-like mode of operation. Unfortunately, these devices did not find any commercial applications.

II. THE STATIC INDUCTION TRANSISTOR (SIT)

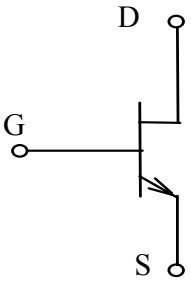
Knowing the pros and contras of the multichannel FETs and gridistors, in 1975 J. Nishizawa and his associates have proposed a device based on previous two. As it was made clear that the main mechanism of operation of this device is based on the static induction it was named "static induction transistor" – SIT [14] – [16]. According to its output *I-V* characteristics SIT can be considered as a solid-state analogy

of the vacuum tube, which, at certain predefined pinch-off voltage, can be operated in triode-like and/or pentode-like mode of operation. Basically, SIT can be considered as a multichannel structure (Fig. 1) which, when no voltage is applied to the gate, is in the on-state.

The main characteristics of the SIT, in triode mode of operation, are: very short channel length, low internal series resistance, low input (gate to source) capacitance, low noise, low harmonic distortions, and possibility to control high power at audio-frequencies. The transient on/off time is very short (typically around 250 ns). SIT exhibits relatively high voltage drop in forward conducting state (90 V for 180 A device, and 18 V for 18 A device). It's current and voltage ratings can exceed 300 A and 1200 V, respectively. SIT's operating frequency can be higher than 100 kHz which makes it very suitable for high-frequency high-power applications (audio, VHF, UHF and microwave amplifiers). The very first commercial SIT's were fabricated by *Tokin Corporation* of Sendui, Japan [17]. Today, SIT became very popular for audio applications as building devices in A-class power audio-amplifiers. Namely, the triode-like SIT becomes an ideal replacement for the vacuum triode, commonly used in these applications [18]. Nowadays, several companies already produce SIT-based audio-amplifiers.



(a)



(b)

Figure 1. Basic structure of SIT and device symbol: a) basic structure [18]; b) device symbol

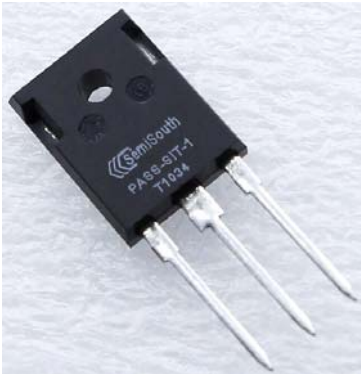


Figure 2. SIT manufactured by SemiSouth [19]

One of them is *Digital Do Main* from Japan, which uses SIT's fabricated by *Yamaha Silicon*. Another one is *First Watt*, that fabricate audio-amplifiers based on SiC SIT's produced by *SemiSouth Company* as an "application specific device" for this purpose (Fig. 2) [19].

The typical SIT's (triode-like and pentode-like) *I-V* output characteristics are presented in Fig. 3.

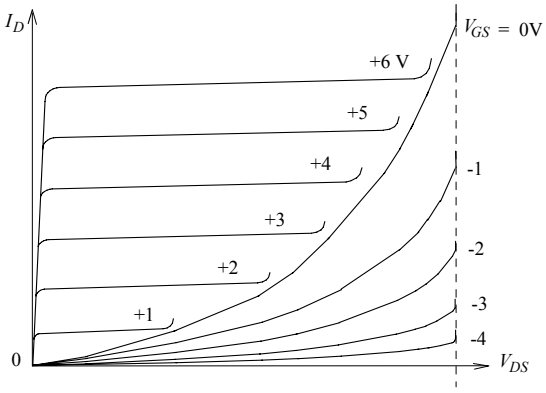


Figure 3. Typical SIT output *I-V* characteristics

III. SIT MODELING

Electronic circuit design requires accurate methods of evaluating circuit performance. Today, because of the enormous complexity of modern electronic circuits, in this process, the computer aided circuit analysis and simulation are essential and can provide information about circuit performance that is almost impossible to obtain with laboratory prototype measurements. Since the appearance of the SIT numerous attempts have been made to develop appropriate device models suitable for theoretical, but also computer aided, analysis and simulation of the SIT based systems [20] − [28]. Some of these models are pure analytical and are based on the physical laws governing the SIT behaviour. However, only few of them are adapted to computer programs for analysis and simulation [22] (PSpice), [25] (Saber). Unfortunately, all of these models are based on pentode-like (bipolar) SIT mode of operation. In lack of triode-like SIT models for circuit analysis and simulation and

having into consideration the increasing interest for this device in designing the output stages of audio-amplifiers [19], a simple parametric model suitable for Spice based simulations is presented below.

A. Spice Models for Vacuum Triode

The first approach was to modify the Spice models for vacuum triode based on the well-known Child-Langmuir law (known also as space-charge conduction law). The basic model was proposed in [29]. As it was not applicable for real triodes some improvements were presented in [30]. The improved model [30] is based on the behaviour of the vacuum triode and is not supported by its physical laws. Considering the fact that the SIT's output characteristics are based on static induction and that they follow the exponential law rather than the space-charge conduction law, the final conclusion was that these approach cannot be easily modified into triode-like SIT model. On the other hand, when trying to use the derived formulas [14], [15] and [31] in SIT modeling the main problem to solve was the model parameter extraction, as these values are dependent of the physical parameters and/or behaviour of the device. Therefore other possibilities had to be investigated.

B. Spice Model for SIT in triode-like mode of operation

In 2013 the simple model for triode-like N-channel SIT was proposed [32]. The model is quite simple and is based on the results in modeling the static induction thyristor [33]. The model was developed using the analog behavioral modeling, available in any modern version of Spice, and is based on SIT's static and dynamic behaviour, rather than on its physical structure.

The schematic structure of this model for an N-channel device is presented in Fig. 4.

The model is consisted of three voltage-controlled voltage sources (VCVS), two diodes, two resistors and one capacitor. Additionally three capacitors (C_{GS}, C_{GD} and C_{DS} – not shown in Fig. 4) can be added for better modeling of the SIT's frequency response.

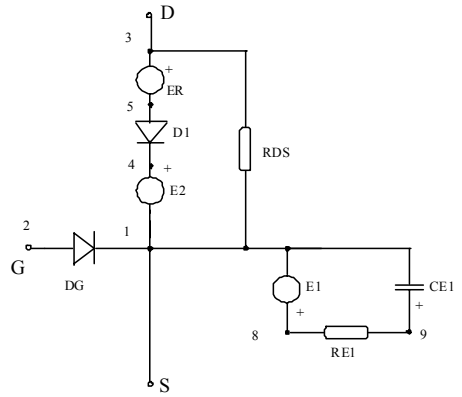


Figure 4. N-channel Spice model for triode-like SIT

Figure 5 presents the subcircuit definition for the model from Fig. 4. The **.param** function allows easy entering necessary model parameters.

```
.param  p=p q=q rs=rs n=n is=is re1=re1 ce1=ce1
.subckt sit     3     2     1
*-------------- D     G     S
er  3 5 value = {v(4,1)*v(5,4)}
d1 5 4 d1
.model d1 d(rs= rs is= is n= n)
e2 1 4 9 1 1
rds 3 1 24meg
dg 2 1 d2
.model d2 d(rs=.3 is=1e-6)
e1 8 1 value = {((v(2,1)* p)*(( q –v(2,1)))}
re1 8 9 re1
ce1 9 1 ce1
.ends
```

Figure 5. Subcircuit definition for the SIT model from Fig. 4

The diode D_G is used to simulate the *pn* junction nature of the gate. While simulating the triode-like mode of SIT operation this diode should be reverse biased. The main current path is composed of two VCVS (E_R and E_2) and the diode D_1.

The VCVS E_R is used for modeling the shape of the SIT's output *I-V* characteristics. Its voltage is actually influenced by the voltages across the VCVS E_2 and the diode D_1.

Besides, the VCVS E_2 serves to define and implement the knee of any single curve in the output characteristics. Namely, although this characteristic value depends on several factors it can be estimated as:

$$V_{B,DS} \approx |V_{GS}| \cdot p \cdot \left(q + |V_{GS}|\right) \tag{4}$$

where p and q are parameters which can be obtained by simple measurements of two curves in the family of the output characteristics for the value of the current of 5mA.

In this model the diode D_1 has several tasks: (1) it provides the forward conducting and the reverse blocking characteristics of the device; (2) by adjusting the parameters **rs** (internal diode resistance), **is** (reverse saturation current) and **n** (diode emission coefficient) in the Spice diode model, the internal resistance of the device and the slope of the output characteristics can be modeled. In [32] the curve fitting was performed by trial and error and after several iterations the values were satisfactory determined. But, knowing that, in the proposed model, the drain to source voltage is defined by E_R, D_1 and E_2 as:

$$\begin{aligned} V_{DS} &= V_{GS} \cdot p \cdot \left(q - V_{GS}\right) + \\ &+ \left(\frac{T}{11605} \cdot n \cdot \ln\frac{I_D}{I_s} + r_s \cdot I_D\right) \cdot \left[1 + V_{GS} \cdot p \cdot \left(q - V_{GS}\right)\right] \end{aligned} \tag{5}$$

which, for V_{GS}=0 (5) becomes:

$$V_{DS} = \frac{T}{11605} \cdot n \cdot \ln\frac{I_D}{I_s} + r_s \cdot I_D \tag{6}$$

It is easy to calculate n as:

$$n = \frac{V_{DS} - r_s \cdot I_D}{T/11605} \cdot (\ln \frac{I_D}{I_s})^{-1} \quad . \tag{7}$$

For all parameter calculations, the working temperature of the device was assumed to be 350 K.

The value of the diode internal resistance r_s, in ohms, can be estimated as:

$$r_s \cong \sqrt{\frac{\Delta V_{DS}}{\Delta I_D}} \tag{8}$$

In all simulations the value of the diode reverse saturation current was chosen to be 10^{-6}A.

The circuit consisted of VCVS E_1, resistor R_{E1} and the capacitor C_{E1} is, essentially, used to model the device AC characteristics. Additionally, in this case, it is used also to adapt the gate control function before including its influence into the main current path through VCVS E_R. It should be noted that during the AC analysis the value of R_{E1} is to be equal to the value of the load drain resistance in the analyzed circuit.

Although the explained model is valid for the N-channel SIT, its modification for simulating P-channel devices should not be a problem.

IV. Simulation Results

Extensive simulations have been performed in order to test the proposed model characteristics. The model has passed all performed tests and can be used to analyze various types of power converters using SI transistors. In lack of real device the simulation results are compared with the measured results for real SITs, presented in [14], [15], [19] and [31]. All simulations were performed using Spice versions produced by two different companies: (1) Pspice Lite v16.6 - Cadence Design Systems Inc. [34] and (2) LTspice v4.12t - Linear Technology Corporation [35]. Only few of the obtained results are presented below.

A. Simulation Results - DC Mode of Operation

Several types of SIT were modeled and their models were tested using both of the above mentioned Spice programs. For each example of SIT its model parameters were extracted from its measured characteristics using equations (4), (7) and (8). In Fig. 6 the measured [14] and simulated I_{DS}-V_{DS} static characteristics (using the model presented in Fig. 4), for 1000 V SIT (fabricated by Tokin Corporation), are presented. Although there is a visible discrepancy between the curves for V_{gs}=0, which is caused by the relatively high value of the parameter **rs** in the diode model, it can be seen that for all other values of V_{gs} the simulated results correspond well with the measured ones.

The simulated I_{DS}-V_{DS} static characteristics, for the 40 V SIT (2SK76) produced by Tokin Corporation, are presented in Fig. 7-b. Model parameters were extracted from the measured characteristics given in Fig. 7-a [15], using the method described above. Again, the simulated results correspond well to the measured ones.

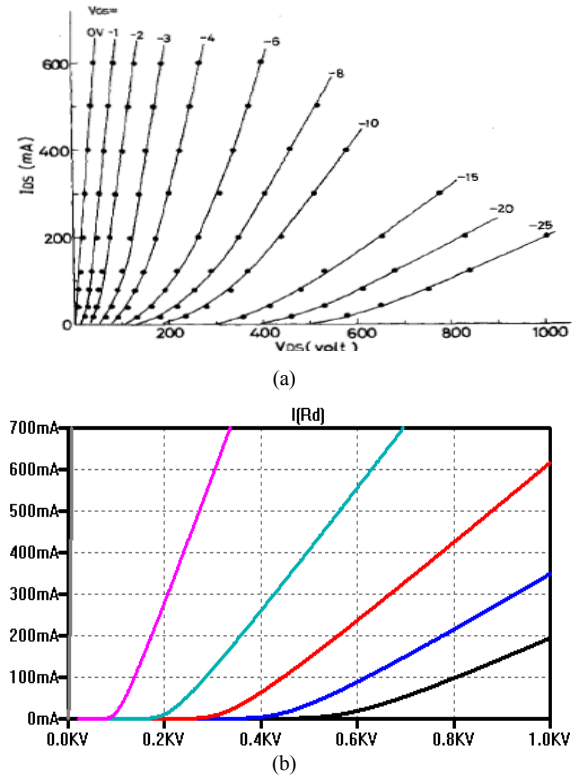


Figure 6. 1000 V SIT (Tokin Corp.) I_{DS}-V_{DS} characteristics: a) measured characteristics [14]; b) simulated characteristics using model from Fig. 4 with p=0.0714, q=100, rs=8Ω, n=6, is=10^{-6}A (V_{DS} is on the horisontal axis; I_D is on the vertical axis; V_{GS} was changed from 0 to -25 V in steps of -5 V.)

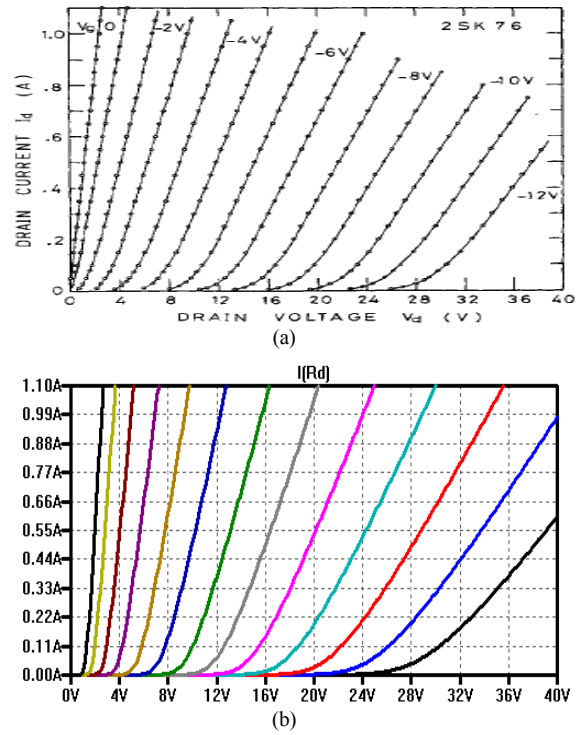


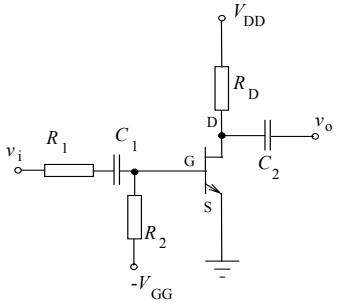
Figure 7. 40 V SIT (2SK76) I_{DS}-V_{DS} characteristics: a) measured characteristics [15]; b) simulated characteristics using model from Fig. 4 with p=0.0694, q=3, rs=1.1Ω, n=4, is=10^{-6}A (V_{DS} is on the horisontal axis; I_D is on the vertical axis; V_{GS} was changed from 0 to -12 V in steps of -1 V.)

B. Simulation Results – Frequency Response

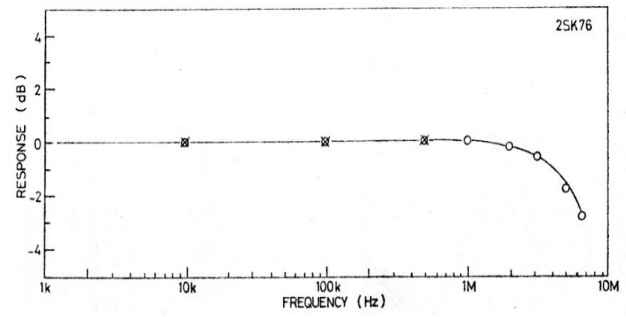
It has already been said that one of today's main applications of triode-like SIT is in building linear audio-amplifiers. Therefore, the frequency response (small signal operation) of the proposed model has, also, been tested. The simulation of the frequency response, for the improved 2SK76 SIT [31], was performed using the same circuit topology (Fig. 8-a) as in [31] with the load resistance of 8 ohms. Equation (9) is used to adjust the high frequency amplitude response of the model to that of the practical device, shown in Fig. 8-b [31]. In (9) R_{E1} should have the same value as R_D so the only unknown remains C_{E1}.

$$f_H = 1/(2\pi R_{E1} C_{E1}) \quad (9)$$

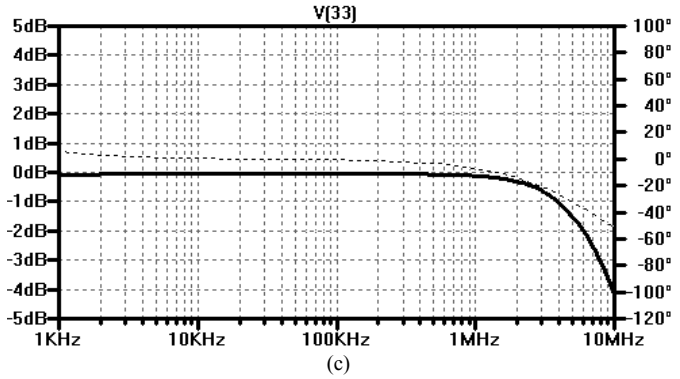
The results are shown in Fig. 8-c. The upper cut-off frequency of 8 MHz has been obtained for C_{E1} =2.5 nF. It can be seen that the simulated curve correspond well to the measured one.



(a)



(b)



(c)

Figure 8. SIT Frequency amplitude response (improved 2SK76): (a) circuit; (b) measured amplitude response from [31]; (c) simulated response using: C_{E1} =2.5 nF, $R_{E1} = R_D$=8Ω, p=0.0694, q=3, rs=1.1Ω, n=4, is=10^{-6}A. (V(33)≡v_o – the solid line shows the amplitude response while the dashed line shows the phase response)

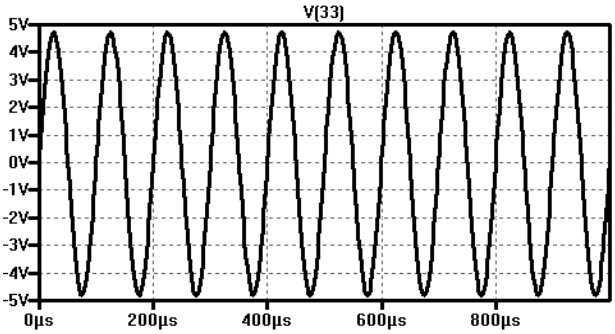


Figure 9. The output voltage (v_o) waveform – the magnitude of the input sinewave voltage is chosen so that the delivered power to the 8Ω load is 1.5W. The SIT model parameters are (100 V SIT): p=0.0714, q=100, rs=8Ω, n=6, is=10^{-6}A (the signal frequency is f_i = 10kHz, while R_D=8Ω, C_1 = 10µF, C_2 = 100µF)

To examine the linearity of the model transfer characteristics the common source configuration from Fig. 8-a was used. The AC voltage source was connected to the gate of the device through the 10 µF capacitor. The load (R_L = 8 Ω) was connected to the drain of the SIT through 100 µF capacitor. The magnitude of the input sine-wave signal was chosen so that the power of 1.5 W was delivered to the load [19]. The simulated output voltage is shown in Fig. 9. The performed Fourier analysis showed that the distortions of the output signal are lower than 1%, which is in full compliance with the data given in [19]. The analysis shows clearly that these distortions are mainly caused by the second order harmonic, while the amplitudes of the higher order harmonics were lower than 0.01% in comparison with the amplitude of the fundamental.

V. CONCLUSIONS

The Spice model for the static induction transistor is presented. The model is based on the device behaviour rather than on the physical structure of the SIT. Simple method for model parameters extraction from the measured characteristics of the real SIT has been developed. Extensive simulations have been performed to examine the model characteristics. The model has passed all performed tests and the results correspond very well with the experimental ones given in [14], [15], [19] and [31]. Although the proposed model is not an ideal solution it will help to start simulating various circuits containing static induction transistors. The first problem to solve is the discrepancy between the measured and simulated curves for V_{gs} = 0, when simulating the high voltage devices, where the value of the internal diode resistance **rs** is relatively high. One possible solution is to add a small DC bias at the gate input of the model. This problem will be investigated in details in developing an advanced SIT model.

At least two other things are to be done in the near future: (1) to develop an algorithm for self extracting model parameters from devices data sheets; (2) to extend the model to covering both mode of SIT operation (triode and pentode-like) in order to enable full implementation of SIT for the Spice simulation program, or even to other circuit simulation program packages.

ACKNOWLEDGMENT

The author would like to thank the reviewers of this paper for their very constructive and useful comments and suggestions.

REFERENCES

[1] J. E. Lilienfeld, "Method and apparatus for controlling electric currents", *US Patent 1745175*, Jan. 18, 1930

[2] J. E. Lilienfeld, "Device for controlling electric current", *US Patent 1900018*, Mar. 07, 1933

[3] J. E. Lilienfeld, "Amplifier for electric currents", *US Patent 1877140*, Sept. 13, 1932

[4] W. Shockley, "A unipolar 'field-effect' transistor", *Proc IRE*, Vol. 40, 11, 1952, pp. 1365-1376

[5] Y. Watanabe, J. Nishizawa, *Japanese Patent 205068*, No. 28-6077, application date Dec 1950

[6] W. Shockley, "Transistor electronics: Imperfections, unipolar and analog transistors" *Proc. IRE*, Vol. 40 , No 11, 1952, pp. 1289 – 1313

[7] Y. Mochida, J. Nishizawa, T. Ohmi, R.K. Gupta, "Characteristics of static induction transistors: Effects of series resistance", *IEEE Trans. Electron Devices,* Vol. 25 , No. 10, 1978, pp. 761 – 767

[8] H. A. R. Wegener, "The cylindrical field-effect transistor", *IEEE Trans. Electron Devices*, 6, 1959, pp 442–449

[9] R. Zuleeg, V.O. Hinkle, "A multichannel field–effect transistor", *Proc. IEEE*, Vol. 52, No. 10, 1964, pp. 1245 – 1246

[10] R. Zuleeg, "Multichannel field–effect transistor; theory and experiment", *Solid-State Electronics*, Vol. 10, 1967, pp. 559-576

[11] R. Zuleeg, "A silicon, planar space-charge-limited current triode", *International Electron Devices Meeting – Book of Abstracts*, 1965 Vol. 11, 1965, pp 63

[12] R. Zuleeg, "A silicon space-charge-limited triode and analog transistor", *Solid-State Electronics*, Vol. 10, 1967, pp. 449-460

[13] S. Teszner, R. Gicquel, "Gridistor—A new field-effect device" *Proc. IEEE*, Vol. 52 , No 12, 1964, pp 1502 – 1513

[14] J. Nishizawa, T. Terasaki, J. Shibata, "Field-effect transistor versus analog transistor (static induction transistor)" *IEEE Trans. Electron Devices,* Vol. 22 , No. 4, 1975, pp. 185 – 197

[15] Y. Mochida, J. Nishizawa, T. Ohmi, R.K. Gupta, "Characteristics of static induction transistors: Effects of series resistance", *IEEE Trans. Electron Devices,* Vol. 25 , No. 10, 1978, pp. 761 – 767 [7]

[16] T. Ohmi, J. Nishizawa, M. Tatsuta, "High power static induction transistor", *Proc. Int. Electron Devices Meeting*, 1978, pp. 703 – 704

[17] SIT Handbook, Tohoku Metal Ind., 1987

[18] B. K. Bose, "Evaluation of modern power semiconductor devices and future trends of converters", *IEEE Trans on Ind. Appl*., Vol. 28, No. 2, 1992, pp. 403-413

[19] N. Pass, "Introduction to static induction transistors", *First Watt,* http://www.firstwatt.com/sitintro.html (2011)

[20] A.G.M. Strollo, P. Spirito, "A two-dimensional analytical model for the output I-V characteristics of the static induction transistor (SIT)", *Proc. 2nd Int. Symp. on Power Semiconductor Devices - ISPSD '90*, 1990. pp 196 - 203

[21] A.M. Ionescu, A. Rusu, C. Postolache," Bipolar static induction transistor (BSIT) static model", *Proc. 6th MELECON*, Vol.1, 1991, pp. 107 – 110

[22] G. Busatto, "Physical Modeling of Bipolar Mode JFET for CAE/CAD Simulation", *IEEE Trans. Power Electronics*, Vol. 8, No. 4, 1993, pp 368-375

[23] A.G.M. Strollo, "Modeling pentode-like characteristics of recessed-gate static induction transistor" IEEE Trans. Electron Devices, Vol. 41, No 4, 1994 , pp 616 - 617

[24] G. E. Bunea, S.T. Dunham and T.D. Moustakas, "Modeling of a GaN based static induction transistor" *MRS Proceedings*, Vol. 537, 1998, http://journals.cambridge.org/article_S1946427400288891

[25] A. S. Kashyap, et al., "Compact Circuit Simulation Model of Silicon Carbide Static Induction and Junction Field Effect Transistors", *Proc. IEEE Workshop on Computers in Power Electronics*, 2004, pp 29-35

[26] E. Platania, et al., "A Physics-Based Model for a SiC JFET Device Accounting for the Mobility Dependence on Temperature and Electric Field" , *IEEE IAS Annual Meeting*, 2008, pp 1-8

[27] Z. Chen et al. "Model for Power SiC Vertical JFET With Unified Description of Linear and Saturation Operating Regimes" *IEEE Trans on Ind. Appl*., Vol. 47, No. 4, July/August 2011

[28] M. J. Kumar and H. Bahl, "New Schottky-gate Bipolar Mode Field Effect Transistor (SBMFET): Design and Analysis using Two-dimensional Simulation," *IEEE Trans. on Electron Devices*, Vol.53, pp.2364-2369, September 2006.

[29] S. Reynolds, "Vacuum Tube Models for PSPICE Simulations," *Glass Audio*, 4/93, p. 17.

[30] N. L. Koren, "Improved Vacuum Tube Models for Spice Simulations," *Glass Audio*, 5/96, p. 18.

[31] J. Nishizawa, "Static induction transistor", *Semiconductor technologies*, 1982, pp 201-219

[32] G. L. Arsov, "SPICE Model for Static Induction Transistor (SIT), Part I: SIT in triode mode of operation", Proc. ETAI 2013, Ohrid, Macedonia, 26-28.09.2013 CDROM (in Macedonian),

[33] G. L. Arsov, "A Pspice model for the static induction thyristor", *Proc. 10th Int. Symposium on Power Electronics - EE99,* 1999, T1-5.1, pp. 1-5

[34] http://www.cadence.com/products/orcad

[35] www.linear.com/**ltspice**

On the measurement methods for dielectric constant determination in Nb/$BaTiO_3$ ceramics

Miloš Marjanović, Vesna Paunović, Zoran Prijić, Aneta Prijić, Danijel Danković, Vojislav Mitić
Department of Microelectronics
University of Niš, Faculty of Electronic Engineering
Niš, Serbia
milos.marjanovic@elfak.ni.ac.rs

***Abstract*—In this paper the measurement methods for dielectric constant determination in Nb/$BaTiO_3$ ceramics have been presented. The experimental results obtained using LCR meter with dielectric test fixture and the influence of ceramic microstructure on the accuracy of methods have been discussed.**

Keywords-ceramics; dielectric constant; measurement methods

I. INTRODUCTION

For accurate measurement of the dielectric constant of a ceramics, a three-terminal system is needed [1]-[4]. The parallel plate measurement methods, also called the three-terminal methods in ASTM D150 [5], involve the use of a parallel plate capacitor as a holder, with sample of the material under the test (MUT) sandwiched between the plates. These methods require an impedance analyzer or LCR meter. The measurements are at low frequencies. Typically, the methods use a wide frequency range from 20 Hz to 1 GHz. The MUT is biased by an AC source and the actual voltage across the material is monitored. The material test parameters, such as dielectric constant, are calculated by knowing the dimensions of the MUT and by measuring its capacitance.

This paper presents experimental results of capacitance measurement and dielectric constant determination of Nb/$BaTiO_3$ ceramics over a wide frequency range. The results are obtained with two methods: contacting electrode and non-contacting electrode method. The dielectric constant values in function of frequency are considered. The influence of samples microstructure on the accuracy of the methods is also discussed in details.

II. PARALLEL PLATE MEASUREMENT METHODS

There are two types of parallel plate measurement methods: contacting electrode method, and non-contacting electrode method. The least accurate measurements are by contacting electrode method without thin film electrode. When thin film electrode is applied onto surfaces of the MUT, the highest accuracy is possible. Therefore, contacting method is the most widely used one [6]-[7]. The main advantage of contacting method is that for obtaining the dielectric constant it is necessary to take only one measurement. On the other hand, the non-contacting electrode method has medium measurement accuracy and involves very simple preparation of sample and setup. MUT should be solid material with a flat and smooth surfaces, but dielectric constant is derived by using the results of two capacitance measurements.

When measuring the dielectric material between two electrodes, stray capacitance or edge capacitance is formed on the edges of the electrodes and consequently the measured capacitance is larger than the capacitance of the MUT. The edge capacitance causes a measurement error, since the current flows through the dielectric material and edge capacitor [8]-[9]. A solution to the measurement error is to use the guard electrode which absorbs the electric field at the edge and the capacitance that is measured between the electrodes is only composed of the current that flows through the dielectric material (Fig. 1). When the main electrode is used with a guard electrode, the main electrode is called the guarded electrode.

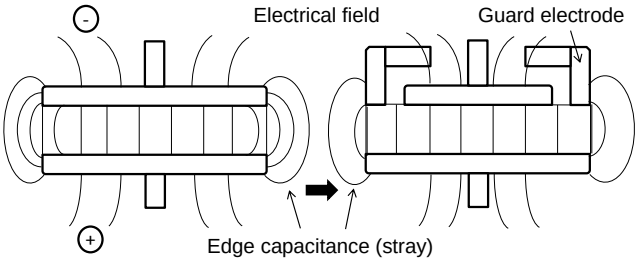


Figure 1. Illustration of guard electrode effect

A. Contacting electrode method

The contacting electrode method determines permittivity by measuring the capacitance of the MUT directly (Fig. 2). The dielectric constant is calculated using the equation:

$$\varepsilon_r = \frac{t \cdot C_p}{S \cdot \varepsilon_0} = \frac{t \cdot C_p}{\pi\left(\frac{d}{2}\right)^2 \cdot \varepsilon_0}, \quad (1)$$

where t is average thickness of MUT, C_p is equivalent parallel capacitance of MUT, S is guarded electrode's surface area, d is diameter, ε_0 is permittivity of vacuum. The contacting electrode method doesn't require material preparation and the operation involved when measuring is simple [7]. When contacting the MUT directly with the electrodes, an airgap is formed between the MUT and the electrodes. Materials with rough surfaces can be affected by airgap, as illustrated in Fig. 3. A measurement error can occur because the measured capacitance will be the series connection of the capacitance of the dielectric material

The authors gratefully acknowledge the financial support of Serbian Ministry of Education, Science and Technological Development. This research is a part of the Project OI 172057 and of the Project TR 32026.

and the airgap. Thin samples are most prone to airgap effect. The methods for reducing error due to airgap are formation of thin film electrodes on a dielectric material and maximization of the pressure on the test fixture to the extent that it doesn't deform the MUT. There is a technique to apply a thin film electrode onto the surfaces of the dielectric material in order to increase the contacting area and to eliminate the airgap that occurs between the MUT and the electrodes. An electrode is manufactured onto the dielectric material using high-conductivity silver paste. Thin film electrode must be thinner than the dielectric material. In this case, it is important to appropriately position the fabricated thin film electrode onto the MUT, to precisely contact the electrodes of test fixture. The measurement error is a function of the relative permittivity (ε_r) of the MUT, thickness of the MUT (t_m), and the airgap's thickness (t_a), and it is determined by:

$$\delta_{r\varepsilon} = \frac{\varepsilon_r - 1}{\varepsilon_r + \frac{t_m}{t_a}} \cdot 100\ [\%]. \quad (2)$$

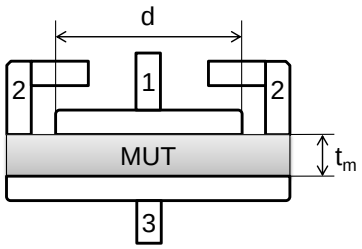


Figure 2. Illustration of contacting electrode method (1 – guarded electrode, 2 – guard electrode, 3 – unguarded electrode)

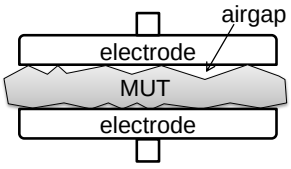


Figure 3. Illustration of rough-surfaced MUT

B. Non-contacting electrode method

The non-contacting electrode method does not require thin film electrodes, but still overcomes the airgap effect, so this method was conceptualized to unite the advantages and eliminate the disadvantages of the contacting electrode method [7]. Two capacitance measurements, obtained with the MUT and without it, are necessary and the results are used to derive dielectric constant as:

$$\varepsilon_r = \frac{1}{1 - \left(1 - \frac{C_1}{C_2}\right) \cdot \frac{t_g}{t_m}}, \quad (3)$$

where C_1 is capacitance without MUT inserted, C_2 is capacitance with MUT inserted, t_g is electrode gap, and t_m is average thickness of MUT. The gap between guarded/guard electrode and MUT (t_g-t_m) should be very small when compared to the thickness of the MUT (t_m) (Fig. 4).

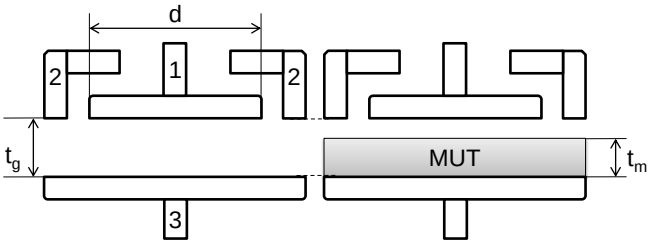


Figure 4. Illustration of non-contacting electrode method (1 – guarded electrode, 2 – guard electrode, 3 – unguarded electrode)

III. SAMPLE PREPARATION AND EXPERIMENTAL PROCEDURE

The samples of modified $BaTiO_3$ ceramics doped with 0.5, 1.0, 2.0 mol % Nb_2O_5 were examined. The samples were prepared by a conventional solid state sintering procedure starting from reagent grade Nb_2O_5 and $BaTiO_3$ powder. Starting powders were ball milled in ethyl alcohol for 24 h. After drying at 200°C for several hours, the powders were pressed into disks under 120 MPa. The pellets were sintered at 1320°C in air atmosphere for 2 h [10]. The microstructure of the sintered samples were examined by scanning electron microscope (SEM) JOEL-JSM 5300 equipped with EDS (QX 2000S) system. Thin samples of 11 mm diameter and nearly 1.7 mm thickness were sandwiched using uniformly coated silver paste on both sides so that a ceramic capacitor structure is formed. The calculation of dielectric constant (permittivity) was performed by using measured values of capacitance, samples thickness and electrode area. The dielectric properties of the samples are measured in the frequency range from 20 Hz to 1 MHz.

The measurement system that employ the parallel plate methods consists of LCR meter Agilent 4284A [11] with the 16451B dielectric test fixture [12], which has capabilities to measure solid materials. Four electrodes are implemented to accommodate the contacting and non-contacting electrode methods and various MUT sizes. The three-terminal configuration with guard electrode eliminates the effect of edge capacitance and prevents occurrence of larger capacitance value than it really is. First electrode is unguarded electrode which is fed to the measurement instrument's high terminal, and the low terminal is connected with guarded electrode. Guard electrode is connected to the outer conductor of the BNC connector and encompasses the main electrode so that absorbs the electric field at the edge of the electrodes. When 16451B test fixture is used, measurements are possible in the frequency range up to 30 MHz, with maximum DC voltage of ±42 V. This test fixture has wide operation temperature range (from 0°C to 55°C). The limitation is cable length (1m). Since the 16451B test fixture introduces errors due to electrical length, residual impedance, and stray admittance, these errors can be entirely removed by open, short and load compensation. Measurement procedure flow chart with test fixture is shown in Fig. 5. The MUT must be a solid sample that is smooth and has equal thickness from one end to the other end.

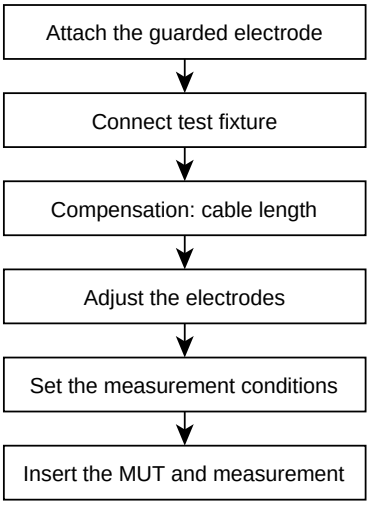


Figure 5. Basic flow chart for measurement with 16451B test fixture

IV. RESULTS AND DISCUSSION

$Nb/BaTiO_3$ samples are stable and don't transform under applied pressure. Because of that, samples are suitable for both methods: contacting and non-contacting. The SEM imaging showed that the samples of Nb doped $BaTiO_3$ ceramics have spherical shaped grains (not flat surface). For 0.5 mol% of $Nb/BaTiO_3$ samples with fairly uniform microstructure, the average grain size is from 0.5 µm to 1 µm (Fig. 6). With increasing of dopant concentration, the grain size of $Nb/BaTiO_3$ samples increases and for the samples doped with 2 mol% of Nb the grain sized is around 7 µm, dispersed in a fine-grained matrix with average grain size in the range from 0.5 µm to 2 µm (Fig. 7). The Nb - rich regions are associated with small grained microstructure. Apart from the fine grains, some local areas have secondary abnormal grains of dopant [10], so the airgaps are deeper.

The dielectric constant evaluation has been made by capacitance measurements with two different methods in frequency range from 20 Hz to 1 MHz. According to the results plotted in Fig. 8, the permittivity for all samples decreases with increase of dopant concentration. At room temperature, the highest value of dielectric constant is ranged from 2600 for samples doped with 2 mol% to 3250 for the investigated samples doped with 0.5 mol% of Nb. The lowest value of dielectric constant (ε_r=2600) measured in samples doped with 2 mol% of Nb can be attributed to the formation of secondary abnormal grains that clearly lead to the decrease of permittivity. Moreover, the inhomogeneous distribution of dopand and formation Nb rich regions causes decreasing dielectric constant as well. After an insignificantly higher value of dielectric constant at low frequencies, ε_r becomes nearly constant at frequencies greater than 400 kHz.

As can be seen from the results show in Fig. 8, measured characteristics of $Nb/BaTiO_3$ samples with contacting and non-contacting electrode methods are quite consistent. It can be concluded that these methods can be equally used for dielectric characterization of $BaTiO_3$ ceramics. The relative deviation of dielectric constant between methods is shown in Fig. 9. The largest relative deviation was observed at low frequencies, where the samples have a maximum of dielectric constant. The deviation is in the range from 0.01 % to 3.2 %. Due to the changes of the instrument measurement range, it can be seen increase in relative deviation at 200 kHz.

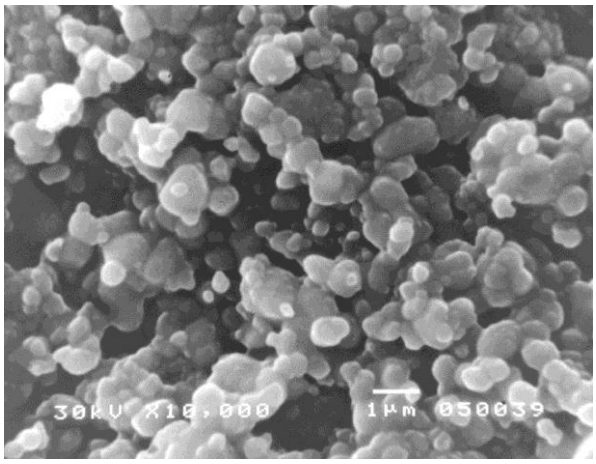


Figure 6. SEM image of 0.5 mol% $Nb/BaTiO_3$ sintered at 1320°C

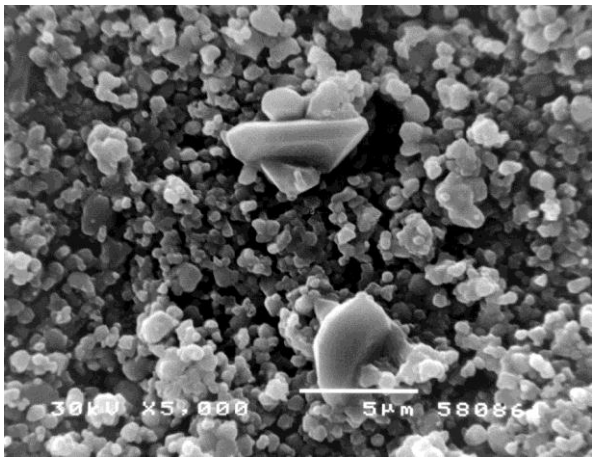


Figure 7. SEM image of 2 mol% $Nb/BaTiO_3$ sintered at 1320°C

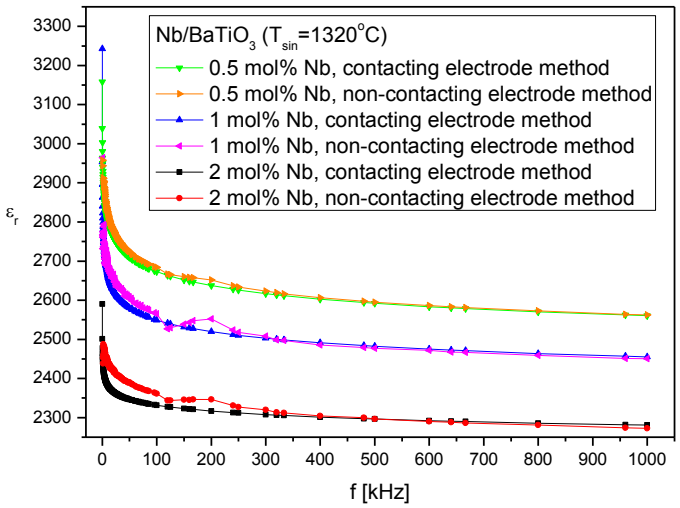


Figure 8. Dielectric constant as function of frequency for $Nb/BaTiO_3$ samples measured with contacting and non-contacting electrode methods

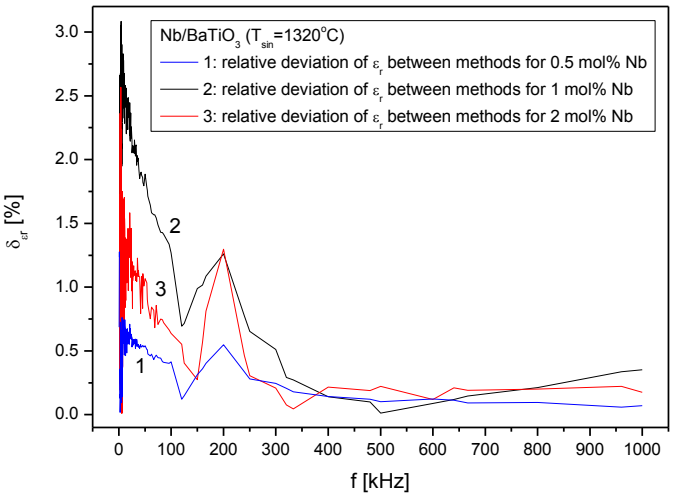


Figure 9. Relative deviation of dielectric constant between methods for $Nb/BaTiO_3$ samples

To achieve stable measurements with contacting electrode method, the pressure should be set at a level that doesn't deform the MUT. The pressure should be as strong as possible in order to minimize the occurrence of the airgap between the MUT and the electrodes. When the non-contacting method is employed, the electrode gap (t_g) is required to be at most 10% larger than the thickness of the MUT. The results of measurement error caused by airgap using (2) have been calculated in Table 1. The relative error is greater with increasing of t_a/t_m ratio and dielectric constant. Notice that the effect is higher with thin materials.

TABLE I. MEASUREMENT ERROR CAUSED BY AIRGAP

ε_r / t_a/t_m	**2500**	**3000**
0.001	71.4 %	74.98 %
0.005	92.56 %	93.72 %
0.01	96.12 %	96.74 %
0.05	99.17 %	99.30 %
0.1	99.56 %	99.63 %

V. CONCLUSION

In this paper the measurement methods for dielectric constant determination in $Nb/BaTiO_3$ ceramics have been investigated. The samples of modified barium titanate ceramics were used to compare contacting and non-contacting electrode methods. These three-terminal methods were used to eliminate a measurement error caused by the edge capacitance. The thin film electrodes were prepared by silver paste to cover the roughness on the samples surfaces and to reduce error due to airgap. Furthermore, it is noticeable that measurement error increases with increasing of airgap thickness. Very good results in a matching of $\varepsilon_r(f)$ plot for both methods are achieved. The maximum relative deviation between methods is 3.2%.

REFERENCES

[1] O. V. Tereshchenko, F. J. K. Buesink, F. B. J. Leferink, "An overview of the techniques for measuring the dielectric properties of materials", IEEE General Assembly and Scientific Symposium, 2011.

[2] T. T. Grove, M. F. Masters, R. E. Miers, "Determining dielectric constants using a parallel plate capacitor", Am. J. Phys. 73(1), 2005, pp. 52 – 56.

[3] H. Kassem, V. Vigneras, G. Lunet, "Characterization techniques for materials' properties measurement" in Microwave and Millimeter Wave Technologies: from Photonic Bandgap Devices to Antenna and Applications, InTech, 2010, pp. 289-314.

[4] D. Mistra, "Permittivity Measurement", CRC Press, 1999.

[5] Standard Test Methods for AC Loss Characteristics and Permittivity (Dielectric Constant) of Solid Electrical Insulation. Available: http://www.astm.org/Standards/D150.htm

[6] Agilent Basics of Measuring the Dielectric Properties of Materials, application note, Agilent Technologies, 2006.

[7] Agilent Solutions for Measuring Permittivity and Permeability with LCR Meters and Impedance Analyzer, application note 1369-1, Agilent Technologies, 2003.

[8] H. J. Wintle, "Capacitor Edge Corrections", IEEE Trans. Of Electrical Insullation, Vol. EI-21, No. 3, 1986, pp. 361 – 363.

[9] A. C. Lynch, "Edge capacitance in the measurement of dielectric properties", Proc. IEE, Vol. 120, No. 8, 1973, pp. 934 – 938.

[10] Lj. Zivkovic, V. Paunovic, M. Miljkovic, M. M. Ristic, „Microstructure Evolution and Dielectric Properties of Nb/Mn and Dy/Mn Doped Barium Titanate Ceramics", Materials Science Forum Vol 518, 2006, pp. 229-234.

[11] Agilent 4284A, technical overview. Available: http://cp.literature.agilent.com/litweb/pdf/5952-1431.pdf

[12] Agilent 16451B dielectric test fixture, technical overview. Available: http://cp.literature.agilent.com/litweb/pdf/16451-90020.pdf

Frequency Dependant Current Distribution and Resistance Coefficient of Aluminium Conductors

Karolina Kasaš-Lažetić, Dejana Herceg,
Dragan Kljajić
Department of power, electronics and telecommunications
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
kkasas@uns.ac.rs, vuletic@uns.ac.rs, dkljajic@uns.ac.rs

Miroslav Prša
Retired professor
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
prsa@uns.ac.rs

***Abstract*—Aluminium conductors, such as ACSR (Aluminium Conductor Steel Reinforced) or ACCR (Aluminium Conductor Composite Reinforced) are widely applied in electric power transmission and distribution systems. For that reason, we decided to investigate the behavior of those conductors, making an accent on current distribution inside them and resistance coefficient determination. In this paper we do present all calculation results for two ACSR and two ACCR, of similar characteristics, at frequencies up to 2500Hz. Obtained results, presented graphically, confirmed our expectations. The entire calculation was carried out applying COMSOL Multiphysics 3.5a computer program package.**

Keywords-Current distribution; graphical presentation; inhomogeneous conductors; resistance coefficient

I. INTRODUCTION

In all time-varying electromagnetic fields, inside conductors and inside all neighboring conducting bodies, skin effect and proximity effects appear. The consequences of those effects are non-uniform current distribution across conductor's cross-section, increased Joule's losses and virtual increasing of conductor's resistance. Non-uniform current distribution and increased conductor's resistance can be presented by several functional dependences. Current distribution is usually presented as dependence of current density vector magnitude on the distance from the conductor's axes. Two of the most common presentations of resistance coefficient, $k_R = R'_{\approx} / R'_{=}$, are the dependence on frequency and on factor $\sqrt{f / R'_{=}}$, where $R'_{\approx}$ denotes the AC resistance per kilometer of conductor's length, $R'_{=}$ is the DC resistance per kilometer of conductor's length and f is the applied frequency. Calculated resistance coefficients for homogenous conductors of different cross-sections, together with both mentioned graphical presentations were shown in [1]-[3]. Special accent on the second resistance coefficient functional dependence and its advantages was given in [1]. It was shown in [1] that, for significant skin effect in any homogenous conductor, resistance coefficient depends linearly on $\sqrt{f / R'_{=}}$ term,

$$\frac{R'_{\approx}}{R'_{=}} = \sqrt{\frac{\mu}{4}}\sqrt{\frac{f}{R'_{=}}} \,.$$

On the other hand, both aluminium conductors, ACSR (Aluminium Conductor Steel Reinforced) and ACCR (Aluminium Conductor Composite Reinforced) are not homogenous materials. The typical cross-section of both conductors is given in Fig. 1, [4]-[6].

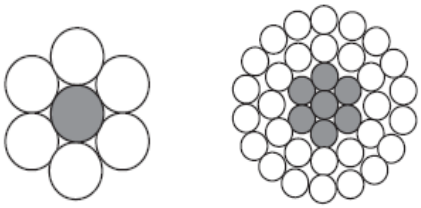
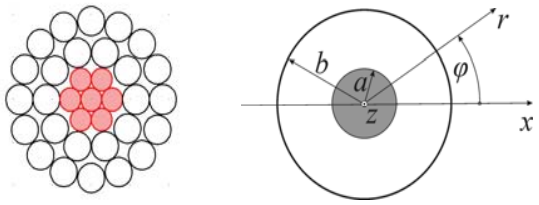


Figure 1. Typical cross-section of an ACSR or an ACCR conductor

Due to complex cross-section geometry, shown in Fig. 1, and even more due to different conductivities of supporting core (steel or aluminium oxide composite) and conductive aluminium wires around the core, we do expect much more complicated relations. Moreover, the nonlinear ferromagnetic core of ACSR, is expecting to play a significant role in all calculations and obtained results.

Investigating the possibility of the conductor's geometry simplification, a simplified model, shown in Fig. 2 b), is adopted for all calculations.



a) b)

Figure 2. a) Typical cross-section of an ACSR or an ACCR conductor, and b) and simplified model

It is shown that the results of calculations on simplified model correspond to the results of the real conductor's investigations [3], [7]. For that reason only the results obtained on simplified model will be presented in this paper.

II. Theoretical Approach

In order to determine conductors' AC resistance, total current distribution, induced electric field, and Joule's losses power must be calculated.

Considering typical cross-section shown in Fig. 2 b), the problem can be presented in cylindrical coordinate system with the conductor's axes positioned in z-axes of chosen coordinate system. The problem can also be considered as two-dimensional and linear. Although ACSR steel core is made of nonlinear, ferromagnetic material, magnetic field inside the core is negligible, so the permeability dependence on magnetic flux density magnitude is negligible as well.

Consequently, current distribution and resistance coefficient calculation can be performed in complex, two-dimensional domain, applying complex magnetic vector potential, which has only z-component and is a solution of following partial differential equation [8], [9],

$$\Delta \underline{A}_z - j\omega\mu\sigma \underline{A}_z = -\mu \underline{J}_z . \tag{1}$$

In cylindrical coordinate system, the above equation can be written as,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \underline{A}_z}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 \underline{A}_z}{\partial \varphi^2} - j\omega\mu\sigma \underline{A}_z = -\mu \underline{J}_z . \tag{2}$$

The calculated values of complex magnetic vector potential define the complex induced electric field strength vector and the complex induced current density vector, both in z-direction,

$$\vec{\underline{E}}_{ind} = -j\omega \vec{\underline{A}} , \qquad \vec{\underline{J}}_{ind} = -j\omega\sigma \vec{\underline{A}} . \tag{3}$$

Total complex current density vector is the sum of imposed complex current density vector, $\vec{\underline{J}}_0$ and induced current density vector, $\vec{\underline{J}}_{ind}$,

$$\vec{\underline{J}}_{tot} = \vec{\underline{J}}_0 + \vec{\underline{J}}_{ind} . \tag{4}$$

Joule' losses power per kilometer of conductor's length can be determined by integration across the conductor's cross-section,

$$P'_J = \int_{Scs} \frac{J_{tot}^2}{\sigma} dS . \tag{5}$$

AC resistance per kilometer of conductor's length can now be calculated as,

$$R'_{\approx} = \frac{P'_J}{\left|\underline{I}_0\right|^2} . \tag{6}$$

In order to define the resistance coefficient,

$$k_R = \frac{R'_{\approx}}{R'_{=}} , \tag{7}$$

the DC resistance must be determine as well. This resistance can be calculated applying the well known formulae for parallel connection of defined number of wires, N_k [3],

$$R'_{=core} = \frac{1}{\sigma_{core} N_{core} S_{core}} , \qquad R'_{=Al} = \frac{1}{\sigma_{Al} N_{Al} S_{Al}} . \tag{8}$$

$$R'_{=} = \frac{R'_{=core} \cdot R'_{=Al}}{R'_{=core} + R'_{=Al}} . \tag{9}$$

III. Model

DC resistances were calculated applying (8) and (9). In ACSR case the conductivities, $\sigma_{core} = \sigma_{Fe}$ and σ_{Al} were measured [2] and those values, applied in all resistance calculations are,

$$\sigma_{Al} = 3.55 \cdot 10^7 \text{ S/m}, \qquad \sigma_{Fe} = 0.559 \cdot 10^7 \text{ S/m}.$$

In an ACCR, aluminium and composite aluminium oxide conductivities are taken from [6],

$$\sigma_{Al} = 3.48 \cdot 10^7 \text{ S/m}, \qquad \sigma_{AlO} = 1.405 \cdot 10^7 \text{ S/m}.$$

AC resistance calculations were carried out applying AC/DC module of COMSOL Multiphysics 3.5a computer program package [10]. A 2D model, representing a cross-section of conductor in surrounding air domain, is set up. The mode "Quasi-static, Magnetic/Perpendicular Induction Currents/Vector Potential" is chosen, together with "Time-harmonic analysis".

As shown in [3], for resistance calculations the simplified model, shown in Fig. 2 b), together with cylindrical coordinate system, can be satisfactory applied.

In all calculating models, rms value of current is chosen to be the same in every conductor, equal, I = 1A. Having a linear problem, this value will not influence the final results. All conductors are surrounded by air, with defined electromagnetic characteristics, ε_r=1, μ_r=1, σ=0.

In order to define boundary conditions, an exterior boundary was involved. This boundary was supposed to be the circle of radius R_b=1m, out of which there is no electromagnetic field. [2], [3]. On the boundary, the surface current density,

$$J_s = -\frac{I}{2\pi R_b} \tag{10}$$

was supposed, providing the same return current in order to ensure zero electric and magnetic field outside the chosen modeling domain.

Magnetic vector potential is continuous across all interior boundaries between all inner conductors and between external aluminium conductors and surrounding air.

The applied software offers extended post processing calculations, so the resistance coefficient could be calculated easily.

All calculation results will be commented in following chapters. Moreover, the calculation results for two types of conductors, ACSR and ACCR, will be compared and commented as well.

IV. Obtained results

In order to enable comparison between two different type of conductors, we chose to investigate three pairs of conductors with similar cross-sections. According to [4] and [6], the conductors with similar cross-sections are: ACSR 240/40mm^2 and ACCR Hawk 477, ACSR 150/25mm^2 and ACCR Ostrich 300, as well as ACSR 680/85mm^2 and ACCR Martin 1351. Some constructive characteristics of both conductors' types are given in Table I.

TABLE I. Constructive Elements of ACSR and ACCR with Similar Cross-Section

Conductor type	**Aluminium outer wires**		Core
	No. & diameter of wires	***No. of layers***	***No. & diameter of wires***
	n x mm		***n x mm***
ACSR 240/40 mm^2	26(10+16)×3.45	2	7(1+6)×2,68
ACCR Hawk 477	26(10+16)×3.40	2	7(1+6)×2,70
ACSR 150/25 mm^2	26(10+16)×2.7	2	7(1+6)×2.1
ACCR Ostrich 300	26(10+16)×2.7	2	7(1+6)×2.1
ACSR 680/85 mm^2	54(10+16+28)×4.0	3	19(1+6+12)×2.4
ACCR Martin 1351	54(10+16+28)×4.0	3	19(1+6+12)×2.4

As an example, the current distribution inside the ACSR 240/40 mm^2 at 50 Hz is given in Fig. 3 and the current distribution at 450 Hz is given in Fig. 4.

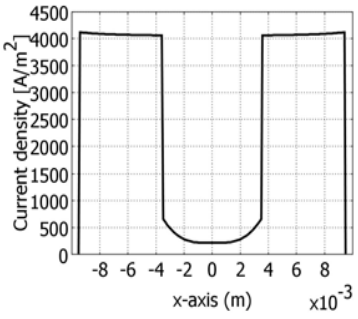


Figure 3. Current distribution across the ACSR 240/40 mm^2, at 50 Hz

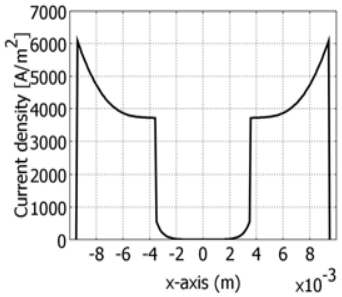


Figure 4. Current distribution across the ACSR 240/40 mm^2, at 450 Hz

Two basic characteristics in ACSR behavior could be observed in Fig 3. First, due to different conductivity values, even at 50 Hz a very small amount of current exists in the steel core. This amount is additionally smaller due to significant skin effect, produced by big permeability of steel part of the conductor. At 50Hz skin effect inside aluminium part is not significant. Hence, current distribution in this part is almost uniform.

At the frequency of 450 Hz the skin effect is much more emphasized, which can be noticed in Fig. 4. First, there is no current inside the steel core and the current distribution inside aluminium part is no longer uniform; the current density vector magnitude increases toward conductor's surface.

The same diagrams for the similar ACCR, Hawk 477, are presented in Fig. 5 and Fig. 6.

In Fig. 5 the main characteristics of ACCR can be seen. The difference in conductivity values between the core and the external conducting wires are much smaller comparing ACSR. Consequently, inside the core much bigger percentage of current exists at all frequencies.

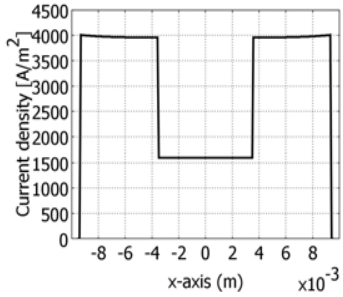


Figure 5. Current distribution across the ACCR Hawk 477, at 50 Hz

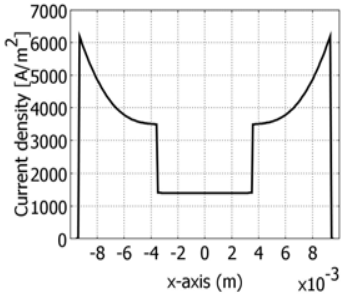


Figure 6. Current distribution across the ACCR Hawk 477, at 450 Hz

Besides, the relative permeability of aluminium oxide composite core is equal to one, so the skin effect is almost negligible at 50 Hz.

In Fig. 6 a significant skin effect can be noticed, presented by smaller amount of current inside the core and by non-uniform current distribution in both parts of a conductor.

Calculated resistance coefficients for those conductors depending on frequency, are presented in Fig. 7.

The upper pair of curves is for the 680/85 and Martin 1351 conductors, while the smaller resistance coefficients represent 150/25 and Ostrich 300 pair of conductors.

Unexpectedly, probably due to smaller conductivity of ACCR aluminium wires, the resistance coefficient of both ACCR is bigger than of the similar ACSR resistance coefficient.

Much more interesting is the other graphical presentation modus, shown in Fig. 8.

The results presented in Fig. 8 are very interesting. Despite the fact that the conductors are inhomogeneous, this resistance coefficient dependence is a straight line again, as it was for homogeneous conductors. Obviously, in this case, resistance coefficient does not depend on conductors' cross-sections or conductors' conductivity values. The difference between ACSR and ACCR remains, but it is smaller that in previous diagram.

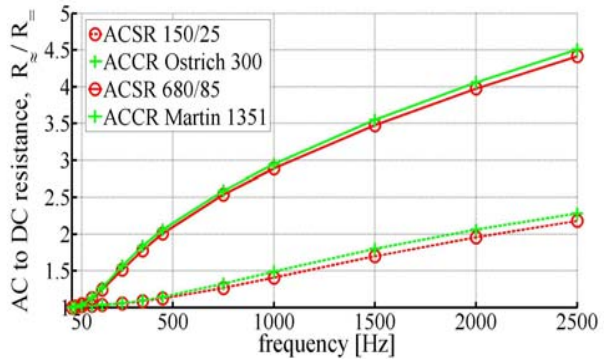


Figure 7. Resistance coefficient as a function of frequency.

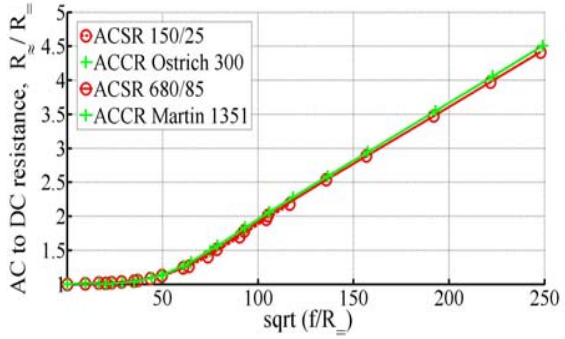


Figure 8. Resistance coefficient as a function of $\sqrt{f / R_{=}}$ term.

V. CONCLUSION

An attempt to explore the behavior of inhomogeneous conductors, frequently applied in energy delivery systems, was successfully performed. Current distribution and resistance coefficient were calculated and graphically presented for six types of aluminium conductors, comparing the pairs of ACSR and ACCR conductors with similar cross-sections.

Most of the results were as expected, regarding the current distribution with more or less emphasized skin effect, but the fact that ACSR resistance coefficient is smaller for all calculated frequencies, was a surprise.

Another goal was achieved as well. It was shown that again, one more time, the dependence of resistance coefficient on the term $\sqrt{f / R_{=}}$ does not depend on conductor's cross-section size, or on conductor's conductivity. For that reason, this functional dependence still is important and convinient for all applications of inhomogenous conductors as well.

All examples were calculated applying COMSOL Multiphysics 3.5a computer program package, which is found one more time as a powerful tool for all electromagnetic fields problems evaluations.

ACKNOWLEDGMENT

This paper has been supported by Ministry of Sciences and Technological Development of the Republic Serbia, under the grant for project TR 32055.

REFERENCES

[1] M. Prša., K. Kasaš-Lažetić, N. Đjurić and B. Vukobratović, "Why Resistance Coefficient, R'≈/R'=, as a Function of sqrt(f/R'=)", Acta Electrotehnica, Proceedings of the 4th International Conference on modern Power Systems MPS 2011, Cluj-Napoca, Romania, pp 382-386, May 2011.

[2] M. Prša, K. Kasaš-Lažetić and N. Đurić, "Determination of Frequency Dependent Radiuses of Steel Cored Aluminium Conductors", 2nd International Conference on EMF-ELF, Paris, France, Paper No A-P-05 p.117-123, March 2011.

[3] K. Kasaš-Lažetić, M. Prša, V. Bajović and B. Vukobratović, "Determination of ACSR's Electrical Characteristics", 10th International Conference on Applied Electromagnetics, IIEC 2011, Niš, Serbia, Session O2 O3, September 2011.

[4] "Ropes for overhead lines", brochure in PDF format, (2010), retrieved January 13, 2014, from: http://www.elka.hr/media/katalog/14/alu-celicna_uzad_za_nadzemne_vodove.pdf, January 2014.

[5] Lj. Gerić, P. Đapić, Distrubition Equipments, 2nd ed., Novi Sad: University of Novi Sad, Faculty of Technical Sciences, Novi Sad, 2006, p.147 (in Serbian).

[6] „3M™ Aluminium Conductor Composite Reinforced (ACCR)", Catalog in PDF format, issued by the manufacturer.

[7] K. Kasaš-Lažetić, D. Herceg, N. Đurić, N. Mučalica and *M.* Prša, "Frekvencijski zavisne radne karakteristike Al-Če i Al-AlO provodnika", XIII Međunarodni naučno-stručni simpozijum Infoteh-Jahorina 2014, Jahorina, BiH, vol 13. Ref ENS 3.7, pp. 165-168, March 2014. http://www.infoteh.rs.ba/rad/2013/ENS-4/ENS-4-6.pdf

[8] B. D. Popović, Electromagnetics, 2nd ed., Belgrade: Gradjevinska knjiga, 1986, pp. 35-37 and p. 269 (in Serbian).

[9] D. Ivanović, Vector Analysis, 2nd ed., Belgrade: Naučna knjiga, 1960, pp. 211-218 (in Serbian).

[10] Documentation COMSOL MULTIPHYSICS, CLS 3.5a, 2009.

Eddy Currents inside Pipeline buried beneath HV Overhead Power Transmission System

Dejana Herceg, Karolina Kasaš-Lažetić,Dragan Kljajić
Miroslav Prša
Department of power, electronics and telecommunications
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
vuletic@uns.ac.rs

Nikola Mučalica
"Elektrovojvodina", doo, Novi Sad

Novi Sad, Serbia

***Abstract*— Conductive pipes for liquids and gas transportation are frequently buried in the ground, in vicinity of a high voltage three-phase overhead electric power delivery system. The heating of the pipe, due to eddy currents, is of crucial importance.**

In this paper investigation of heating effect is performed at a real problem. First of all it was determined which voltage level power transmission system produces the most significant eddy currents. After that a model of real zinc coated steel pipe was positioned in vicinity of electric system. For nominal currents in electric power delivery system, external magnetic field and eddy currents were calculated, together with power of heating losses inside the pipe's wall.

All calculations were performed numerically, applying COMSOL Multiphysics 3.5a computer program, for the worst pipe position and minimal height of power delivery conductors from the ground, at eight different frequencies, up to 450 Hz. The calculation results, magnetic field distribution, induced current distribution and frequency dependent heating power, are given graphically.

All calculated results show that, for investigated type of power delivery system, currents induced in the pipe's walls are negligible, except in case when the pipe is extremely closed to the power delivery system, which cannot happen in practice.

Keywords—Time varying magnetic field; current induced in pipe's wall; eddy currents losses

I. INTRODUCTION

Metal, conductive pipes for liquids or gas transportation are always placed at previously determined corridors, on the ground surface, above the ground surface or buried into the ground, on the depth defined by appropriate standards. In all those situations it might occur that pipe transportation system approaches or crosses the corridor of electric power transmission or distribution system. Time varying magnetic field, produced by time varying currents in electric system conductors, induces eddy currents in pipe's wall. As any current, those currents are followed by Joules losses and pipe heating, which may provoke even an accident if some inflammable materials are transporting.

For that reason it is very important to be well acquainted with impact of any electric system on neighboring system containing conductive elements.

Time varying magnetic field and eddy currents calculations for some of electrical transmission and distribution systems were performed and presented in [1] and [2], while its influence on all neighboring conductive elements was shown in [3].

Before the main calculations of magnetic flux density vector and eddy current distribution take place, it has to be predicted the electrical system voltage level which produces highest eddy currents. Considering line ending transformers' power, nominal currents inside the system's conductors is defined and magnetic field can be determined. Taking into account minimal standardized conductors' heights above the ground as well, it was concluded that the most critical situation can occur at 400kV transmission system. In this case magnetic field on the ground surface has maximal values, producing maximal induced electromotive forces and maximal current induced in the pipe's wall.

In order to determine and present an impact of overhead high voltage electrical power transmission system on conductive pipe, a real problem, with standardized zinc coated steel pipe, buried in the ground was investigated. The geometry of calculated problem is presented in Fig. 1.

All calculations will be performed for the worst pipe position related to the electric system; the pipe is buried on 1m depth, parallel to the electric system, exactly below the central conductor and for phase arrangement 0-4-8.

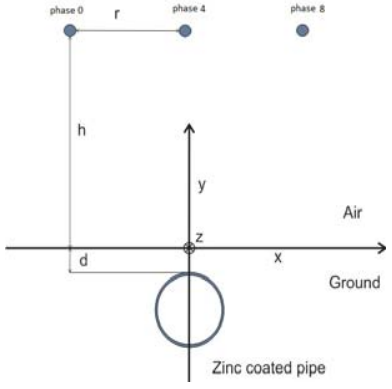


Figure 1. Geometry of investigated problem.

II. Theoretical Approach

Supposing that all currents inside three-phase system conductors are harmonic and that the entire system is linear, complex analysis can be applied. Actually, steel pipe is not linear material, but it can be taken as linear, with relative permeability much higher than 1, not dependent on magnitude of complex magnetic flux density vector. As a matter of fact it can be expected that magnetic field inside the pipe's wall has very low magnitude and for this reason the pipe's permeability can be treated as initial static or initial dynamic, but constant value.

The calculation starts from well known partial differential equation in complex domain involving complex magnetic vector potential [4], [5],

$$\Delta\underline{\vec{A}} - j\omega\mu\sigma\,\underline{\vec{A}} = -\mu\underline{\vec{J}}\,. \tag{1}$$

In chosen Cartesian coordinate system, presented in Fig. 1, for imposed complex current density vector having only z component, the above equation can be written as,

$$\frac{\partial^2 \underline{A}_z}{\partial x^2} + \frac{\partial^2 \underline{A}_z}{\partial y^2} - j\omega\mu\sigma A_z = -\mu\underline{J}_z\,. \tag{2}$$

Z component of complex magnetic vector potential defines z component of complex induced electric field strength vector,

$$\underline{E}_{indz}(x,y) = -j\omega\underline{A}_z(x,y)\,, \tag{3}$$

and z component of complex induced current density vector (complex eddy current density vector),

$$\underline{J}_z(x,y) = \sigma\underline{E}_{indz}(x,y)\,. \tag{4}$$

Power of Joules' losses inside pipe's wall is defined per kilometer of the pipe's length and calculated as an integral over the pipe's wall cross section, S_{pwcs},

$$P'_J = \int_{Spwcs} \frac{|\underline{J}_z(x,y)|^2}{\sigma} dS\,. \tag{5}$$

III. Model

As said in introduction, the most intensive eddy currents will appear in the pipes buried in vicinity of 400kV high voltage overhead transmission line. For that reason in this paper a real problem was investigated. Below a real power transmission line, with characteristics defined in [6], 1m below the earth surface, real zinc coated steel pipe, with Φ 1020mm (40″) and the wall thickness 8.2mm is buried, as shown in Fig. 1.

All calculations were carried out for minimal conductor's height allowed by Serbian standards; 7.5m (+40˚C or -5˚C plus due to weight of ice on conductors).

It should be mentioned that, according Serbian standard SRPS N.CO.105:1987, the conductive pipe cannot be placed so close to HV power delivery system, especially not beneath the system, parallel to the system conductors. Nevertheless, the problem presented in Fig. 1 was investigated, in order to become sure that higher values of heating losses could not appear in practice.

In all calculations, exploring the real case, the RMS modulus of complex current in each conductor of the symmetric three-phase power delivery system was supposed to be the nominal one, $|\underline{I}|$=430A for the power transformer's nominal power of 300MVA.

Several different pipe positions related to electric system were investigated, but only the worst pipe's position results, are presented in this paper.

Other relevant values, taken into consideration in entire calculations are as follows: The phase arrangement in electrical system is (0-4-8), the distance between the nearest conductor and pipe's surface is 8.5m, ground resistivity is 50Ωm and relative permeability of steel pipe is, μ_r=4000.

Moreover, in order to explore the behavior of the problem even in the case when higher current harmonics are present as well, the calculation was performed for eight chosen frequencies; 0.01Hz, 16⅔Hz, 50Hz, 100Hz, 150Hz, 250Hz, 350Hz and 450Hz.

All calculations were carried out applying COMSOL Multiphysics 3.5a computer program package [7], based on finite elements method.

IV. Obtained results

The magnetic flux density distribution, produced by a HV overhead electric power transmission system, along y-axis, in a case of steel pipe buried in the ground, for the pipe positioned as presented in Fig. 1, at 50 Hz, is shown in Fig. 2.

The diagram shown in Fig. 2 has an expected shape. Due to low frequency and smaller influence of induced currents on magnetic field distribution, magnetic flux density vector magnitude is practically equal in the entire pipe's wall cross-section. From Fig. 2 it can be noticed that the magnitude is a little higher in the pipe's wall closer to ground surface, at y = -1m, compared to the part of pipe's wall the most distant from earth surface, y = -2m. Inside the pipe magnetic flux density vector magnitude is negligible, due to shielding effect of ferromagnetic pipe.

For the same model, magnetic flux density vector magnitude distribution at 450Hz is presented in Fig. 3.

The shape of magnetic field distribution inside the ground, at 450Hz, shown in Fig. 3, differs from previous case. Magnetic flux density vector magnitude is much higher in the pipe's wall part closer to the earth surface. Higher frequency produces higher induced currents and its more significant impact on both magnetic field and eddy current distribution. Currents induced in the closer pipe's wall part produce magnetic field which decrease magnetic field in distant pipe's wall parts, acting as an electromagnetic shield.

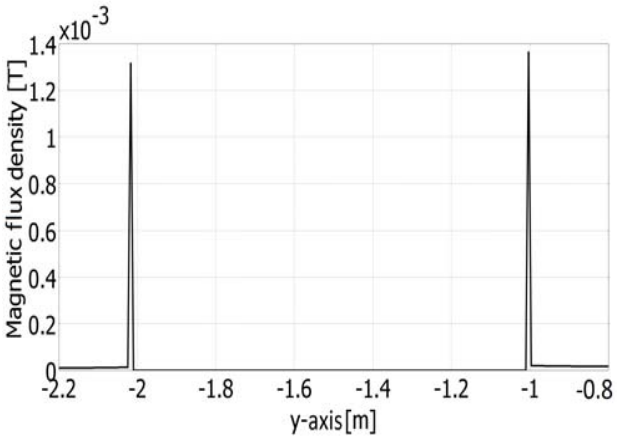


Figure 2. Magnetic flux density distribution along y-axis, at 50Hz.

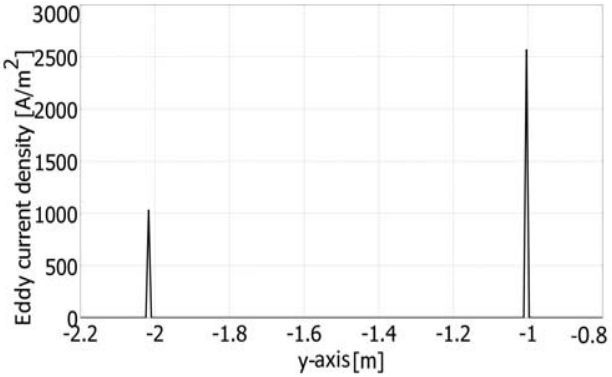


Figure 4. Induced currents density distribution along y-axis, at 50Hz.

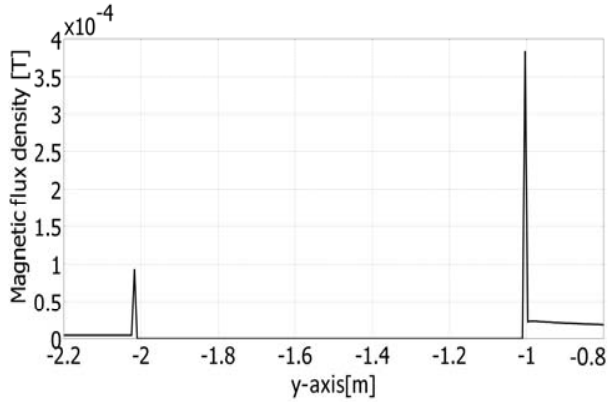


Figure 3. Magnetic flux density distribution along y-axis, at 450Hz.

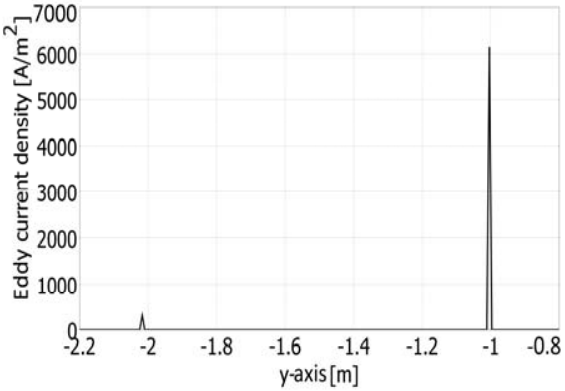


Figure 5. Induced currents density distribution along y-axis, at 450Hz.

Time varying magnetic field, produced by currents in electrical system conductors, provokes induced electromotive force, which cause the motion of free charges inside pipe's wall, forming induced currents (eddy currents).

Induced current density vector magnitude distribution inside pipe's wall, along y-axis, at 50 Hz, is shown in Fig. 4.

Similar to the magnetic field distribution, induced currents at 50Hz have smaller value and less significant effect on magnetic field distribution. Currents induced in pipe parts closer to the earth surface (y = - 1m) are stronger and produce magnetic field which decreases total magnetic field dipper in the ground. This effect can be noticed in Fig. 3, showing smaller magnetic flux density vector magnitude at y = - 2m and consequently smaller value of eddy currents.

Induced current density vector magnitude distribution at 450Hz is shown in Fig. 5.

Due to higher frequency and more significant shielding effect, smaller magnetic field in the pipe's parts dipper in the ground, which can be noticed in Fig. 3, produces almost negligible eddy currents in a pipe's wall parts at y = - 2m, presented in Fig. 5.

As said in introduction, part of electrical transmission system energy through induced electric field and induced currents in pipe's wall, will transform to thermal energy inside the pipe's wall. Considering thorough energy transmitting by electrical power delivery system, this loss of energy is negligible. Nevertheless, due to the thermal energy, pipe temperature may increase, which can be significant especially in the cases of inflammable material transportation inside the pipe. In that case, even a fire brake out or an explosion can occur.

For that reason the knowledge of Joules' losses power, defined by (5), is very important. As emphasized before, this power is determined per kilometer of pipe's length and its dependence on frequency is presented in Fig. 6.

At industrial frequency, f=50Hz, the Joules' losses power per kilometre of pipe's length are,

$$P'_{J50Hz} = 51W / km .$$

At the highest examined frequency, f=450Hz, the Joules' losses power per kilometre of pipe's length are,

$$P'_{J450Hz} = 131W / km .$$

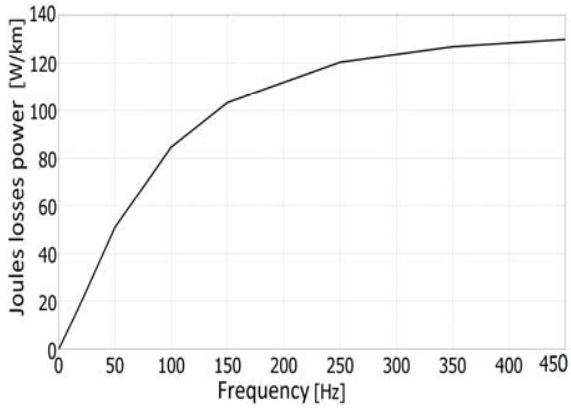


Figure 6. Frequency dependent eddy currents losses power inside pipe's wall, per kilometer.

Obviously, even in the worst case which can not appear in practice, Joules' losses power is insufficient to produce a significant heating of the fluid inside the pipe.

V. CONCLUSION

Prior to all calculations, the worst case had to be defined. Investigating magnetic field, produced by overhead high voltage three-phase power delivery systems, with nominal system currents, at the earth surface, the transmission system at 400kV was determined as the most significant one.

After that, we started with magnetic field calculations, in electrical system vicinity and eddy currents inside the buried zinc coated pipe's wall determination. Several different positions of electrical system and pipe were explored but in this paper only the worst case was presented. The entire calculation has been repeated for 8 different frequencies, up to 450Hz, in order to take into account possible presence of higher current harmonics.

The calculation results were as expected. Even in the worst case, which, according to existing standards, can not appear in practice, calculated Joules' losses power is not sufficient to provoke significant heating of the pipe, nor of fluid transporting inside the pipe.

Nevertheless Serbian standard, SRPS N.C0.105:1987, suggests that, if there is a need for crossing of two systems, the pipe system should be positioned orthogonal to the electrical system's conductors. In that case the currents induced in pipe's wall are minimal and the fluid heating is minimal as well.

The problem was successfully resolved applying COMSOL Multiphysics 3.5a computer package. For all calculations the AC/DC Module was applied, together with materials library data.

ACKNOWLEDGMENT

This paper has been supported by Ministry of Sciences and Technological Development of the Republic Serbia, under the grant for project TR 32055.

REFERENCES

[1] M. Prša, N. Mučalica, K. Kasaš-Lažetić, "Determination of electromagnetic field in a three-phase system with three single core power electric cables 110kV", Proc. of the IEEE region 8 EUROCON 2009 Conference, Saint-Petersburg, Russia.

[2] M. Prša, K. Kasaš-Lažetić, N. Mučalica, "Skin effect and proximity effect in a real, high voltage double three-phase system", Proc. of the IEEE region 8 EUROCON 2011 Conference, Lisbon, Portugal.

[3] N. Mučalica, M. Prša, K. Kasaš-Lažetić, D. Herceg, "Impact of HV Cable System on Neighboring Conductive Pipes", 16th International Symposium on Power Electronics – Ee 2011, Novi Sad, Serbia, October

[4] B. D. Popović, *Electromagnetics*, 2nd ed., Belgrade: Gradjevinska knjiga, 1986, pp. 35-37 and p. 269 (in Serbian).

[5] M. Stojaković, *Mathematical Analysis 2 – second part*, "Fakultet tehničkih nauka Novi Sad", 2007, 101-120 (in Serbian).

[6] Lj. Gerić, P. Djapić, *Distribution* equipments, Univerzitet u Novom Sadu, "Fakultet tehničkih nauka Novi Sad", 2006, 137-142 (in Serbian).

[7] Documentation COMSOL MULTIPHYSICS, CLS 3.5a, 2009.

Session T-02
POWER ELECTRONICS

The use of microcontrollers in modern solutions of power electronics

Miroslav Lazić, Dragana Petrović, Skender Miodrag
Iritel a.d. Beograd
Belgrade, Serbia
mlazic@iritel.com, titelac@iritel.com,
zorancvejic@iritel.com

Sandra Sovilj Nikić
Iritel a.d. Beograd
Belgrade, Serbia
sandrasn@eunet.rs

***Abstract*— The role of modern power electronics is to provide reliable and uninterruptible power supply for the consumers. Electric power stations are the most commonly used primary sources of energy. That energy is distributed to the consumers by large power distribution grid. Secondary power sources are usually some sort of alternative power sources or batteries. All power sources have to be controlled all the time. There have to be a system of power electronic devices that provides reliable and continuous power supply operation. Modern power electronic devices are switching type. The basic elements of a switching power supply are power electronics and control electronics. The control electronics are suitable for controling with logical circuits. Because of that, microcontrolles are now part of every modern power electronic devices. If the system has more than one power electronic devices controlled by the microcontroller then multiply microcontroles have to be coordinated to work as one. This paper describes an implemented solution of uninterrupted power supply for telecommunications equipment that uses either power distribution grid either alternative power source either storage battery. The microcontroller is a part of every power source that equipment uses. Also it controls and adjust all the power sources so that equipment has reliable and continuous power supply.**

Keywords- alternative power sources, microcontrolles, batteries

I. INTRODUCTION

The application of a switching mode in power electronic devices has enabled an achievement of the efficiency level over 95% as well as reduction in size of devices. In addition, it has enabled easy monitoring and control at the digital level [1]. Microcontrollers have become an integral part of power electronic devices. In the beginning they were used only in the specific solutions and when microcontrollers become widely available they become an integral element of any complex power electronic device. There are different approaches to the use of a microcontroller. It is common that the microcontroller only monitors the operation of power electronic devices. There are solutions where the microcontroller, in addition to monitoring, is used to define the mode of power electronic devices (source of constant voltage or source of constant current, setting the output size value, setting the level of protection, etc.). There is also a third group of power electronic devices, in which a microcontroller is used for generating signals for control of the power switches. In the third group of solutions, power electronic devices, different in function, could be generated by the same power electronic assembly. The system of power electronics consists of more power electronic devices which together provide the necessary electricity for operation of a consumer. If the individual elements of the system have a microcontroller, it is logical that there is a system microcontroller which would control and direct the operation of each power electronic device as well as the entire system of power electronics. This paper presents a solution for supplying telecommunication facility from multiple power sources. Each energy source has an associated power electronic device, and the system microcontroller adjusts the operation of all elements in the system according to the given operation algorithm.

II. ORGANIZATION OF UNINTERRUPTED SUPPLY OF A MODERN TELECOMMUNICATION FACILITY

Uninterruptable power supply of telecommunication facilities is based on the principle of parallel operation of a chemical power source (batteries) and DC voltage obtained from the electric power distribution grid. The nominal value of the DC voltage is 48V. This value can be changed in the range from 42 to 56V. The device which generates DC voltage from AC voltage is called a rectifying unit. In modern solutions of power supply organization in addition to the voltage obtained from electric power distribution grid alternative energy sources are also used as power sources. These are solar panels, wind turbines and fuel cells. Figure 1 shows a block diagram of the organization of continuous operation of the modern telecommunication facilities. The energy system has to provide reliable power for telecommunication and IT equipment, with the well-defined electrical characteristics and precisely defined the autonomy of operation. The system has to be able to optimize the power consumption, in order to maximally extend the autonomy of the most important devices at a given time. Power system has to efficiently use the energy from alternative energy sources. The solution has to be configurable, in order to could be applied for different types of telecommunication centers.

A. Power sources

It is accepted that the main voltage for power supply of stationary and mobile telecommunications devices is DC voltage. Energy sources could be:

- electric power distribution grid,
- electric power unit (aggregat) and,
- storage batteries (DC voltage).

The voltage of a electric power distribution grid and a aggregat is an AC voltage whose nominal value is 230 V.

Rectifying unit is a power electronic device which generates DC voltage, needed to power telecommunications devices, from the AC voltage. In addition, the rectifying unit has a possibility to charge a storage battery. The most common solutions of uninterrupted power supply of telecommunication systems use all three energy sources (electric power distribution grid, electric power unit and storage battery). There are also solutions that rely on two sources, one of which is always storage battery. Switching power supply from power distribution grid to power supply from the aggregat is not immediate, and in the transitional period storage batteries are used. Telecommunication centers are usually installed in stationary objects, but often users such as the military and police, have telecommunication centers installed in vehicles (mobile telecommunication centers). Stationary telecommunication facilities are often located in urban areas where the voltage from electric power distribution grid is available. The main power source is the voltage from electric power distribution grid. Storage battery and power unit are backup power sources. In the telecommunication facilities in which the power supply from electric power distribution grid is reliable, they do not use power unit as a backup power source. They use storage battery as a backup power source. Of course, the capacity of batteries has to be sufficient to achieve the required operation autonomy of telecommunication station. For the stationary telecommunication facilities, the nominal value of the power supply voltage of telecommunications center is 48V. This means that the voltage changes in the range from 42V to 55V, depending on the level of charge in the storage battery, and temperature. It is standardized that positive pole of the DC voltage to be grounded. In the Figure 1 block diagram of the organization of uninterrubtable power supply for stationary telecommunication facilities is shown.

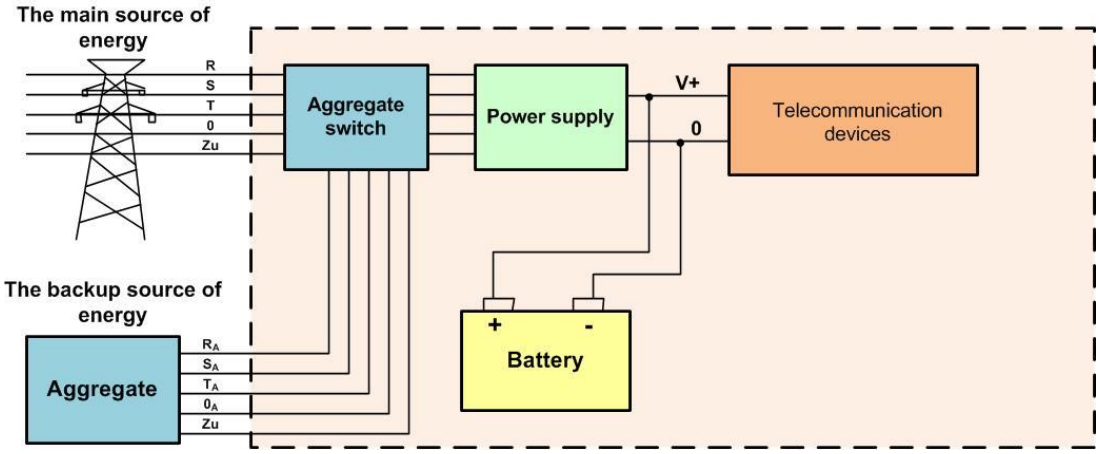


Figure 1. Organization of uninterruptable power supply of stationary telecommunication center

B. Organization of power supply of telecommunications device

In practice, there are two ways of organizing of uninterruptable power supply of telecommunications device:

- centralized,
- distributed.

The Figure 2 presents the solution of centralized power supply. A rectifying unit and a set of storage battery powers a greater number of telecommunications devices. The advantage of this implementation is simple realization and maintaining. The lack of this conception is that one failed power electronic device can cause cancellation of all consumers. In addition, the rectifying unit has to be dimensioned for the potential consumers who will be installed in the future. This means that the rectifying unit and battery are oversized which increases the cost and difficulty of the power supply. It also reduces the efficiency of power electronic devices. That is why these solutions are mainly used in stationary telecommunication facilities. DC distribution enables connection of a larger number of consumers to the same rectifying unit [2].

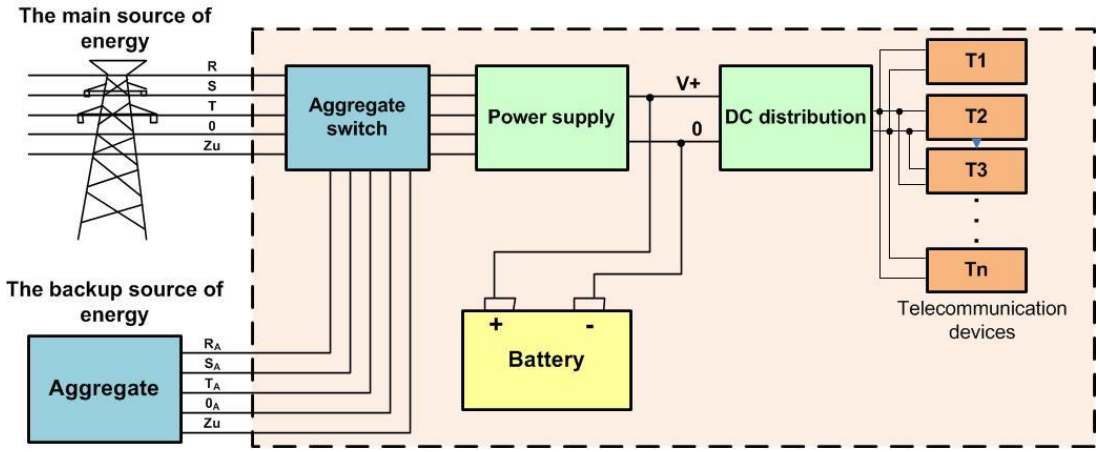


Figure 2. Centralized power supply of telecommunications devices

In a distributed solution, each telecommunications device has its own rectifying unit and own battery as shown in Figure 3.

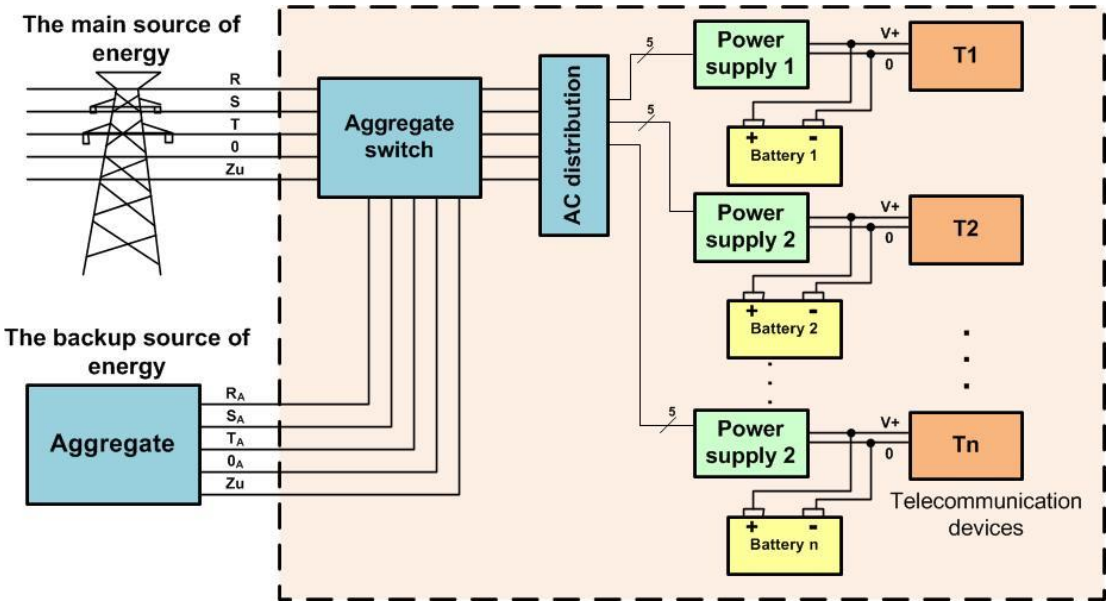


Figure 3. Distributed power supply of telecommunications devices

Distributed solution of organizing of power supply of telecommunication systems eliminates the disadvantages of centralized solution. Cancellation of a rectifying unit will cause interruption of operation of only one telecommunications device. However, the price of power supply from distributed concept is much more than the cost of centralized power supply, which also applies to maintenance costs.

Having in mind the disadvantages of the previous two configurations, the solution shown in Figure 4 was created. It is a modified version of distributed power supply. Only one set of batteries is used. AC/DC converters are used instead of rectifying units. These converters generate stable DC voltage with a nominal value of 48V. They are simpler and less expensive than rectifying units. The value of the output voltage is adjustable through a microcontroller that is an integral part of the converter. In addition to setting the value of the voltage microcontroller can set the mode so that the converter can operate as a constant voltage source and a constant current source. The same converter is used to charge the battery - battery chargers. Then they operate as a constant current. This approach for solving the problem of uninterruptable power supply can be applied in both stationary and mobile telecommunication systems. However, it is better to use it in mobile telecommunication centers as it allows optimization of power to the consumers. The voltage converters supply groups of consumers. To charge the battery the same converter is used, but it is set to operate as a battery charger (IU characteristic). Only one set of batteries is used. The configuration in Figure 4 is configurable. The number of voltage converters is defined by the real needs.

Figure 4 The modified configuration of distributed power supply optimized for mobile telecommunication centers

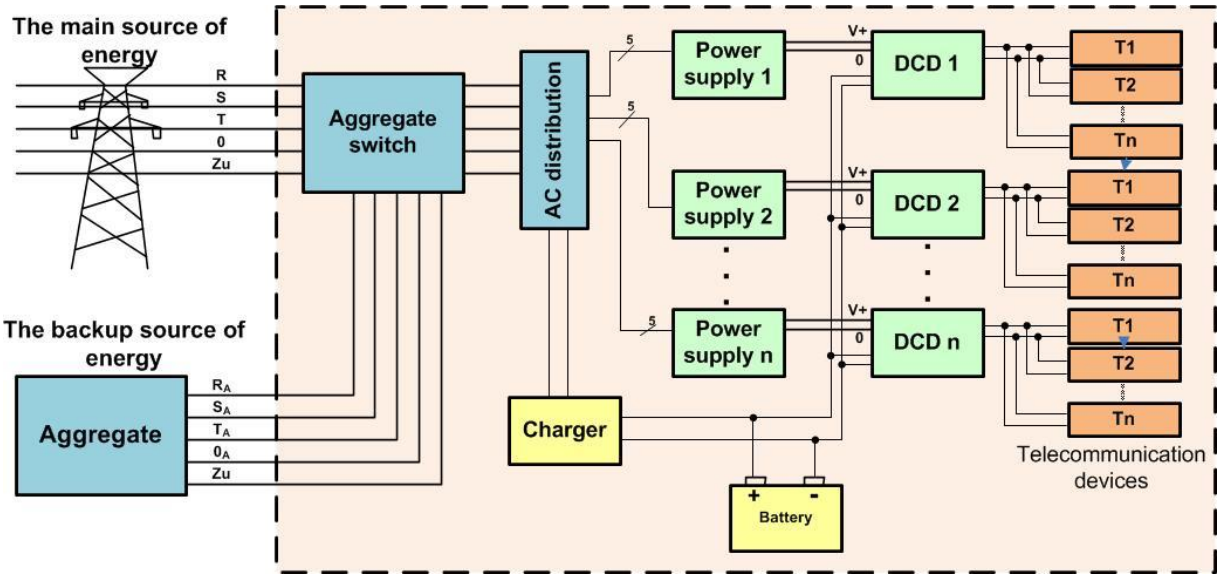


Figure 4. The modified configuration of distributed power supply optimized for mobile telecommunication centers

The configuration shown in Figure 4 is suitable for connection to alternative (renewable) energy sources. Solar panels are DC sources and they can be used to charge the system battery without conversion to AC voltage. The same is true for less power (up to 10kW) wind turbines. Reducing the number of voltage conversion the efficiency of use of alternative sources increases.

The occurrence of renewable energy sources telecommunications companies are interested in implementation them primarily in remote, inaccessible telecommunication facilities. The use of fuel cells, solar panels and wind turbines has started. The aim is to avoid the power units, reduce operating costs of renewable energy sources, and certainly reduce environmental pollution. Simultaneous use of energy sources based on fossil fuels and renewable energy sources can achieve the desired goal.

III. CONNECTION OF ALTERNATIVE ENERGY SOURCES

A usual configuration for connecting renewable energy sources is shown in Figure 5. Electricity produced from solar or wind is connected in parallel with existing energy sources: power generator, power unit and storage battery. It is not realistic to expect that renewable energy sources at once completely replace energy sources based on fossil fuels, primarily due to the low efficiency. Interruptions in the power supply of telecommunication facilities cause disruption of telecommunications traffic, and that means big losses for the telecommunications companies. It is realistic to expect that in the first phase of the implementation of renewable energy sources, renewable energy sources will be connect in parallel with the existing power systems. Renewable energy sources will supply telecommunication system, and if there is not enough energy from renewable energy sources in that case power supply from the resources based on fossil fuels will be used. In following stages the sources based on fossil fuels will be slowly replaced with renewable energy sources [3-5].

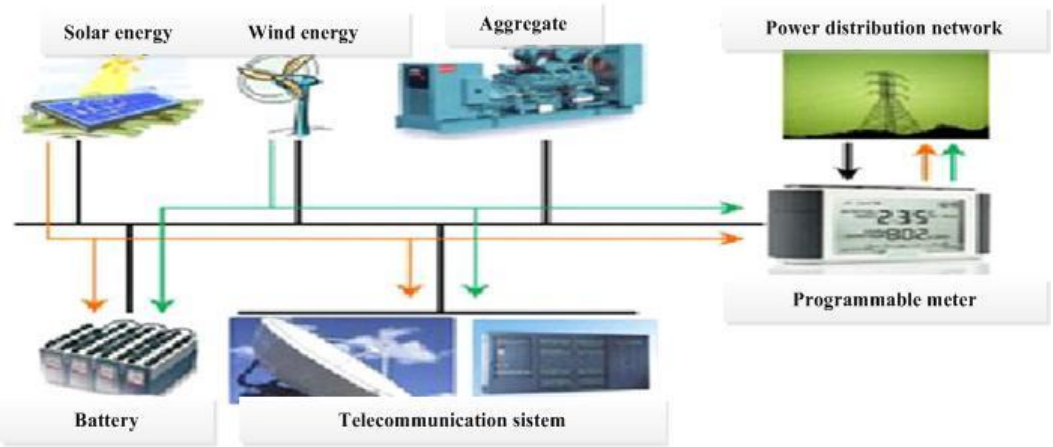


Figure 5. Connecting and control of telecommunications consumers power supply

A. Distribution at the AC voltage level

Figure 6 shows the connection of unconventional sources of energy and electric power distribution grid for power supply of consumers. Switch P1 is the power unit relay, which in the case of power failure in the electric power distribution grid includes power unit and connects it with consumers. Switch P2 enables parallel connection of voltage generated by the solar panels, fuel cells and wind turbines, with energy sources based on fossil fuels. Thus, adding is at AC voltage level. To achieve this, inverter is a necessary element. Inverter generates AC voltage from the DC voltage from the output of renewable energy source generated based on fossil fuels. There are two possible operation modes of configuration shown in Figure 6.

- Parallel operation of electric power distribution grid and inverter power generated from renewable energy sources,
- alternative operation - telecommunication facility is supplied from the electric power distribution grid or from the voltage generated by the inverter.

The first method allows the excess of energy produced by renewable energy sources to be placed in power distribution grid. However, in order to connect two sources of AC voltage in parallel a necessary condition is to have the same effective value, the same frequency and the same phase angle. Of course, it is feasible, but economically difficult profitable. Due to its complexity, parallel operation is less reliable and it is not really to be used for supply of telecommunication facilities.

An alternative manner of power supply is easier for realization, and therefore it is more reliable. The voltage at the output of the inverter does not have to be synchronous and have the same phase to the voltage in the power distribution grid. When switching from one to another source of energy, the power interruption appears. Interruption is necessary because the voltage of the power distribution grid is not synchronous and it does not have the same phase to the voltage at the output of the inverter. The storage battery will power supply consumers in commuting time. This approach saves energy, which is supplied by the electric power distribution company, hence energy from renewable energy sources will be used and when there is a voltage from power distribution grid. The requirement for supply from alternative sources is to produce enough energy to power supply consumers. Lack of this concept is that the lack of energy from renewable energy sources cannot be recovered from the electric power distribution grid. Telecommunication facility is supplied from the electric power distribution grid or from renewable energy sources. There is no possibility to operate in parallel.

The advantage of the processes described is a simple realization. The use of energy sources based on fossil fuels is reduced, but it has many disadvantages. First of all, it has a low degree of efficiency. Inverters are devices that generate DC voltage from the input AC voltage. Voltage conversion is not lossless. It consumes the energy of renewable energy sources. Then rectifying unit generates DC voltage from AC voltage at the output of the inverter and thereby also consumes some energy. Any conversion of energy consumes a part of the energy from the renewable energy sources. Overall efficiency of energy use of renewable energy sources (1) is the product of efficiency of individual converters (2). The efficiency of renewable energy sources can be calculated as the ratio of power required to telecommunications equipment and the total power of renewable energy sources (1).

$$P_{OI} = P_{VG} + P_{SP} + P_{GC}. \quad (1)$$

$$\eta_{OI} = P_{TU}/P_{OI} = P_{TU}/P_{OINV} * P_{OINV}/P_{OSS} * P_{ISS}/P_{OI}$$

$$= \eta_{ISP} * \eta_{INV} * \eta_{SS}, \quad (2)$$

where in:

P_{TU}, - Power required for operation of telecommunications devices,

P_{OI}, - Available capacity of renewable energy sources,

P_{ISS}, - Power from the output of serial stabilizer,

P_{OINV}, - Power at the output of the inverter,

η_{INV}, - Inverter efficiency,

η_{SS}, - Serial stabilizer efficiency,

η_{ISP}, - Rectifying unit efficiency.

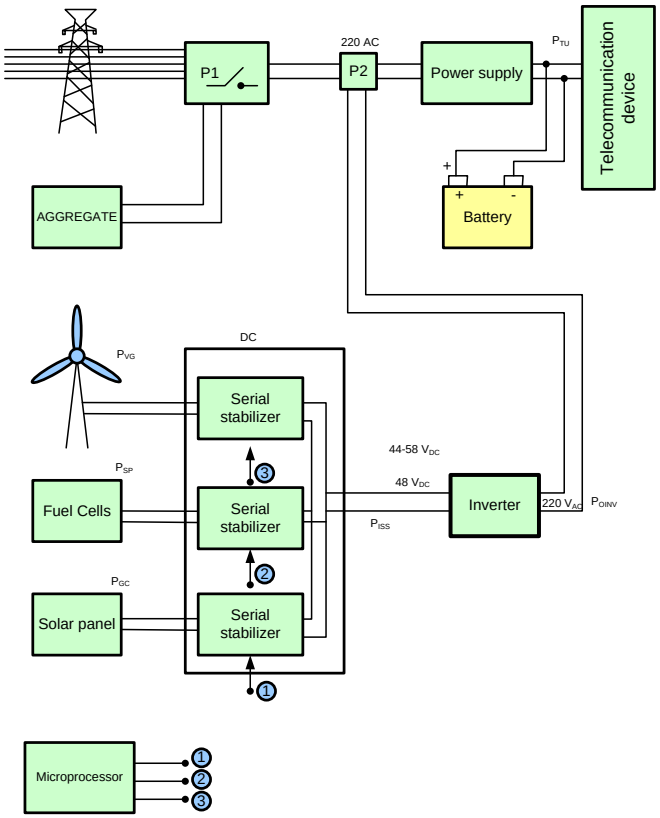


Figure 6. Using of inverter for generating an AC voltage from voltage of renewable energy sources

By choosing high-quality serial stabilizer configuration the effectiveness of 95% may be achieved. It is not possible to achieve such a performance when using inverter. The efficiency of inverters for which a condition of galvanic separation satisfied (necessary in relation to telecommunications equipment) is less than 90%, realistic about 80%. Efficiency of rectifying unit according to applicable telecommunications regulations cannot be less than 90%. The total efficiency of this solution is the product of the efficiencies of all three devices, and it is smaller than 75%. So, we should strive for solutions with minimum number of conversions. [6]

B. Distribution at the DC voltage level

Qualitative solution for connecting various sources is shown in Figure 7. Renewable energy sources are connected in parallel with rectifying unit. Thus, the distribution of energy is at the DC level. Between telecommunications devices and energy sources, there is only one converter - serial stabilizer. Thus, with this configuration maximum efficiency of use of energy of renewable energy sources is achieved.

The parallel operation of renewable energy sources and energy sources based on fossil fuels is enabled. The same microcomputer can control serial stabilizer and rectifying unit. Microcomputer should be programmed so that the voltage at the output of the serial stabilizers is greater than the voltage at the output of rectifying unit. Telecommunications device will be powered from renewable energy sources. If the amount of energy from renewable energy sources is insufficient to supply telecommunications equipment (a measure is the value of the output voltage), a microcomputer will increase the voltage of rectifying unit. In addition, the current value of all energy sources is controlled. By changing the value of the reference voltage, it is possible to realize parallel operation of both types of energy sources. It should seek to make maximum use of renewable energy sources, and minimum energy of sources based on fossil fuel and/or electric power distribution grid. [6]

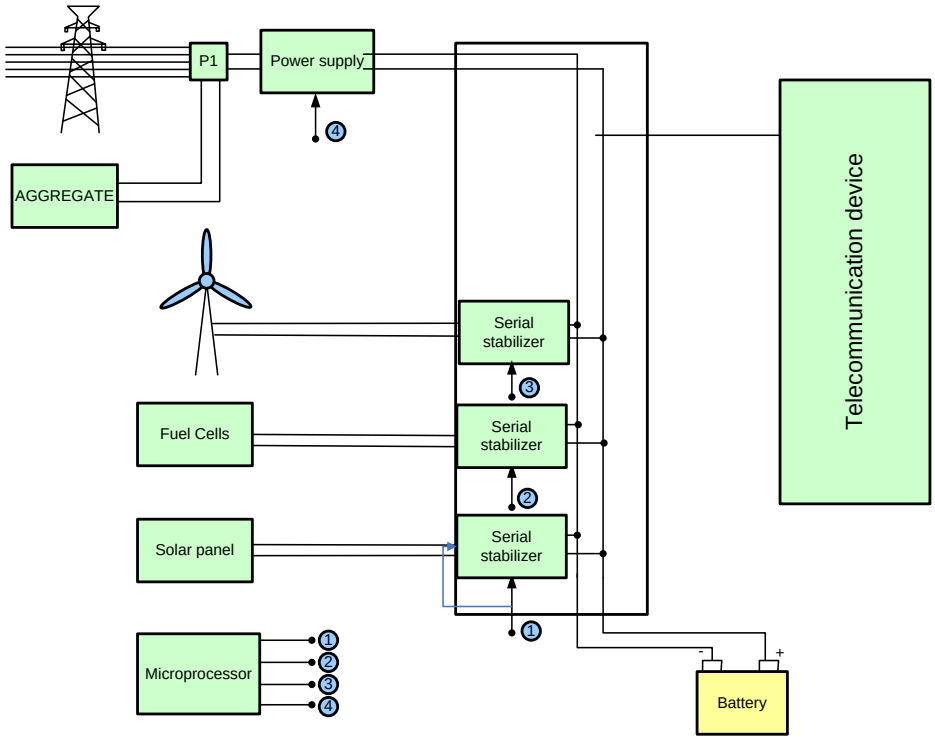


Figure 7. The distribution of energy of the DC level

IV. THE SYSTEM OF POWER ELECTRONIC DEVICES

In the previously described concept, a common feature of all power electronic devices is to provide uninterrupted power supply of telecommunications device. The power sources that are available at the moment of decision will be used. Nevertheless, additional criterion could be included as for instance the lowest price of energy. Therefore, there has to be coordination between power electronic devices. The choice of energy source which powers consumers depends on the availability of energy resources and the needs of consumers. Thereby, the only energy source that is completely controlled is storage battery. The presence of voltage of power distribution grid and energy of alternative sources are not under the control of power electronic devices. Because of that, the batteries are heart of the system. However, chemical energy sources are sensitive to external conditions; their capacity depends on the temperatures, and they have to be maintained by well-defined procedures. This means that it is necessary to develop special power electronic devices that would control the storage battery. Starting from the described concept, three groups of devices were developed:

- voltage converters (rectifying modules and serial converters)
- voltage distribution (DC and AC)
- control devices for batteries

All devices that are used to provide uninterruptable power supply are the parts of the system of power electronic devices. Each unit of power electronics have to be manageable. That basically means that it must have a microcontroller. In addition to measuring the parameters of energy sources microcontrollers control an operation of the voltage inverter. They can include it or exclude it as needed. In this way, the energy source for supplying consumers is defined. Of course, in such a way parallel operation of the various energy sources is enabled. Thereby, the power sources which operate in parallel do not have to provide the same amount of energy. The goal is to set maximum power from the alternative sources at their disposal at any given moment. Figure 8 shows the block diagram of the organization of uninterruptable power supply of mobile telecommunication center. The solution was implemented and characteristics are tested. The roles of some elements are:

- Programmable AC distribution connects the individual consumer with voltage from power grid. Connection is through a switch and, if necessary, the link with power grid could be disconnected. Practically, if there is enough energy in alternative sources, consumers will be disconnected from the power grid and they will be power supplied from alternative energy sources. Of course, it is possible to use the combined solution - the part of consumers uses power grid and the other part uses alternative sources.
- Rectifying modules are controllable AC/DC converters. At the implemented solution nominal value is 24V. Rectifying modules power supply telecommunication devices in the case when the voltage from power grid is present. Rectifying modulus -charger has the same configuration as the other rectifying modules, but it is set to charge the battery optimally from the power grid

- Programmable serial converter controls alternative source and on the basis of the measured parameter values sets the output voltage of the converter in order to maximize use of alternative energy sources. Practically, it is non-insulated DC/DC converter controlled by a microcomputer.
- DC distributions provide uninterrupted power supply of telecommunications equipment. Possible sources of energy are electric power distribution grid and storage battery. It is possible that the modules operate in parallel.
- A system for controlling the storage battery is a circuit that controls the correctness of each cell of the storage battery. It has the ability to correct the value of the voltage of each cell and thus it allows optimal charging and discharging of the storage battery.

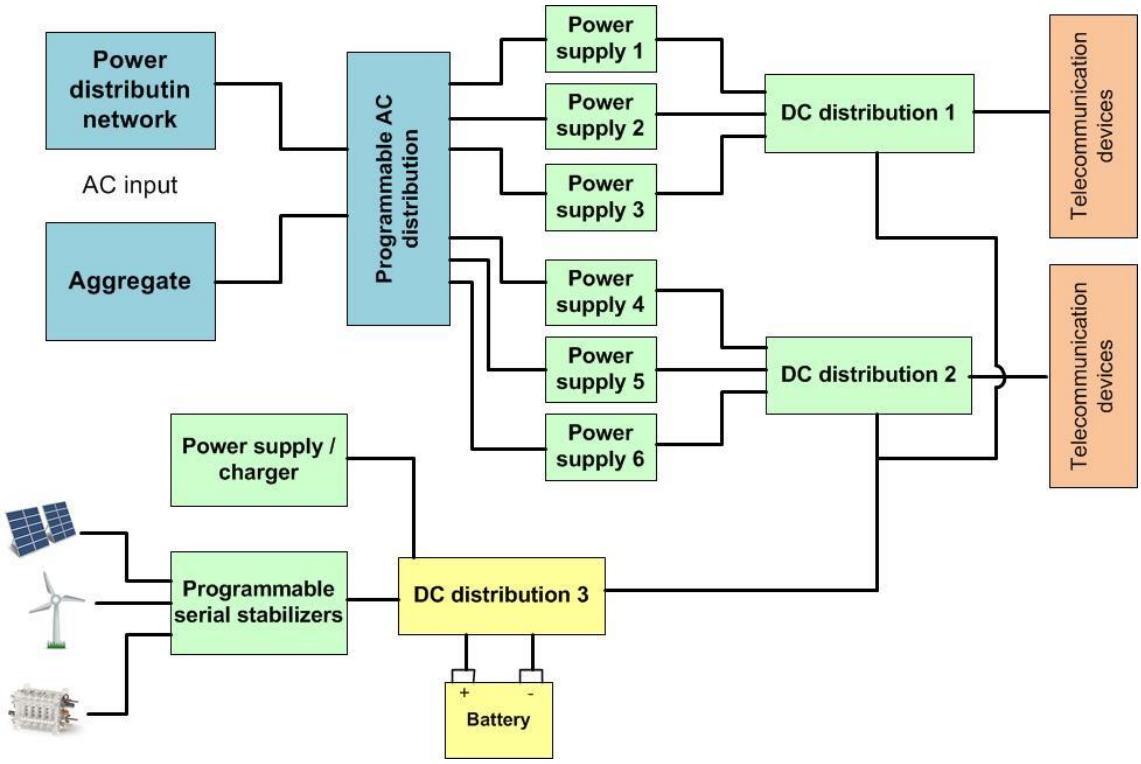


Figure 8. Organization of the power supply of mobile telecommunication center

The coordination of operation of the microcontrollers is accomplished through a microcontroller system. Figure 9 shows how microcomputers of power electronic devices are connected in the solution of uninterruptable power supply of mobile telecommunications center.

The main program is installed on a personal computer. Through the same computer remote monitoring is organized as well as the control of power electronic devices. Thus, PC is at the top of the hierarchy of power electronics system. Via Ethernet computer communicates with the distributions. Power converters communicate with distributions via RS 485 interface. The power converters collect data on energy sources and consumer needs. The data are processed in the distributions and forward to the computer. Based on the collected data and the pre-set algorithm, the computer adjusts which power sources will be used.

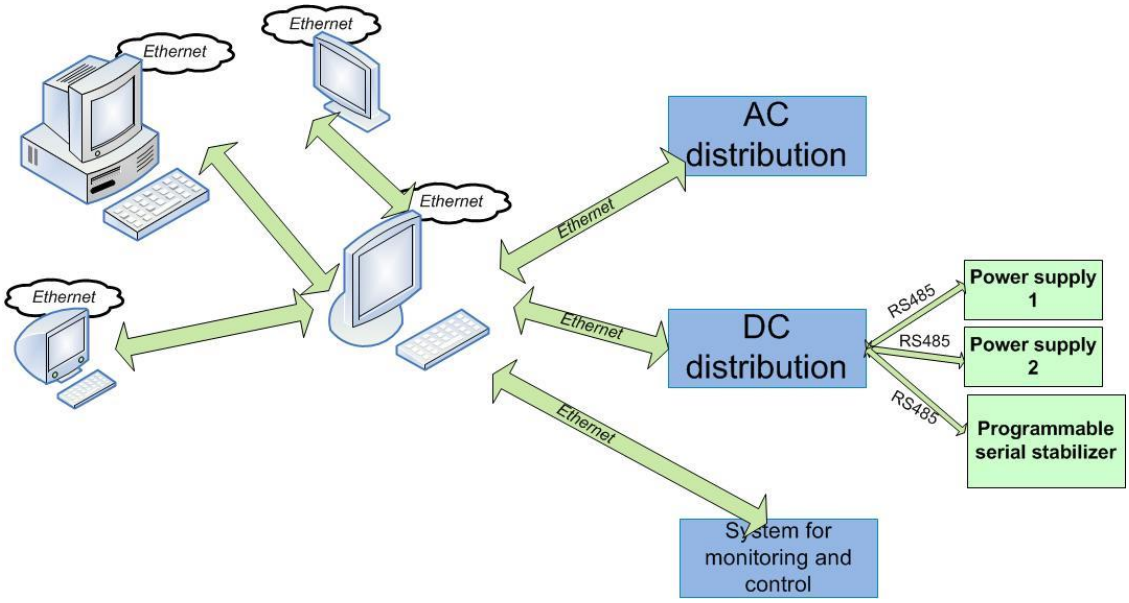


Figure 9. Microcomputers connecting of the power supply of mobile telecommunication center

CONCLUSION

The paper describes the solution implemented for power supplying mobile telecommunication facility. The power electronic devices are connected to each other and together form a system of power electronics. Since microcontrollers have become integral elements of almost every power electronic device a solution n that allows a reliable and uninterruptable operation of the telecommunications device with the maximum use of alternative sources was created. The procedure described is called "DC distribution". In the method described the efficiency of the use of alternative sources increases more than 30%. The storage battery was chosen as basic source of energy. In this way, the operation autonomy of the device in the case of absence of power sources is correctly defined. Also, the solution described allows parallel operation of alternative energy sources and electric distribution power grid.

Until now, a model of the solution described was realized. In the next years it is expected a production of zero series, and later serial production of procedure of the uninterrupted supply of telecommunication facilities.

ACKNOWLEDGMENT

This paper is part of a project supported by the Ministry of Education and Science of the Republic of Serbia, under the heading TR32016.

REFERENCES

[1] D.G. Holmes, T.A. Lipo, "Pulse Width Modulation for Power Converters", IEEE Press, Wiley Interscience, Piscataway, 2003.

[2] Dragana Petrović, Miroslav Lazić, Zoran Cvejić, Bojana Jovanović, Dragan Jekić, Organization of uninterrupted power supply of mobile telecommunication facilities ISBN 978-99938-624-2-8, ENS 1-12, Vol. 13, p. 118-121 Infoteh Jahorina 19. mart - 21. mart 2014.

[3] Z. Čorba, V.A. Katić, M. Mišić, G. Vamović, "A Hybrid System Based on Renewable Energy Sources", 15th International Symposium on Power Electronics – Ee 2009, www.dee.uns.ac.rs, Novi Sad, Serbia, 28-30th Oct. 2009, CD-ROM, Paper No.T7-1.5

[4] Ph.Delarue, A.Bouscayrol, E.Semail, "Generic control method of multileg voltage-source-inverter for fast practical application," IEEE Trans. on Power Electronics, vol. 18, no. 2, 2003, pp. 517-526.

[5] B.Francois, A.Bouscayrol, "Design and modeling of a five-phase voltage-source inverter for two induction motors," Proc. European Power Electronics and Applications Conf. EPE, Lausanne, Switzerland, 1999, CD-ROM, Paper No.626.

[6] Miroslav Lazić, Dragana Petrović, Zoran Cvejić, Bojana Jovanović, Optimizing use of the alternative energy sources - DC distribution, ISBN 978-86-80509-70-9, 1st International Conference on Electrical, Electronic and Computing Engineering IcETRAN 2014, Vrnjačka Banja, Serbia, June 2 – 5, 2014. Invited paper. P.p. ELI1.1.1-6

PMSG Synchronization Control Algorithm based on the Active Damping Principle

Marko Vekić, Stevan Grabić, Evgenije Adžić, Zoran Ivanović, Vlado Porobić
Faculty of technical sciences
Novi Sad, Serbia
vekmar@uns.ac.rs

Abstract— In this paper the modified active damping law was used to damp oscillations during the synchronization process of PMSG cascade. Due to the imposed capacity limit of the converter and exclusive reactive power injection, the damping ability is limited. Boundaries of the area which defines possible synchronization depend on PMSG initial speed and the difference between the grid and PMSG angle at the moment of connection. Exhaustive tests by means of the state of the art HIL emulation were performed to prove both the modified damping law and its area of application.

Keywords-PMSG (Permanent Magnet Synchronous Generator), PMSG cascade, active damping, modified active damping, area of operation, boundaries,

I. INTRODUCTION

Active Damping Principle is proposed as the part of control algorithm strategy in various grid connected applications. The aim of active damping is very similar or the same as the purpose of passive damping structures (based on passive elements) – to damp oscillations during the transition period and improve stability. In contrast to passive dampers, active damping topology contain power electronics converters as active elements. It should be noted that active damping is much different from one similar structure - active filter. The purpose of active filtering is to suppress or eliminate undesired harmonics in the system, while the active damping primarily damps transient oscillations. Modus operandi of the active damping is to use the converter (say 3-phase inverter) and control it in such manner to emulate the "programmable" impedance, effectively increasing the system original damping level. The concept is tested in many applications. The active damping principle is employed in FACTS to improve the transient stability of the system in case of power flow fluctuations [1]–[6], and grid voltage disturbances [7]–[9]. Recently a number of papers propose the active damping to improve the stability of LLCL-filter based grid connected inverters [10] – [12].

In [13] the concept of permanent magnet synchronous generator (PMSG) cascade is proposed (Fig. 1) to actively damp the system subjected to the input power fluctuations (changes in the input load torque). This is essentially achieved through the series converter connected in the open winding of the PMSG which effectively modulates the overall reactance. In this paper the similar approach with modified algorithm is employed to help the PMSG direct connection to the grid. The results are verified using state – of – the art Hardware In the Loop (HIL) platform [14].

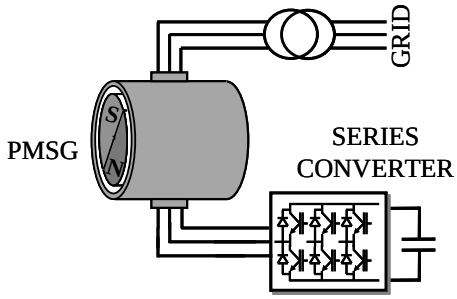


Figure 1. PMSG Cascade

II. ACTIVE DAMPING ALGORITHM FOR PMSM CASCADE

The main components of a PMSG cascade system (see Fig. 1) are: a grid-connected PMSG and a small series converter (20% of rated power) in the star point of the open winding PMSG) [13].

In Fig. 2, the synchronous generator is modeled as a three phase voltage source u_{EMF} in series with synchronous inductance x_S, while the grid and the series converter are modeled as two additional three-phase voltage sources, u_{GRID} and u_{SC}, respectively.

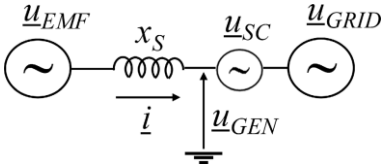


Figure 2. Simplified model of PMSG Cascade

Power flow through the system is controlled by injecting reactive voltage u^r_{SC} whose phase is shifted 90◦ with respect to the line current vector. The additional requirement is that injected reactive voltage component u^r_{SC} should not exceed 0.2 [p.u.].

The grid is assumed to be ideal (the infinite power) where U_{GRID} = 1 [p.u.], while the PMSG rotor speed varies only slightly around its nominal value, i.e. n = 1 [p.u.]. It also implies U_{EMF}=1. So simplified system can be described with following equations:

$$m_{el} = \frac{\sin\theta}{x_S}\left(1 - \frac{u^r_{SC}}{\sqrt{2\cdot(1-\cos\theta)}}\right) \quad (1)$$

$$\tau_{mec}\frac{dn}{dt} = M_m - m_{el} \quad (2)$$

$$\frac{1}{\omega_{grid}}\frac{d\theta}{dt} = n - 1 \quad (3)$$

The first part of (1) is well known expression which describes (in [p.u.]) the PMSG with rounded rotor connected to the ideal grid. Generally this equation defines the power flow between two voltage sources connected through the reactance (x_s here). However, the second part of the equation depends on the injected voltage of series converter, u^r_{SC}. Controlling this voltage, the electrical torque m_{el} can be modulated and hence, this modulated torque could bring the damping ability. Using the electro-mechanical analogy, it can be shown [13] that the voltage u^r_{SC} should be generated as follows:

$$u^r_{SC} = -k_{dmp}\cdot\frac{d\theta}{dt}\cdot\mathrm{sgn}(\theta) \quad (4)$$

where k_{dmp} is the damping factor and θ is the angle difference between the grid voltage and induced PMSG electromotive force. After the start of synchronization this angle difference is usually called the power angle. The upper equation is actually the variation of well-known spring law applied to electromechanical phenomena [15]. From (1) and (4) it can be obtained:

$$u^r_{SC} = -k_{dmp}\cdot\frac{d\theta}{dt}\cdot\mathrm{sgn}(\theta) = -k_{dmp}\cdot\omega_{GRID}\cdot(n-1)\cdot\mathrm{sgn}(\theta) \quad (5)$$

which is basically the core of the control law for damping oscillations caused by the input (mechanical) torque changes.

III. SYNCHRONIZATION

A. Introduction

It is well known that for successful PMSG connection to the grid, appropriate grid voltage and emf vectors (magnitudes, angles and frequencies) should match. Even when voltage magnitudes ($U_{grid} = U_{gen}$) and frequencies ($n_{grid} = n_{gen}$) are the same, but not the phase angle ($\theta_{grid} \neq \theta_{gen}$), oscillations will occur as simulation results clearly confirm (Fig. 3). Of course these results are completely expected. Parameters of the simulated machine are given in [16]. The overall damping level of the system is even worse when we speak about the high power machine with relatively small stator resistance.

B. Modified active damping principle

The process of synchronization to the grid when voltage vectors are different essentially resembles the case when grid connected PMSG suffers load disturbances.

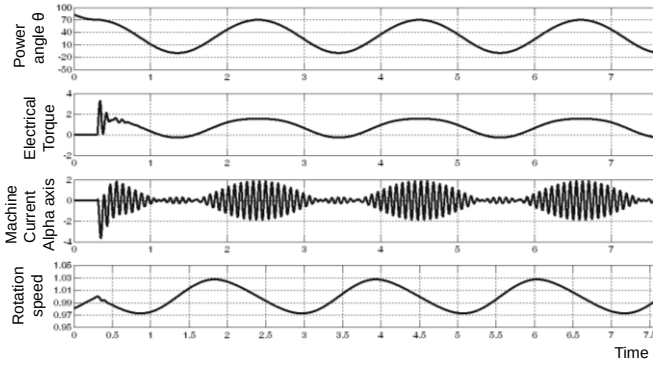


Figure 3. Synchronization attempt when the difference in phase angles is $\theta = 70°$ and $n_{grid} = n_{gen}$

The both processes include oscillations. Therefore, the idea to employ the active damping algorithm described in the section II. is actually very intuitive. However, the described control law (5) has to be modified, because in the case of the synchronization it cannot be assumed that PMSG speed is nominal, i.e. $n \neq 1$ [p.u.]. Therefore, PMSG speed appears in the further consideration as a variable which also implies variable $U_{EMF} = n\cdot\psi$, where ψ is permanent magnets flux. Following these modifications, the equation (1) becomes:

$$m_{el} = \frac{\psi}{n\cdot x_s}\cdot\sin\theta - \frac{\psi}{n\cdot x_s}\cdot\sin\theta\cdot\frac{U^r_{SC}}{\sqrt{1+n^2\cdot\psi^2-2\cdot n\cdot\psi\cdot\cos\theta}} \quad (6)$$

Accordingly, the control law accordingly is given by:

$$U^r_{SC} = -\frac{a}{2}\cdot n\cdot(n-1)\cdot\frac{\sqrt{1+n^2\cdot\psi^2-2\cdot n\cdot\psi\cdot\cos\theta}}{\sin\frac{\theta}{2}} \quad (7)$$

In the upper equation a is the damping coefficient instead of k_{dmp} from (5). The torque expression (6) now becomes:

$$m_{el} = \frac{\psi}{x_s}\cdot\frac{\sin(\theta)}{n} + a\cdot\frac{\psi}{x_s}\cdot(n-1)\cdot\cos\left(\frac{\theta}{2}\right) \quad (8)$$

In order to complete the control law, the coefficient a has to be determined. The control structure is given at the Fig. 4.

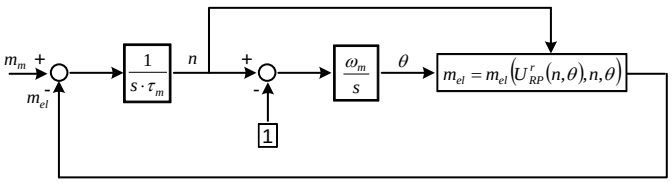


Figure 4. Synchronization control scheme

Although this control diagram looks pretty simple, the block which represent torque calculation is non-linear as (8) clearly shows. Therefore, in order to obtain coefficient a the linearization around the operation point has to be performed. It is justified to adopt the PMSG nominal value ($n = 1$ [p.u.]) for operation point because the process of synchronization should start when the speed approaches grid (nominal) speed. The linearized control diagram is presented on Fig. 5.

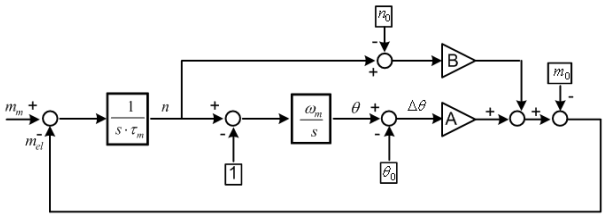


Figure 5. Linearized synchronization control scheme

The characteristic equation derived from Fig. 5 is the second order equation:

$$f(s) = s^2 + s \cdot \frac{B}{\tau_m} + \frac{A \cdot \omega_m}{\tau_m} \tag{9}$$

where: $A = \frac{\partial m_{el}}{\partial \theta}_{\theta_0, n_0}$, $B = \frac{\partial m_{el}}{\partial n}_{\theta_0, n_0}$ (10)

and m_{el} is defined by (8).

Observing (9), the coefficient ξ and natural frequency ω can be easily obtained from:

$$2 \cdot \xi \cdot \omega_n = \frac{B}{\tau_m} \text{ и } \omega_n^2 = \frac{A \cdot \omega_m}{\tau_m} \tag{11}$$

i.e: $\xi = \frac{B}{2 \cdot \sqrt{A \cdot \tau_m \cdot \omega_m}}$ (12)

Finally, from (10) and (12) desired damping coefficient a can be calculated as the function of ξ, and operation point θ_0, n_0 (the initial angle difference and PMSG speed value at the very moment of synchronization respectively). The operation point (θ_0, n_0) itself depends on the input mechanical torque M_m which accelerates PMSG. At the Fig. 6 dependency of the coefficient a as the function of the driving input torque in the full range of M_m from 0.1 to 1 [p.u.] is presented.

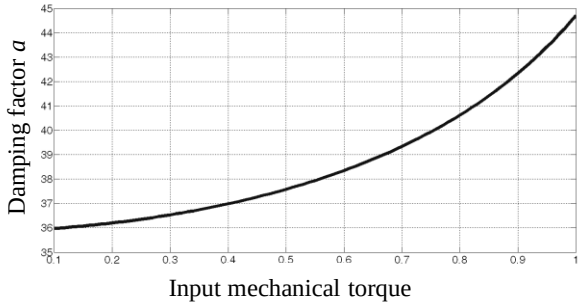


Figure 6. Damping factor a

From the upper diagram we can see that the coefficient a varies from 35.98 to 44.71 in the whole driving torque range. Therefore, we can adopt the average value, i.e. a = 38.7 in further calculations. Results presented later will justify such choice.

C. Synchronization feasibility

Let us apply the modified active damping law in the case from Fig. 3. The new situation is illustrated at the Fig. 7.

Obviously, the synchronization is accomplished successfully. Oscillations in the power angle, PMSG currents, speed and torque fade away completely after 2 s, while the transient process is properly damped.

However, keeping in mind that the capacity of the series converter is limited to 0.2 [p.u.], i.e:

$$\left|u_{sc}^r\right| \le 0.2\left[p.u.\right] \tag{13}$$

it is clear that the synchronization cannot be achieved in all instances, i.e. for all values of initial PMSG speed n_{init} and angle difference θ_{init}.

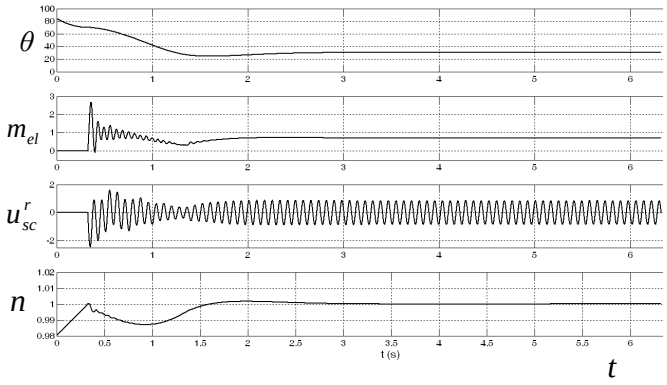


Figure 7. Synchronization attempt when the difference in phase angles is θ_{init} = 70°, n_{grid} = n_{gen} and modified active damping law applied

Indeed, if the simulation is performed under the assumption that θ_{init} = - 150°, the situation is much different as shown at Fig. 8.

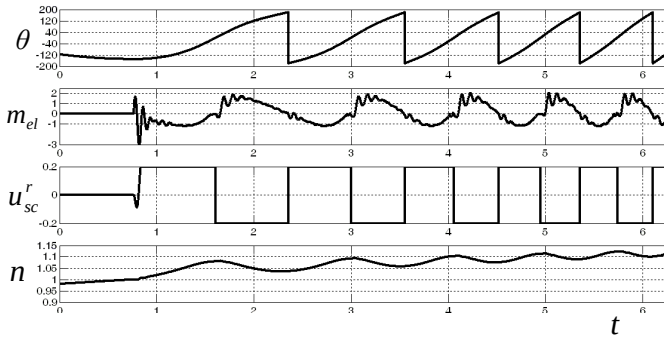


Figure 8. Synchronization attempt when the difference in phase angles is θ_{init} = 150°, n_{grid} = n_{gen} and modified active damping law applied

Obviously, the synchronization was not successful. That means the application of the proposed modified active damping law has certain limits which are the consequence of (13).

D. Modified active damping law application area

Boundaries of the area where the modified active damping control principle is applicable have to be determined. For that purpose, a detailed mathematical and simulation model are developed [17] in order to find the range of the input mechanical torque M_m, initial angle difference θ_{init} and speed n_{init} for which the synchronization is possible. The simulation results are shown at the Fig. 9.

Results from the Fig. 9 are obtained under assumption that the initial PMSG speed at the moment of synchronization was n_{init} = 0.98 [p.u.], while the power angle (the difference between the grid and PMSG angle) and input mechanical torque are arbitrary values. It can be concluded that synchronization is possible for any value of the mechanical input torque in the range of [0.1,..,0.9, 1] when the initial power angle θ_{init} is in the range of -30º to 100º. These boundaries are essentially the consequence of the limited series converter voltage (13).

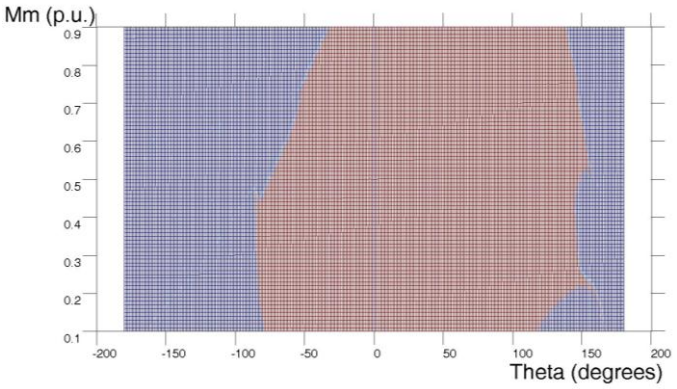


Figure 9. The area where synchronization is possible

After the area is obtained by means of the detailed simulation, we need to perform exhaustive testing in order to verify the proposed modified active damping law.

IV. EXPERIMENTAL RESULTS

In order to check the control law and its borders, we need to perform tests in all operation points, particularly on area border (Fig. 9) itself and a little beyond. Testing with the real hardware, even when using small power test bench would be dangerous and too laborious because we need to test the grid tied PMSG application. A very efficient alternative would be real time, hardware in the loop (HIL) emulation with extremely high simulation step, i.e. overall latency of 1 μs [14].

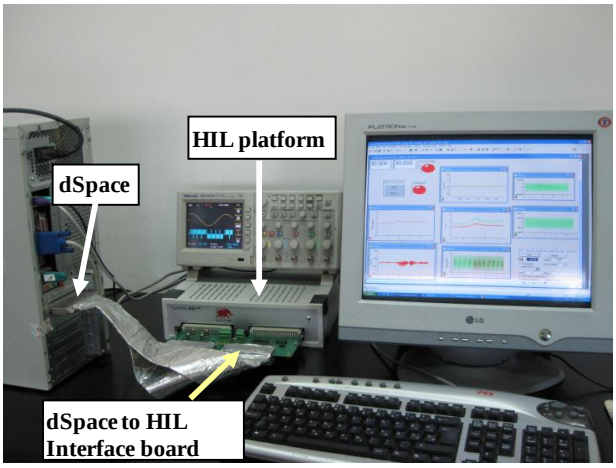


Figure 10. Experimental setup

Fig. 10 shows the experimental setup for evaluation of the proposed synchronization algorithm. *Dspace dS1104* is employed as the controller stage (Fig.5), while *Typhoon HIL400* platform is used to emulate the power stage (Fig. 1). Tests were performed for nine cases of input mechanical torque: M_m =(0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0) [p.u.]. For each case, three operation points were chosen regarding initial power angle θ_{init} – the upper boundary, the lower boundary and region between them. Here, only several characteristic results are shown.

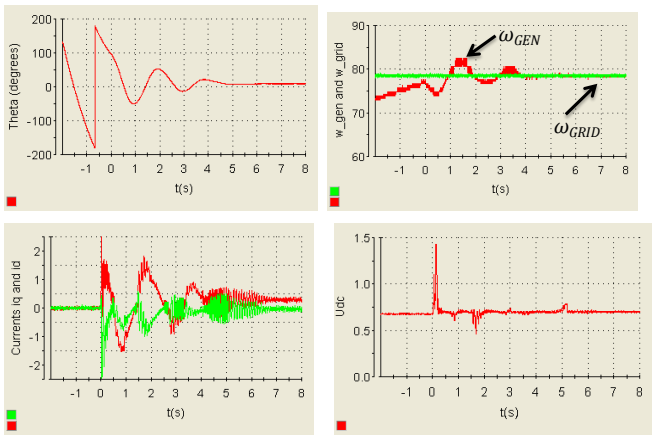


Figure 11. PMSG Synchronization when M_m=0.3 [p.u] and θ_{init}=115º

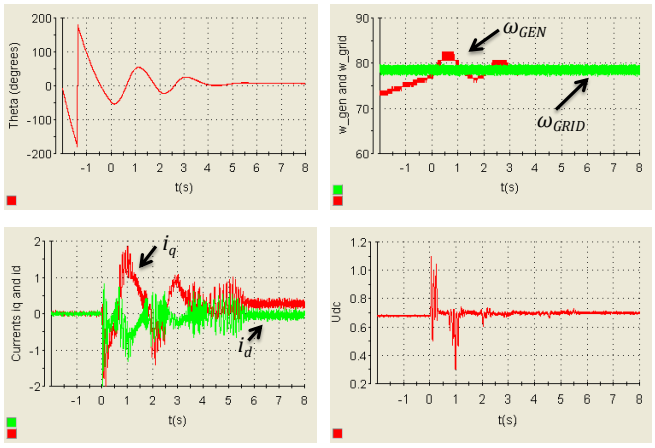


Figure 12. PMSG Synchronization when M_m=0.3 [p.u] and θ_{init}=-45º

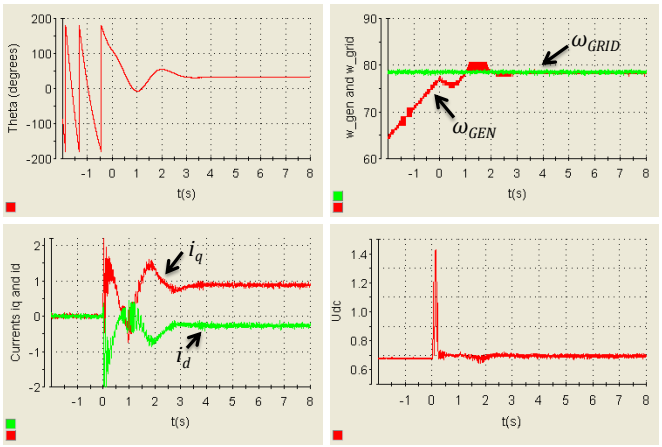


Figure 13. PMSG Synchronization when M_m=0.9 [p.u] and θ_{init}=120º

At displayed figures, ω_{GEN} represents an electrical speed of PMSG, ω_{GRID} is grid frequency, i_d and i_q are d and q components of the current in synchronous rotating frame where i_q is the current component proportional to the electrical torque.

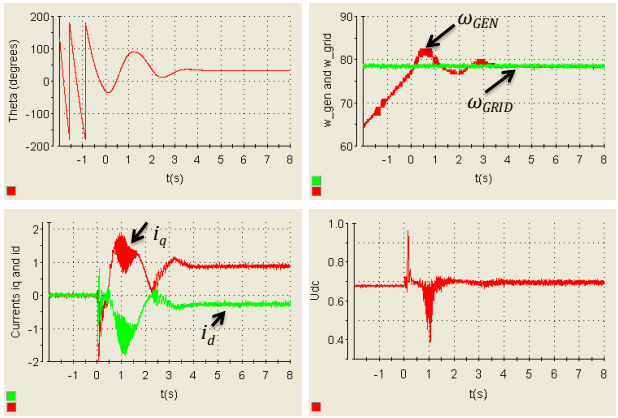


Fig. 14. PMSG Synchronization when M_m=0.9 [p.u] and θ_{init}=-35º

The upper results represent two characteristic cases: small input driving torque (M_m = 0.2 [p.u.]) and big driving torque (M_m = 0.9 [p.u.]) in two extreme cases; at the lower and upper border of the area (Fig. 9). From Fig. 11 – Fig.14 it can be concluded that synchronization is performed successfully.

At the Fig. 15 results are presented when the operation point is deeply into the safe area.

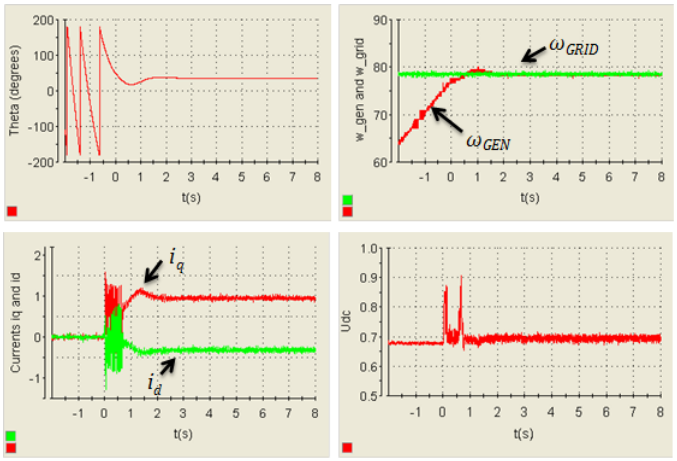


Figure 15. PMSG Synchronization when M_m=1.0 [p.u] and θ_{init}= 45º

According to expectation, the system response is very clear and smooth.

Further investigation was performed for cases when θ_{init} belongs to the area well beyond designated borders (Fig. 15).

From the figure bellow it is clear that the PMSG connection to the grid is not possible when the initial angle difference is significantly displaced from the pre-defined boundaries. However, the PMSG connection to the grid is still possible slightly beyond borders of the area defined by Fig. 9. In the purely theoretical scenario where no limits are imposed to the voltage injected by series converter, the synchronization would be possible for any instant in the range of θ_{init} = [-180°, …, 180°].

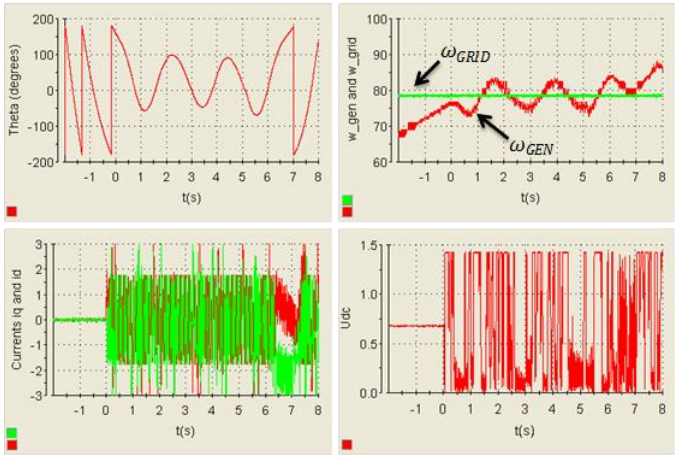


Figure 16. PMSG Synchronization when M_m=0.6 [p.u] and θ_{init}=160 º

V. CONCLUSIONS

The active damping principle is used successfully to cope with transient oscillations in grid-tied applications. In this paper the modified algorithm for PMSG cascade was employed to help PMSG direct connection to the electrical grid. It is assumed that converter injects only reactive voltage component equal to 0.2 [p.u.]. Therefore, the validity of algorithm is limited by initial synchronization conditions: angle difference, PMSG speed and driving input torque. The area of successful connection to the grid is determined using simulation and exhaustively tested by means of HIL emulation in all operation points of interest. The future research should deal with further modifications in the damping law principle.

ACKNOWLEDGMENT

This research was partially co-funded by the Ministry of Education, Science and Technological Development of Republic of Serbia under contract No. III 042004 and by the Provincial Secretariat for Science and Technological Development of AP Vojvodina under contract. No. 114-451-3508/2013-04.

REFERENCES

[1] F. A. L. Jowder, "Influence of mode of operation of the SSSC on the small disturbance and transient stability of a radial power system," *IEEE Trans. Power Syst*., vol. 20, no. 2, pp. 935–942, May 2005.

[2] B. T. Ooi, M. Kazerani, R. Marceau, Z. Wolanski, F. D. Galiana, D.McGillis, and G. Joos, "Mid-point sitting of facts devices in transmission lines," IEEE Trans. Power Delivery, vol. 12, no. 4, pp. 1717–1722, Oct. 1997.

[3] M. Noroozian, L. Angquist, M. Ghandhari, and G. Andersson, "Improving power system dynamics by series-connected facts devices," *IEEE Trans. Power Delivery*, vol. 12, no. 4, pp. 1635–1641, Oct. 1997.

[4] L. Gyugyi, C. D. Shauder, S. L. Williams, T. R. Rietman, D. R. Torgerson, and A. Edris, "The unified power flow controller: A new approach to power transmission control," *IEEE Trans. Power Delivery*, vol. 10, no. 2, pp. 1085–1097, Apr. 1995.

[5] F. A. Rahman and B. T. Ooi, "VSC-HVDC station with SSSC characteristic," IEEE Trans. Power Electron., vol. 19, no. 4, pp. 1053–1059, Jul. 2004.

[6] M. Noroozijan and G. Andersson, “Power flow control by use of controllable series components,” *IEEE Trans. Power Delivery*, vol. 8, pp. 1420–1429, Jul. 1993, no.3.

[7] U. Gabriel and R. Mihalic, “Direct methods for transient stability assessment in power systems comprising controllable series devices,” *IEEE Trans. Power Syst*., vol. 17, no. 4, pp. 1116–1122, Nov. 2002.

[8] B. S. Rigby, N. S. Chonco, and R. G. Harley, “Analysis of power oscillation damping scheme using a voltage-source inverter,” *IEEE Trans. Ind. Appl*., vol. 38, no. 4, pp. 1105–1113, Jul./Aug. 2002.

[9] M. Poshtan, B. N. Singh, and P. Rastgoufard, “A nonlinear control method for SSSC to improve power system stability,” in *Proc. Int. Conf. Power Electron., Drives Energy Syst. (PEDES'06)*, 2006, pp. 1–7.

[10] J. Xu, S. Xie, T. Tang, “ Active Damping-Based Control for Grid-Connected LCL-Filtered Inverter With Injected Grid Current Feedback Only”, IEEE Trans. Ind. Electronics, vol. 61, no. 9, pp. 4746-4758, Sept. 2014.

[11] J.C. Wiseman, B. Wu, “Active Damping Control of a High-Power PWM Current-Source Rectifier for Line-Current THD Reduction”, IEEE Trans. Ind. Electronics, vol. 52, no. 3, pp. 758-764, June. 2005.

[12] X. Wang, F. Blaabjerg, M. Liserre, Z. Chen, J. He, Y. Li, “An Active Damper for Stabilizing Power-Electronics-Based AC Systems”, IEEE Trans. Power Electron., vol. 29, no. 7, pp. 3318–3329, Jul. 2014.

[13] S. Grabic, N. Celanovic, V. A. Katic, “*Permanent Magnet Synchronous Generator Cascade for Wind Turbine Application”,* IEEE Trans. on Pow. Electron. vol.. 23, no. 3, May 2008, pp. 1136-1142.

[14] M. Vekic, S. Grabic, D. Majstorovic, N. Celanovic, V. Katic, “Ultralow Latency HIL Platform for Rapid Development of Complex Power Electronics Systems”, IEEE Trans. on Pow. El. vol.. 27, no. 11, Nov 2012, pp. 4436-4444.

[15] Grabic, N. Celanovic, and V. Katic, “Series converter stabilized wind turbine with permanent magnet synchronous generator,” in Proc. IEEE 35th Annu. Power Electron. Spec. Conf., Aachen, Germany, 2004, pp. 464–468.

[16] Marko Vekic, Milan Rapaic, Stevan Grabic Nikola Celanovic, VladimirKatic, "HIL Evaluated New Control Algorithm For PMSG Grid Connection", 17th International Symposium on Power Electronics - Ee 2013, Novi Sad, Serbia, October 30th-November 1st, 2013 Paper No. Sps-1.5, pp 1-5, ISBN: 978-86-7892-550-4

[17] Marko Vekic, “New real time emulation based procedure for Power Electronics controllers development”, PhD thesis, Faculty of technical sciences, Novi Sad, February 2014, in Serbian.

Conduction Modes of a Peak Limiting Current Mode Controlled Buck Converter

Predrag Pejović and Marija Glišić

***Abstract*—In this paper, analysis of a buck converter operated applying a peak limiting current mode control is performed, focusing regions where the limit cycle is unstable. Normalized discrete time converter model is derived. Chart of operating modes is presented, and it is shown that the converter exhibits an infinite number of discontinuous conduction modes in an area where the continuous conduction mode would be expected assuming stable limit cycle. The converter is analyzed applying numerical techniques to determine period number of different discontinuous conduction modes and dependence of the output current on the output voltage and the limiting current. The numerical results agree with the analytical results in areas where the limit cycle is stable, and differ in regions where the limit cycle is unstable. Two different notions of stability, the limit cycle stability and the converter open loop stability, are clarified.**

***Index Terms*—Buck Converter, Current Mode Control, Discontinuous Conduction Mode.**

I. Introduction

Buck converter, shown in Fig. 1, operated applying the peak limiting current mode control is analyzed in this paper, aiming proper understanding and modeling of regions characterized by unstable limit cycle. Analytical models of buck converters are presented in [1], where both the continuous and the discontinuous conduction mode are analyzed, as well as instability of the limit cycle in the continuous conduction mode for the output voltage greater than one half of the input voltage. The results presented there indicate open loop instability of the converter in the discontinuous conduction mode for the output voltage greater than one half of the input voltage, although the limit cycle should be stable in this case. Comprehensive presentation of [1] aggregates results of many previous publications. Although the topic is more than 30 years old [2], it occasionally attracted attention over decades [3-7]. Detailed analysis is presented in [6], where fairly general case is analyzed, resulting in equations that are sometimes hard to follow. Discrete time model of the converter and bifurcations are analyzed in [7]. In [8], open loop instability of the converter operated in the discontinuous conduction mode for the output voltage greater than one half of the input voltage is analyzed, and as an auxiliary result multiple period discontinuous conduction modes are noticed. Besides, such waveforms have been occasionally observed in practice, but remained without a deeper analysis. This paper is a continuation of [8], focusing regions of unstable limit cycle operation.

Purpose of this paper is to clarify operation of the peak limiting current mode controlled buck converter in regions where the limit cycle is unstable in an easy-to-follow manner. To achieve this goal, both analytical and numerical techniques are applied. Analysis of the buck converter is reduced to analysis of a switching cell. The circuit voltages, currents, and time are normalized to generalize the results. A program to analyze the switching cell operation is given, and the results obtained using the program are presented, analyzed and discussed. It is shown that the converter exhibits an infinite number of period-n discontinuous conduction modes, and boundaries of the region where such modes exist are analytically derived. Regions where period-n discontinuous conduction modes occur are determined numerically and presented for $n \leq 10$. Dependence of the converter output current on the output voltage and the limiting current is computed numerically, and the results are equal to theoretical predictions in the regions where period-1 operation occurs, while different elsewhere. Open loop stability of the converter, being a concept different than the limit cycle stability, is analyzed numerically, and unstable regions are identified.

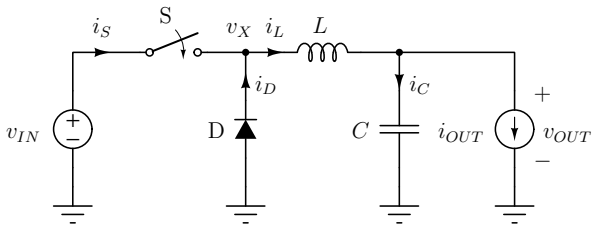


Fig. 1. Buck converter.

II. Reduction to a Switching Cell Model

A. Approximation

To simplify the analysis and to focus attention to the part of the circuit that causes the complex behavior, let us assume that the converter output voltage is constant. This assumption at least has to hold over one switching period, to justify linear ripple approximation. The assumption results in a simplified equivalent circuit of the buck converter, as shown in Fig. 2, frequently named as a "switching cell".

B. Circuit Equations

After the converter is reduced to the switching cell introducing the approximation that the output voltage is constant, solving the circuit essentially reduces to determination of the inductor current waveform. Governing equation to determine the inductor current waveform is

$$L\,\frac{d\,i_L}{dt} = v_L \tag{1}$$

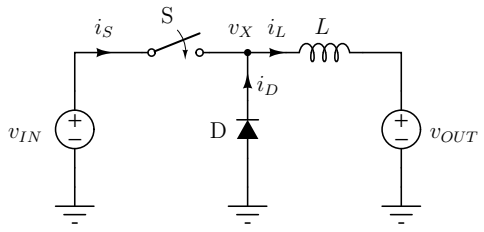


Fig. 2. Switching cell.

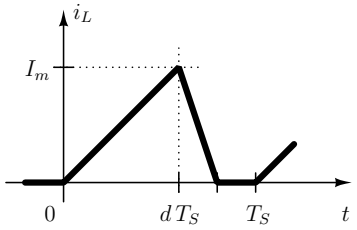


Fig. 3. Aimed waveform of i_L in the DCM.

which is a differential equation that requires an initial condition to provide a unique solution.

Voltage across the inductor depends on the states of the switching elements. For all three (out of four theoretically possible) combinations of states of switching elements that occur during the converter operation, voltages across the inductor are

$$v_L = \begin{cases} V_{IN} - V_{OUT}, & \mathrm{S-on, D-off} \\ -V_{OUT}, & \mathrm{S-off, D-on} \\ 0, & \mathrm{S-off, D-off} \end{cases} \tag{2}$$

which results in the waveform of the inductor current in the discontinuous conduction mode as depicted in Fig. 3. In the case of the discontinuous conduction mode, later to be labeled as period-1 discontinuous conduction mode, initial value of the inductor current at the beginning of the switching period is zero. The waveform consists of linear segments, since all of the voltages that might appear across the inductor are assumed as constant in time.

C. Normalization

The approach to analyze the switching cell in this paper is primarily based on numerical simulations. To generalize the results, normalization is performed. Choice of the normalization variables is typical, such that all of the voltages are normalized taking the input voltage as the base quantity,

$$m \triangleq \frac{v}{V_{IN}} \tag{3}$$

which results in $M_{IN} = 1$. Chosen notation is such that indexes of the voltages in general remain, while the quantity is labeled as m instead of v. The only exception from this rule is the output voltage, labeled just as M, according to

$$M \triangleq \frac{V_{OUT}}{V_{IN}}. \tag{4}$$

Currents are normalized using $V_{IN}/(f_S\,L)$ as the base quantity

$$j \triangleq \frac{f_S\,L}{V_{IN}}\,i \tag{5}$$

while the time intervals are normalized taking the switching period as the base quantity

$$\tau \triangleq \frac{t}{T_S}. \tag{6}$$

After the normalizations are performed, the differential equation that governs the inductor current reduces to

$$\frac{d\,j_L}{d\tau} = m_L \tag{7}$$

and the set of possible voltages across the inductor is

$$m_L = \begin{cases} 1 - M, & \mathrm{S-on, D-off} \\ -M, & \mathrm{S-off, D-on} \\ 0, & \mathrm{S-off, D-off} \end{cases}. \tag{8}$$

III. Discrete Time Model of the Switching Cell

Discrete time model of the switching cell is aimed towards obtaining the average the output current of the switching cell during one switching period (j_{OUT}), as well as the final value of the inductor current at the end of the switching period, $j_L(1)$, which is the initial value for the next switching period. Both of the variables are dependent on the initial value of the inductor current, $j_L(0)$, specified maximum of the inductor current when the switch turns off, J_m, and the output voltage, M. To perform the analysis, it is assumed that the output voltage of the switching cell, M, is constant during a switching period.

Operation of the switching cell during a switching period is analyzed following state changes of the switching elements in time, and thus related changes in the inductor voltage.

Each switching period begins with turning the switch on, and the switching cell enters the state when the switch is on, while the diode is off. Assuming this state throughout the switching period, the final value of the inductor current would be

$$j_X = j_L(0) + (1 - M)\,. \tag{9}$$

In the case $j_X \le J_m$, this completes the switching cycle, and

$$j_L(1) = j_X \tag{10}$$

while contribution of the switching period to the output charge, i.e. the average of the output current is

$$j_{OUT} = \frac{j_L(0) + j_X}{2}. \tag{11}$$

This situation is depicted in Fig. 4(a). In the case $j_X > J_m$, at

$$\tau_1 = \frac{J_m - j_L(0)}{1 - M} \tag{12}$$

the inductor current reaches J_m, and the switch is turned off. The switching cell enters state when the switch is off, and the

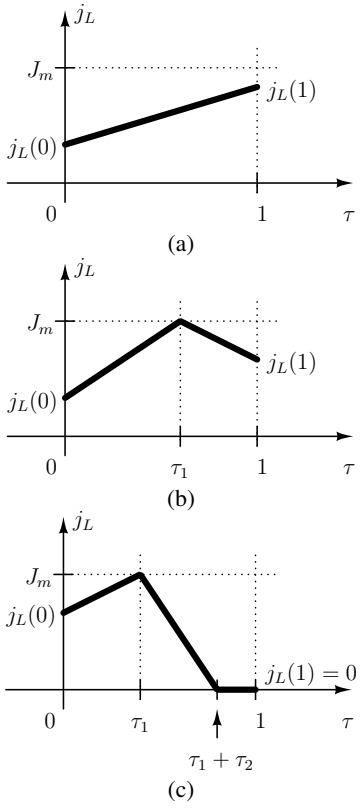


Fig. 4. Switching patterns.

diode is on. Contribution of this interval to the output charge is

$$q_1 = \frac{j_L(0) + J_m}{2}\,\tau_1. \tag{13}$$

In the second interval, if entered, the inductor current is determined by

$$j_L(\tau) = J_m - M\,(\tau - \tau_1)\,. \tag{14}$$

Assuming the switching cell remains in this state till the end of the switching period, the final value of the inductor current would be

$$j_Y = J_m - M\,(1 - \tau_1)\,. \tag{15}$$

In the case $j_Y \geq 0$, obtained value really is the final value

$$j_L(1) = j_Y \tag{16}$$

and contribution of this interval to the output charge is

$$q_2 = \frac{J_m + j_Y}{2}\,(1 - \tau_1) \tag{17}$$

resulting in the output current

$$j_{OUT} = q_1 + q_2. \tag{18}$$

This situation is depicted in Fig. 4(b). In the opposite case, when $j_Y < 0$, at

$$\tau_2 = \tau_1 + \frac{J_m}{M} \tag{19}$$

the diode turns off. Contribution of the time interval when the diode was on to the output charge is

$$q_3 = \frac{J_m}{2}\,(\tau_2 - \tau_1)\,. \tag{20}$$

The third interval, if entered, is characterized by $j_L = 0$, since both the switch and the diode are open. This results in the final value of the inductor current

$$j_L(1) = 0 \tag{21}$$

and the output current

$$j_{OUT} = q_1 + q_3. \tag{22}$$

This concludes the switching cell discrete time model. The model should be considered as a mapping of $j_L(0)$, M, and J_m to $j_L(1)$ and j_{OUT}. Although it is possible to provide these two expressions in a closed form, it is avoided here, since the approach that involves auxiliary variables is more convenient to be programmed, and the aim of this paper is semi numerical analysis of the converter, being performed by a computer program. The program used to analyze the buck converter switching cell operated using peak limiting current mode control is given in the Appendix.

IV. Conduction Modes

Under the term "discontinuous conduction mode", period-1 discontinuous conduction mode is frequently assumed [1]. This mode results in the inductor current waveform as shown in Fig. 3, and in the $(M,\, J_m)$ plane occurs for

$$J_m < M\,(1 - M) \tag{23}$$

as detailed in [8] using the same notation as in this paper. Besides, the open loop averaged model is unstable for $M > 1/2$, but in the whole region the DCM is characterized by period-1 operation, with stable limit cycle.

As an auxiliary result of [8], occurrence of multiple-period discontinuous conduction modes is described. Such modes are characterized by a waveform which repeats after an integer number of periods, and ends with a period of the type depicted in 4(c), ending by an interval in which the inductor current is equal to zero. Actually, this interval enforces periodical behavior, since the initial condition is exact and fixed.

A question that naturally arises is a range in $(M,\, J_m)$ plane in which multiple period discontinuous conduction modes occur. A boundary is determined with J_m value which is large enough to guarantee that the inductor will not get discharged over a switching period. Since the inductor is being discharged with the slope $-M$, and the change in the inductor current is $-J_m$ over normalized time $\Delta\tau = 1$, the upper boundary is determined by $J_m = M$. To summarize, conditions for the multiple period discontinuous conduction mode to occur are:

1) $J_m > M\,(1 - M)$, to exclude period-1 discontinuous conduction mode;
2) $M > \frac{1}{2}$, to provide unstable limit cycle of period-1 continuous conduction mode, as discussed in [8];
3) $J_m < M$, to allow the inductor current to discharge.

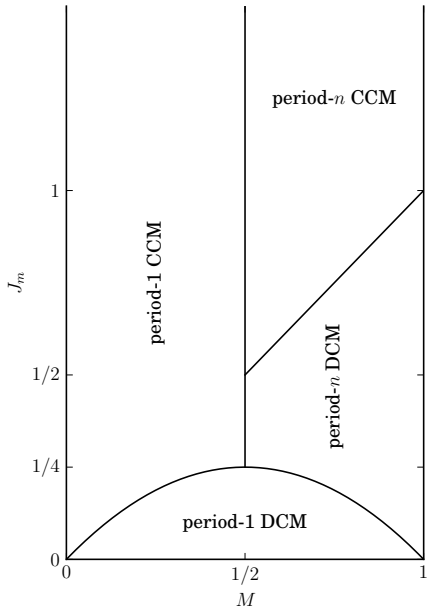


Fig. 5. Chart of the conduction modes.

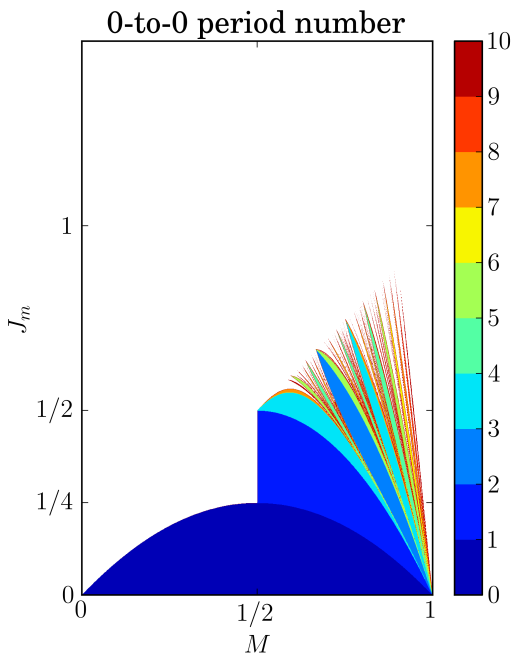


Fig. 6. Dependence of the period number on M and J_m.

Described region of operation is depicted in Fig. 5. Besides the described multiple period discontinuous conduction mode, common period-1 discontinuous conduction mode operating region is presented, as well as the continuous conduction modes:

1) period-1 continuous conduction mode, for $J_m > M\,(1-M)$ (to avoid the discontinuous conduction mode), and $M < \frac{1}{2}$, to provide stable period-1 limit cycle.
2) period-n continuous conduction mode, for $J_m > M\,(1-M)$, and $M > \frac{1}{2}$, characterized by an unstable limit cycle; it should be noticed that n in "period-n" might go to infinity ($n \to \infty$), which corresponds to chaotic operation.

V. Output Current and Stability

Essential function to be modeled is dependence of the switching cell output current, which is the average of the inductor current, $j_{OUT} = \overline{j_L}$, on the output voltage M, and the control parameter J_m. Under the assumption that the limit cycle is always period-1 and stable, there are two results, depending on the conduction mode:

1) in the discontinuous conduction mode the output current is given by

$$j_{OUT} = \frac{J_m^2}{2\,M\,(1-M)} \tag{24}$$

2) in the continuous conduction mode the output current is given by

$$j_{OUT} = J_m - \frac{1}{2}\,M\,(1-M)\,. \tag{25}$$

both of the equations reduce to $j_{OUT} = \frac{1}{2}\,J_m$ at the boundary between the modes. It should be underlined here that in both of the assumed modes the equations are symmetric to the $M = \frac{1}{2}$ line, since the same output current is obtained for M_x and $1-M_x$, assuming the same value of J_m. This result is revisited in this paper and it is shown that it is wrong, due to the limit cycle instability for $M > \frac{1}{2}$.

For $M > \frac{1}{2}$ the result for the output current for $M > \frac{1}{2}$ and $J_m > M\,(1-M)$, is different than given by (25), since the expected inductor current waveforms do not correspond to a stable limit cycle and do not appear in practice. Due to the unstable limit cycle in the continuous conduction mode, the switching cell might operate either in theoretically unlimited number of different period-n discontinuous conduction modes, or in period-n continuous conduction modes and chaotic, i.e. aperiodic, continuous conduction modes, in regions specified in Fig. 5. Dependence of j_{OUT} on M and J_m in such cases is convenient to be analyzed numerically, applying a program given in the Appendix.

The program to analyze the switching cell normalized model numerically, given in the Appendix, iterates the discrete time model of the switching cell for a given set of M and J_m values, starting from the initial value of the inductor current equal to zero. The number of iterations is limited to 500 in the given example, but any larger number would provide even more accurate results. The output current is computed during the simulation, and the output charge accumulated. If at a switching period the converter enters an interval when both the switch and the diode are off, as depicted in Fig. 4(c), the simulation stops, records the period number, and computes the output current according to the charge passed to the output and the number of periods. In the case converter operates in the continuous conduction mode, the simulation

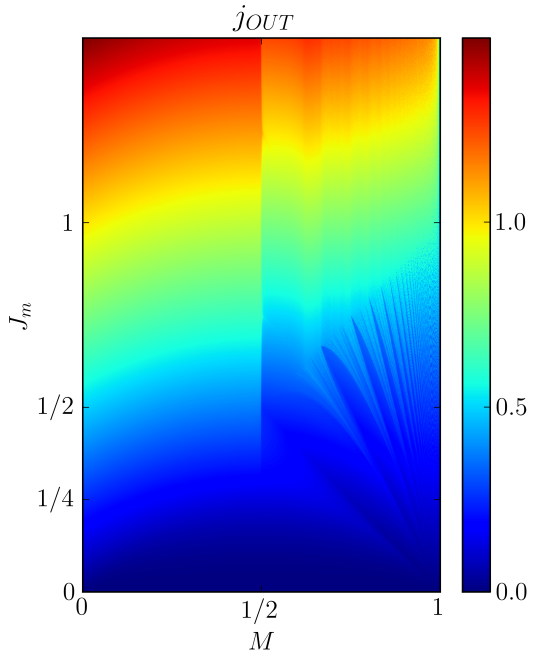


Fig. 7. Dependence of j_{OUT} on M and J_m.

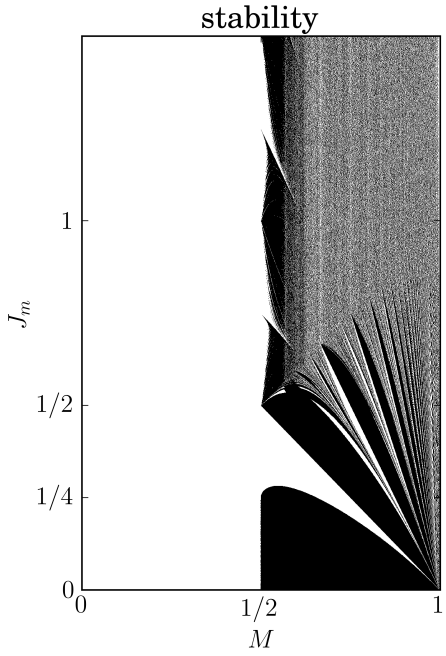


Fig. 8. Dependence of the switching cell open loop stability on M and J_m: white—stable; black—unstable.

stops after the maximal number of periods elapsed, and the output current is determined as the charge supplied to the output over the interval of time determined by the maximal number of periods. In the continuous conduction modes, the larger the maximal number of periods is, the more accurate value for the output current is obtained, since the startup transient is included in the simulation, not just the steady state operation. Program given in the Appendix is written in Python programming language, using PyLab environment, and presents a core part of the package used to generate results for this paper: it computes and stores the period number and average of the output current. The results are presented in Fig. 6 for the period number and in Fig. 7 for the output current. The diagram of the period number indicates boundary of the continuous conduction mode as predicted in Fig. 5 and by the analysis behind it. The dependence of j_{OUT} on M and J_m does not follow the symmetry predicted by (25) and (24), since instability of assumed fixed points is taken into account.

In [8], focus of the analysis was in period-1 discontinuous conduction mode, where the limit cycle is stable, but the converter is open loop unstable. Term "open loop" assumes constant J_m, not dependent on M. Open loop instability occurs when

$$\frac{d\, j_{OUT}}{dM} > 0 \tag{26}$$

as detailed in [8]. It is a phenomenon different than the limit cycle instability. For example, for $M > \frac{1}{2}$ and $J_m < M\,(1-M)$, the limit cycle is stable, in period-1 discontinuous conduction mode, while the converter is open loop unstable. To analyze stability for $M > \frac{1}{2}$ and $J_m > M\,(1-M)$ numerical results provided by the program given in the Appendix are analyzed to determine open loop stability of the converter. Such analysis requires derivative of j_{OUT} over M assuming J_m constant. The analysis is performed over the entire set of simulation results applying finite differences to estimate the derivative numerically. The result is presented in Fig. 8, where open loop stable regions are shown in white, while unstable regions are in black. The results match the theoretical predictions for regions where period-1 operation occurs, both for the continuous and for the discontinuous conduction mode, and provide new insight for limit-cycle unstable regions. Frequent changes from stable to unstable operation indicate that small-signal parameters, commonly used in regulator design, are of limited value in limit cycle unstable regions, due to the sensitivity on parameter variations.

VI. Conclusions

In this paper, operation of peak limiting current mode control operated buck converter is analyzed, aiming understanding of its operation in unstable limit cycle regions. The converter is analyzed assuming constant output voltage, effectively reducing the converter to the switching cell. Equations of the switching cell are normalized, to generalize the numerical results. A discrete time model of the switching cell is obtained in the form of two functions that map the initial value of the inductor current at the beginning of the switching cycle, the output voltage and the limiting current to the final value of the inductor current at the end of the switching cycle and the average of the inductor current during the switching cycle.

After the normalized model of the switching cell is obtained, it is analyzed assuming period-1 stable limit cycle operation, and boundary between the continuous and the discontinuous

conduction mode is derived, as well as the output current in both of the modes. The result indicates that the dependence of the output current is symmetric over the line determined by the output voltage being equal to one half of the input voltage. Knowing that the limit cycle is unstable for the output voltage greater than one half of the input voltage, boundary of the region where period-n discontinuous conduction modes might occur is determined analytically. Chart of modes is presented in the limiting current versus output voltage plane.

After the analytical approach, numerical methods are applied to analyze the converter. Obtained normalized model of the switching cell is simulated using the program given in the Appendix, and dependence of the output current and the period number on specified values of the output voltage and the limiting current are obtained. Open loop stability of the converter is analyzed, and complex behavior in regions characterized by unstable limit cycle is observer. In regions characterized by stable limit cycle, numerical and analytical results match perfectly. Expected symmetry of the output current assuming stable limit cycle is shown not to occur, and the output current obtained in regions where period-1 limit cycle is unstable is different from the value obtained assuming stability.

Two notions of stability, the limit cycle stability and the open loop stability, are clarified to be different. Open loop stability chart of the converter is presented, showing that all four combinations of the open loop and the limit cycle stability actually occur in the analyzed converter. In regions where the limit cycle is unstable the converter exposes complex behavior, and small signal analysis is of small practical value due to the sensitivity of parameter variations.

REFERENCES

[1] R. W. Erickson, D. Maksimović, *Fundamentals of Power Electronics*, 2nd Ed., Kluver, Norwell, MA, 2001.

[2] C. W. Deisch, "Simple switching control method changes power converter into a current source," *IEEE PESC'78*, 1978, pp. 300–306.

[3] F. D. Tan, R. D. Middlebrook, "A unified model for current-programmed converters," *IEEE Transactions on Power Electronics*, vol. 10, pp. 397–408, July 1995.

[4] J. Sun, D. M. Mitchell, M. F. Greuel, P. T. Krein, R. M. Bass, "Modeling of PWM converters in discontinuous conduction mode — A reexamination," *IEEE PESC'98*, 1998, pp. 615–622.

[5] D. Maksimović, S. Ćuk, "A unified analysis of PWM converters in discontinuous modes," *IEEE Transactions on Power Electronics*, vol. 6, pp. 476–490, May 1991.

[6] T. Sunito, "Analysis and modeling of peak-current-mode-controlled buck converter in DICM," *IEEE Transactions on Industrial Electronics*, vol. 48, February 2001, pp. 127–135.

[7] C.-C. Fang, "Unified discrete time modeling of buck converter in discontinuous mode," *IEEE Transactions on Power Electronics*, vol. 26, pp. 2335–2342, August 2011.

[8] M. Glišić, P. Pejović, "Stability Issues in Peak Limiting Current Mode Controlled Buck Converter," *17th International Symposium on Power Electronics Ee 2013*, Novi Sad, October-November 2013

VII. APPENDIX

```
from pylab import *


nitmax = 500
scale = 1000

nm = 1 * scale
deltam = 0.5 / nm
nj = 3 * scale / 2
deltaj = 1.5 / nj

M = linspace(deltam, 1 - deltam, nm)
Jm = linspace(deltaj, 1.5, nj)

Jout = empty((nm, nj))
Pern = empty((nm, nj))



mout = []
jout = []
pern = []



ijm = - 1
for jm in Jm:
    ijm += 1

    im = - 1
    for m in M:
        im += 1

        print jm, m

        # initialization
        tau = [0]
        jln = [0]
        dcmflag = 0
        jl0 = 0
        q = 0
        nper = 0

        count = 0
        while (dcmflag == 0 and count < nitmax):
            count += 1
            nper += 1
            tau1 = (jm - jl0) / (1 - m)
            if tau1 > 1:
                jl1 = jl0 + 1 - m
                tau.append(tau[-1] + 1)
                jln.append(jl1)
                q += (jl0 + jl1) / 2.0
                jl0 = jl1
            else:
                tau.append(tau[-1] + tau1)
                jln.append(jm)
                q += (jl0 + jm) / 2.0 * tau1
                tau2 = jm / m
                if tau1 + tau2 > 1:
                    tau3 = 1 - tau1
                    jl1 = jm - m * tau3
                    tau.append(tau[-1] + tau3)
                    jln.append(jl1)
                    q += (jm + jl1) / 2.0 * tau3
                    jl0 = jl1
                else:
                    tau.append(tau[-1] + tau2)
                    jln.append(0)
                    q += jm / 2.0 * tau2
                    dcmflag = 1
                    tau4 = 1 - tau1 - tau2
                    tau.append(tau[-1] + tau4)
                    jln.append(0)
                    jl0 = 0
        jout = q / nper

        Jout[im, ijm] = jout
        Pern[im, ijm] = nper



np.save('M.npy', M)
np.save('Jm.npy', Jm)
np.save('Jout.npy', Jout)
np.save('Pern.npy', Pern)
```

Modeling and Analysis of New Adaptive Dual Current Mode Control

Srđan Lale, Milomir Šoja, Slobodan Lubura
Faculty of Electrical Engineering
University of East Sarajevo
East Sarajevo, Bosnia and Herzegovina
srdjan.lale@etf.unssa.rs.ba, milomir.soja@etf.unssa.rs.ba,
slubura@etf.unssa.rs.ba

Milan Radmanović
Faculty of Electronic Engineering
University of Niš
Niš, Republic of Serbia
milan.radmanovic@elfak.ni.ac.rs

Abstract—**This paper proposes a new adaptive dual current mode control (ADCMC) approach which is modification of existing dual current mode control (DCMC). ADCMC introduces several significant advantages over DCMC, such as no peak-to-average error in the inductor-current signal, better transient response of inner current loop, improved line regulation and easier adjustment to different types of power electronics converters. Besides description of the working principles of ADCMC, this paper presents the development of small-signal model and transfer functions of ADCMC on the example of buck converter. Simulation results are presented which prove the derived analysis.**

Keywords- buck converter; current-mode control (CMC); line regulation; peak-to-average error; transient response.

I. INTRODUCTION

The control of power electronics converters can be divided into two main groups: voltage mode control (VMC) and current mode control (CMC). CMC is frequently used instead of VMC, because it has several important advantages over VMC. The one of them is overcurrent protection, for example. Further, it is well known that CMC improves converters' transient response by reducing the order of their transfer functions. Also, it improves line regulation by inherently built-in feed-forward property. CMC methods are first introduced in 1970s [1]. After that they have been increasingly used. Usually, CMC methods are divided into fixed-frequency and variable-frequency methods. In literature there are several basic fixed-frequency CMC methods and their modifications, including peak CMC (PCMC) [2]-[4], valley CMC (VCMC) [2], average CMC (ACMC) [5], [6] and charge control [7], [8].

PCMC and VCMC are the most commonly used among fixed-frequency methods. They have some excellent features, including constant switching frequency, simple implementation and good dynamic response. However, they have several drawbacks. The most important one is appearance of subharmonic oscillations when the duty cycle is above 0.5 (for PCMC) or below 0.5 (for VCMC) [9]. In order to eliminate these subharmonic oscillations, the slope compensation must be used. Also, they have large peak-to-average error in the inductor-current signal.

Variable-frequency CMC methods overcome these issues by operating in free-running mode. One of the most popular variable-frequency methods is hysteresis CMC [10], [11]. It has several advantages, including no slope compensation, no subharmonic oscillations and zero peak-to-average error. However, hysteresis CMC is not so suitable in practice due to its variable frequency. There are some modifications of hysteresis CMC, as in [12], [13], which rely on its fixed-frequency operation.

In [14] a fixed-frequency dual current mode control (DCMC) is proposed, which ensures stable operation of power converters for the entire range of duty cycle. Unlike PCMC and VCMC, DCMC needs dual boundaries (peak and valley) for the inductor current and two clock signals phase shifted for 180 degrees. In this way the converter naturally crosses from PCMC to VCMC and vice versa. DCMC is more complex to implement, but it offers important advantages, including no slope compensation, no subharmonic oscillations and fixed-frequency operation. However, DCMC has one important drawback. The gap voltage (width between two boundaries) must be chosen in advance properly to be larger than the maximum peak-to-peak ripple of the inductor current. This can adversely affect the waveforms of converter's inductor current and output voltage, especially in the cases of power factor correction circuits (PFC) and inverters, where peak-to-peak current ripple always changes during their work. Also, significant peak-to-average error exists.

This paper proposes a new adaptive dual current mode control (ADCMC), which improves the qualities of DCMC by introducing an adaptive gap voltage, which is equal to the measured instantaneous value of peak-to-peak current ripple. The result is close to hysteresis operation, but with fixed-frequency, which is ensured with two clock signals. ADCMC offers several advantages over DCMC which will be demonstrated in this paper.

This paper is organized as follows. ADCMC's principles of operation are presented in Section II. A small-signal model on the example of buck converter is developed in Section III. Section IV presents the simulation results. Finally, the conclusion is given in Section V.

II. PRINCIPLES OF OPERATION OF ADCMC

The basic principles of operation of DCMC proposed in [14] are presented in Fig. 1 for buck converter, along with characteristic operating modes. DCMC has two main parts: outer voltage loop (negative feedback with compensator $G_c(s)$ for regulation of output voltage v_o) and inner current loop (two comparators, clocks clk_A and clk_B and driving logic). The output of voltage compensator v_c is actually the reference inductor current. Upper and lower boundaries for the inductor current i_L are formed by summing and subtracting the control signal v_c with the voltage V_a. For correct and stable operation of DCMC the voltage band $2V_a$ must be larger than the maximum peak-to-peak ripple of the inductor current. Therefore, the voltage V_a must satisfy the following condition:

$$2V_a > K_{iL}\Delta i_{Lpp\max} = K_{iL}\frac{v_g}{4f_sL}, \tag{1}$$

where: K_{iL} is measuring gain and Δi_{Lppmax} is maximum peak-to-peak ripple of the inductor current, v_g is input voltage, $f_s=1/T_s$ is switching frequency and L is the inductor value.

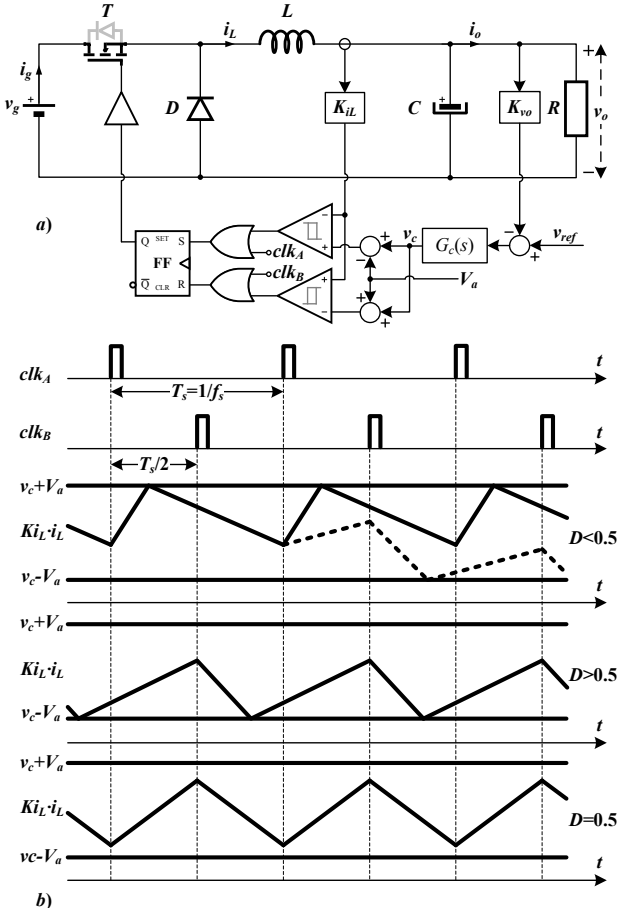


Figure 1. a) DCMC of buck converter, b) Characteristic operating modes.

It is obvious from Fig. 1.b (except for duty cycle D=0.5) that there is significant peak-to-average error in the inductor current (the difference between the reference current v_c and the average value of the inductor current), similarly as in PCMC and VCMC. Therefore, DCMC isn't able to directly control the average value of the inductor current. This can be an issue especially in PFC or inverter topologies. In these topologies with high voltage band $2V_a$ the transient response of the inductor current in boundary case (D=0.5) can be too slow, which results in distortion of waveforms of the inductor current and the output voltage. Also, every time when operating conditions change, such as input voltage, load, switching frequency, etc., a new value of voltage V_a must be determined.

In order to resolve these issues, this paper proposes a new ADCMC method, which is based on computation of voltage band $2V_a$ for the instantaneous operating conditions of converter. The voltage band $2V_a$ can be calculated using the instantaneous value of peak-to-peak current ripple Δi_{Lpp} in the following way:

$$2V_a = K_{va}K_{iL}\Delta i_{Lpp}, \tag{2}$$

where $K_{va}\geq 1$ is scaling factor. If K_{va}=1, ADCMC becomes similar to the hysteresis CMC, but with constant switching frequency. The great benefit from this feature is that ADCMC directly controls the average value of the inductor current, so no peak-to-average error exists. Also, the current transient response will be faster when some of the converter's parameters, especially input voltage, take sudden changes.

Equation (2) applies for general case regardless to the type of power converter. However, it is very difficult to directly measure the current ripple Δi_{Lpp}. It is commonly calculated by measuring the input and output voltage of converter. For the buck converter, the voltage band $2V_a$ is equal to:

$$2V_a = K_{va}K_{iL}\frac{v_o}{Lf_s}\left(1-\frac{v_o}{v_g}\right). \tag{3}$$

The scheme from Fig. 1.a remains the same for ADCMC.

III. SMALL-SIGNAL MODEL ANALYSIS

A small-signal AC model of the buck converter with ADCMC can be derived using standard procedure described in [15]. The small-signal averaged equations for buck converter from Fig. 1.a operating in the continuous conduction mode (CCM), under duty cycle control, can be expressed as:

$$\begin{aligned} L\frac{d\hat{i}_L(t)}{dt} &= D\hat{v}_g(t)-\hat{v}_o(t)+\hat{d}(t)V_g \\ C\frac{d\hat{v}_o(t)}{dt} &= \hat{i}_L(t)-\frac{\hat{v}_o(t)}{R} \\ \hat{i}_g(t) &= D\hat{i}_L(t)+\hat{d}(t)I_L, \end{aligned} \tag{4}$$

where $\hat{v}_g(t)$, $\hat{v}_o(t)$, $\hat{i}_g(t)$, $\hat{i}_L(t)$ and $\hat{d}(t)$ are small AC variations superimposed to the quiescent values V_g, V_o, I_g, I_L and D of input and output voltage, input and inductor current, and duty cycle, respectively. The quiescent values are equal to:

$$V = DV_g,\ I_L = \frac{V_o}{R},\ I_g = DI_L. \tag{5}$$

For modeling of ADCMC, a simple first-order approximation is employed, under assumption that the

measured average inductor current $K_{iL}\langle i_L(t)\rangle_{T_S}$ is equal to the reference current $v_c(t)$. This approximation is valid if the factor K_{va} is equal or approximate to one, while for DCMC this isn't the case. Using this approximation, following applies:

$$K_{iL}\hat{i}_L(t) = \hat{v}_c(t), \tag{6}$$

where $\hat{v}_c(t)$ is a small AC variation of current reference $v_c(t)$.

Substitution of (6) in (4) leads to the small-signal model of buck converter with ADCMC:

$$\begin{aligned} \frac{L}{K_{iL}}\frac{d\hat{v}_c(t)}{dt} &= D\hat{v}_g(t) - \hat{v}_o(t) + \hat{d}(t)V_g \\ C\frac{d\hat{v}_o(t)}{dt} &= \frac{\hat{v}_c(t)}{K_{iL}} - \frac{\hat{v}_o(t)}{R} \\ \hat{i}_g(t) &= \frac{D}{K_{iL}}\hat{v}_c(t) + \hat{d}(t)I_L. \end{aligned} \tag{7}$$

The Laplace transforms of (7) are:

$$\begin{aligned} \frac{L}{K_{iL}}s\hat{v}_c(s) &= D\hat{v}_g(s) - \hat{v}_o(s) + \hat{d}(s)V_g \\ Cs\hat{v}_o(s) &= \frac{\hat{v}_c(s)}{K_{iL}} - \frac{\hat{v}_o(s)}{R} \\ \hat{i}_g(s) &= \frac{D}{K_{iL}}\hat{v}_c(s) + \hat{d}(s)I_L. \end{aligned} \tag{8}$$

The control-to-output transfer function $G_{vc}(s)$ can be obtained from (8) as:

$$G_{vc}(s) = \left.\frac{\hat{v}_o(s)}{\hat{v}_c(s)}\right|_{\hat{v}_g(s)=0} = \frac{R}{K_{iL}\left(1+sRC\right)}. \tag{9}$$

The simple transfer function (9) is used for design of the output voltage compensator $G_c(s)$. It is obvious from (8) that the line-to-output transfer function $G_{vg}(s) = \hat{v}_o(s)/\hat{v}_g(s)\big|_{\hat{v}_c(s)=0}$ is zero, which offers great line regulation. The identical transfer functions can be derived for DCMC, however due to the fact that the first-order approximation isn't valid for DCMC, the nonzero line-to-output transfer function exists, which is proved in [15] using a more accurate model.

IV. Simulation Results

The performances of ADCMC and its advantages over DCMC in the same operating conditions were tested with simulations for the buck converter with following parameters: V_g=28 V, L=120 µH, C=1000 µF, R=4 Ω and f_s=20 kHz. The all scaling and measuring gains K_{iL}, K_{va} and K_{vo} are set to one. According to (1) for DCMC the value of V_a was set to 2 V. The output voltage compensator $G_c(s)$ is a simple proportional-integral (PI) regulator. There will not be discussion about its design, because this paper is focused only on the inner current loop. Dynamics of the output voltage loop is too slow and doesn't affect the faster inner current loop.

A. Peak-to-average Error in Stationary State

In order to show the difference in peak-to-average error for DCMC and ADCMC, stationary state was analyzed for two values of the output voltage: 10 V and 20 V. The waveforms of inductor current, control signal, upper and lower current limit in stationary state are shown in Fig. 2 for DCMC and Fig. 3 for ADCMC.

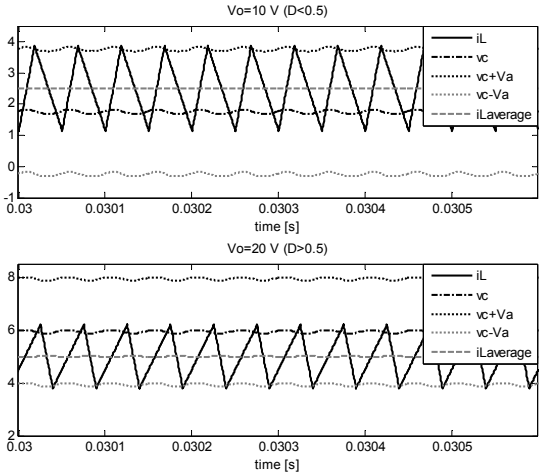


Figure 2. Simulation waveforms in stationary state for DCMC.

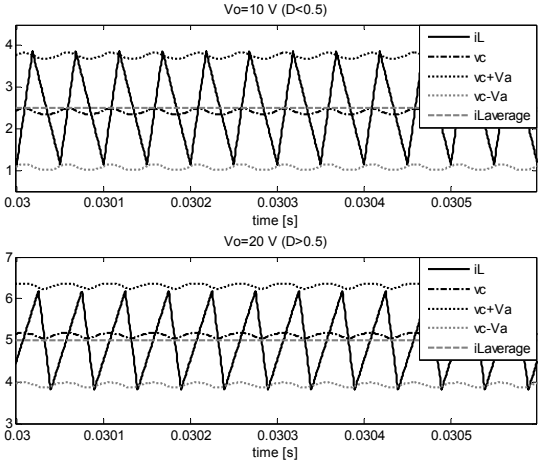


Figure 3. Simulation waveforms in stationary state for ADCMC.

It is obvious from Fig. 3 for ADCMC that the average value of the inductor current excellently follows the current reference (control signal v_c). However, if noted carefully, there is a small error between these two signals, which can be attributed to the delays in numerical calculation (simulation's solver) and fact that the voltage V_a from (3) applies for simplified model of buck converter in ideal conditions.

B. Line Regulation

Step changes of input voltage from 28 V to 16 V and vice versa are introduced in simulation in order to check the quality of line regulation of ADCMC. The output voltage was regulated to the value of 10 V. The results are shown in Fig. 4, Fig. 5 and Fig. 6.

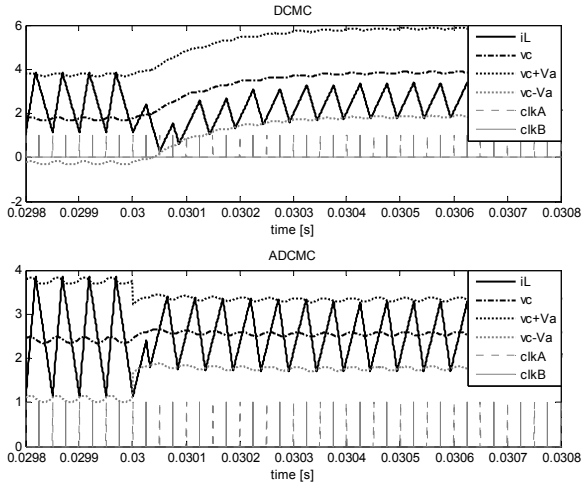


Figure 4. Simulation waveforms for a step change in the input voltage from 28 V to 16 V.

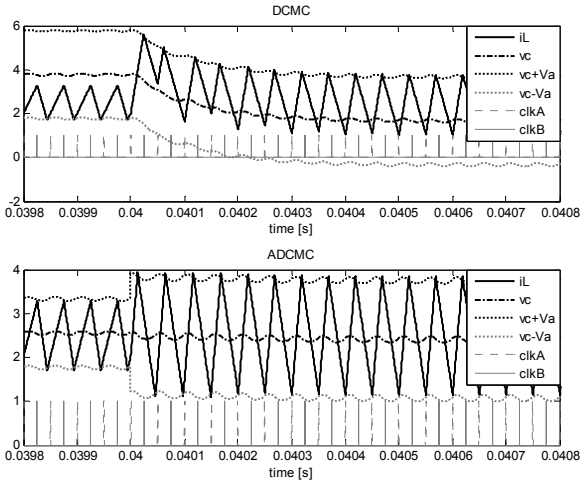


Figure 5. Simulation waveforms for a step change in the input voltage from 16 V to 28 V.

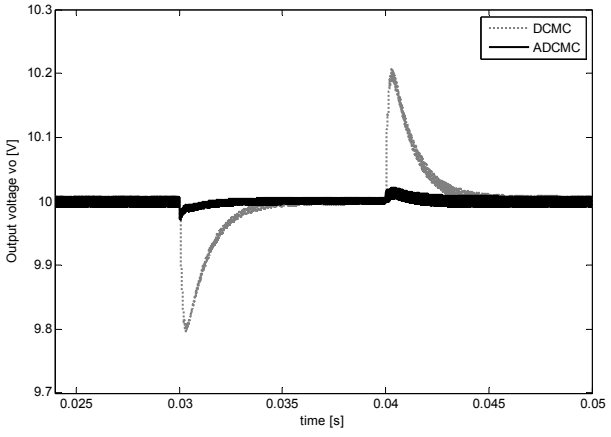


Figure 6. Output voltage for a step change in the input voltage from 28 V to 16 V and vice versa.

It is obvious from Fig. 4 and Fig. 5 that ADCMC has much better transient response of the inductor current, which results in a very improved line regulation, as shown in Fig. 6.

V. CONCLUSION

In this paper, a new adaptive dual current mode control method was proposed. A detailed analysis of ADCMC's principles of operation and small-signal model was presented. The performed simulations proved the analysis and showed some significant advantages of ADCMC over DCMC.

REFERENCES

[1] C. W. Deisch, "Simple switching control method changes power converter into a current source," in Proc. IEEE Power Electron. Spec. Conf., 1978, pp. 135–147.

[2] S. S. Hsu, A. Brown, L. Rensink, and R. D. Middlebrook, "Modeling and analysis of switching dc-to-dc converters in constant-frequency current programmed mode," in Proc. IEEE Power Electron. Spec. Conf., 1979, pp. 284–301.

[3] T. V. Papathomas, "On the stability of peak current-controlled converters: Analysis, simulation, and experiments," IEEE Trans. Ind. Electron., vol. IE-33, no. 2, pp. 176–184, May 1986.

[4] M. Hallworth and S.A. Shirsavar, "Microcontroller-Based Peak Current Mode Control Using Digital Slope Compensation," IEEE Trans. Power Electron., vol. 27, no. 7, pp. 3340-3351, July 2012.

[5] L. Dixon, "Average current mode control of switching power supplies," Unitrode Corp., San Jose, CA, Unitrode Appl. Note U-140, 1999, pp. 356–369.

[6] Y. Qiu, H. Liu, and X. Chen, "Digital average current-mode control of PWM DC–DC converters without current sensors," IEEE Trans. Ind. Electron., vol. 57, no. 5, pp. 1670–1677, May 2010.

[7] W. Tang, F. C. Lee, R. B. Ridley, and I. Cohen, "Charge control: Modeling, analysis, and design," IEEE Trans. Power Electron., vol. 8, no. 4, pp. 396–403, Oct. 1993.

[8] H. S. Maeques, B. V. Borges, and A. V. Anunciada, "Transient behavior and frequency response of constant frequency one cycle charge control modulator (O3C)," IEEE Trans. Power Electron., vol. 23, no. 5, pp. 2507–2515, Sep. 2008.

[9] A.V. Anunciada and M.M. Silva, "On the stability and subharmonic susceptibility of current-mode controlled converters," Power Electronics Specialists Conference, PESC '92 Record., 23rd Annual IEEE, vol. 1, no., pp. 345-353, 29 Jun.-3 Jul. 1992.

[10] R. Redl, "Small-signal high-frequency analysis of the free-running current-mode-controlled converter," in Proc. IEEE Power Electron. Spec. Conf., Jun. 1991, pp. 897–906.

[11] A. Borrell, M. Castilla, J. Miret, J. Matas, and L. G. de Vicuna, "Simple low-cost hysteretic controller for multiphase synchronous buck converters," IEEE Trans. Ind. Electron., vol. 58, no. 6, pp. 2355–2365, Jun. 2011.

[12] S. Buso, S. Fasolo, L. Malesani, and P. Mattavelli, "A dead-beat adaptive hysteresis current control," IEEE Trans. Ind. Applications, vol. 36, no. 4, pp.1174-1180, Jul/Aug. 2000.

[13] S. C. Huerta, P. Alou, J. A. Oliver, O. Garcia, and J. A. Cobos, "Nonlinear control for DC–DC converters based on hysteresis of the C_{OUT} current with a frequency loop to operate at constant frequency," IEEE Trans. Ind. Electron., vol. 58, no. 3, pp. 1036–1043, Mar. 2011.

[14] A. V. Anunciada and M. M. Silva, "A new current mode control process and applications," IEEE Trans. Power Electron., vol. 6, no. 4, pp. 601-610, Oct. 1991.

[15] R. W. Erickson and D. Maksimović, Fundamentals of Power Electronics, 2nd ed. Boston, MA: Kluwer, 2001, pp. 439–487.

Session T-03
CIRCUITS AND SYSTEMS

Python Application for Analyzing Analog Filters' Transfer Function

Miljana Milić and Vančo Litovski
University of Niš, Faculty of Electronic Engineering,
Niš, Serbia
miljana.milic@elfak.ni.ac.rs, vanco.litovski@elfak.ni.ac.rs

***Abstract*— Analog filter represents an inevitable block in modern telecommunications and electronics. They suppress the undesired spectrum components, regardless of their nature. An interactive Python application, developed for analyzing the analog filters' transfer function is described in this paper. The application calculates and draws most of the filter's characteristics such as: step and impulse response of the filter in the time domain, and, module, phase and group delay in the frequency domain. The application accepts transfer function of the filter described both using poles and zeros of the function, or polynomials of the numerator and denominator as inputs. A Python programming language is used as a development platform, since it is free, and offers large available online support.**

Keywords- analog filters, circuit analysis, Python

I. INTRODUCTION

Almost every electronic device today has a filter. Filtering is an essential process in every data processing block. Its task is to suppress unwanted spectral components regardless of their origin.

The transfer function of the filter is often defined in domain of complex frequencies (s-domain). When designing and analyzing digital filters, calculations are performed in the z-domain.

The intention of this work is to develop an interactive software package that can analyze transfer functions of analog filters [1]. It involves the calculation and graphical representation of the module, phase and group delay of the filter in the frequency domain, as well as the response of the filter in the time domain to pulse and step stimuli. This can later be used for the extraction of the most important parameters of the function that's being analyzed. The application for these analyses accepts the analog filter transfer functions that can be described with its poles and zeros, or with the coefficients of the numerator's and denominator's polynomials. We are aware of the fact that there are many similar applications available on the internet. Nevertheless, in our experience, many of them do not show correct results of transfer function analysis.

Python is chosen as a programming platform for developing the application. It is a widely used programming language that supports object-oriented, imperative and functional programming styles. It is free and available for majority of operating systems and can be packed into stand-alone executable programs, allowing the distribution of Python-based software for use on those environments without requiring the installation of a Python interpreter. It has a large and wide-ranging standard library. As of October 2014, the Python Package Index [2], which is the official repository of the third-party software for Python, contains more than 49730 packages offering a variety of functionality, including:

- graphical user interfaces, web frameworks, multimedia, databases, networking and communications
- test frameworks, automation and web scraping, documentation tools, system administration
- scientific computing, data processing, image processing.

In our case there were two major reasons for choosing Python programming language. First, it is C/C++ compatible, and the second, it is also compatible with a platform for software radio development GNU [3].

The theoretical basics of analog filters analysis will be given first, as well as formulas for calculating their properties. After that, some most important features of the Python programming environment will be described, including all necessary packages. This is followed by description of the application development, results of filter analysis example and the conclusion.

II. FILTER'S TRANSFER FUNCTION AND ITS PROPERTIES

Analog filters that we intend to analyze here can be passive and active circuits. They meet the requirements of causality and stability. Their transfer functions are expressed as a rational function of two polynomials where the order of the numerator, *m*, is less than or equal to the order of the denominator *n*.

In this section, we will give basic definitions related to the transfer function of signal filters and formulas for calculating the characteristics that are extracted from them [4], [5], [6]. This analysis is performed in the frequency and in the time domain. Since the response of the filter in the time domain can be calculated only using the Residue Calculus, it is necessary to know poles and zeros of the filter transfer function. Similar requirements stand for the group delay and phase characteristics calculation. Namely, calculation of the filter's phase characteristic using only the coefficients of the

polynomials can cause erroneous results. Therefore, our application applies formulas that use only zeros and poles of the transfer function. It is assumed that polynomials of the transfer function are solved first, by the designer, or alternatively by this Python application.

The factorized form of the transfer function in the s-plane can be represented with:

$$T(s) = T_0 \cdot \frac{\prod_{k=1}^{m}(s - z_k)}{\prod_{i=1}^{n}(s - pi)}. \quad (1)$$

Here, s stands for the complex angular frequency: $s = \sigma + j\omega$. The imaginary part $\omega = \mathrm{Im}\{s\}$, is referred to as real angular frequency: $\omega = 2\pi f$, while f stands for the actual frequency, expressed in (Hz). Zeros of the transfer function are $z_k = \alpha_k + j\beta_k$, $k = 1, 2, ..., m$, while the poles are $p_i = \gamma_i + j\delta_i$, $i = 1, 2, ..., n$. where n represented the order of the filter and should satisfy that $n \geq m$. Using the formula:

$$T(0) = T_0 \cdot \frac{\prod_{k=1}^{m}(-z_k)}{\prod_{i=1}^{n}(-p_i)}, \quad (2)$$

one can calculate gain of the filter at zero angular frequency, while using:

$$T(j1) = T_0 \cdot \frac{\prod_{k=1}^{m}(j1 - z_k)}{\prod_{i=1}^{n}(j1 - p_i)}, \quad (3)$$

one can calculate gain of the filter at unit angular frequency. Since the angular frequency is usually normalized so that $\omega_{ctitical} = 1$, (for low pass filters this is usual bandwidth upper cutoff frequency, and for bandpass filters, this is usual bandwidth center frequency), value expressed by (3) is often of great importance.

On the axis of real frequencies, i.e. for s = jω, the amplitude characteristic is calculated as:

$$\left|T(s)\right|_{s=j\omega} = \left[T(s)T(-s)\big|_{s=j\omega}\right]^{1/2}, \quad (4)$$

or,

$$\left|T(s)\right|_{s=j\omega} = \left|T_0\right| \frac{\prod_{k=1}^{m}\left[\alpha_k^2 + (\omega - \beta_k)^2\right]^{1/2}}{\prod_{i=1}^{n}\left[\gamma_i^2 + (\omega - \delta_i)^2\right]^{1/2}}. \quad (5)$$

If presented in the *semilog* scale, an amplitude characteristic is calculated with (6):

$$A(\omega) = 20 \cdot \log\left\{\left|T(s)\right|_{s=j\omega}\right\} \text{ [dB]}, \quad (6)$$

while the attenuation can be calculated using:

$$a(\omega) = 20 \cdot \log\left\{\frac{1}{\left|T(s)\right|_{s=j\omega}}\right\} \text{ [dB]}. \quad (7)$$

It is extremely important that the phase characteristics are calculated using:

$$\varphi(\omega) = \arg\left\{T(s)_{|s=j\omega}\right\} = \frac{1}{2j}\ln\left[\frac{T(s)}{T(-s)}\right]_{s=j\omega} \quad (8)$$

or:

$$\varphi(\omega) = \sum_{k=1}^{m}\mathrm{arctg}\left[\frac{\beta_k - \omega}{\alpha_k}\right] - \sum_{i=1}^{n}\mathrm{arctg}\left[\frac{\delta_i - \omega}{\gamma_i}\right]. \quad (9)$$

During the *arctg* function calculation, for each value of the angular frequency, one should first determine the quadrant of the angle by observing the signs of the denominator and the numerator. After that it is possible to calculate the desired angle. For example: if signs of the numerator and the denominator are both negative, the observed angle lays in the third quadrant of the complex plane, so that *arctg* function can be calculated as:

$$\pi + \mathrm{arctg}\left(\left|numer./denom.\right|\right). \quad (10)$$

The group delay is calculated with the formula:

$$\tau(\omega) = -\frac{d\varphi(\omega)}{d\omega} = \sum_{k=1}^{m}\frac{\alpha_k}{\alpha_k^2 + (\omega - \beta_k)^2} - \sum_{i=1}^{n}\frac{\gamma_i}{\gamma_i^2 + (\omega - \delta_i)^2}. \quad (11)$$

The general form of the response in the time domain at the impulse function is shown in the following expression:

$$h(t) = \sum_{i=1}^{n}\lim_{s \to p_i}\frac{1}{(q-1)!}\left\{\frac{d^{q-1}}{ds^{q-1}}\left[(s - p_i)^q T(s) e^{st}\right]\right\} \quad (12a)$$

where q represents the order of the denominator's polynomial. If the transfer function has only simple poles, than we have:

$$h(t) = \sum_{i=1}^{n}\left(\lim_{s \to p_i}\left\{\left[(s - p_i)\cdot T(s)\cdot e^{st}\right]\right\}\right). \quad (12b)$$

Substituting (1) into (12b), we obtain:

$$h(t) = \sum_{i=1}^{n}\lim_{s \to p_i}\left\{T_0 \frac{\prod_{k=1}^{m}(s - z_k)}{\prod_{\substack{r=1 \\ r \neq i}}^{n}(s - p_r)} e^{st}\right\} = \quad (13)$$

$$= \sum_{i=1}^{n}\left\{T_0 \frac{\prod_{k=1}^{m}(p_i - z_k)}{\prod_{\substack{r=1 \\ r \neq i}}^{n}(p_i - p_r)}\left[\cos(\delta_i t) + j\cdot\sin(\delta_i t)\right] e^{\gamma_i t}\right\}$$

where $p_i=\gamma_i+j\delta_i$ and $p_r=\gamma_r+j\delta_r$ represent vectors of poles (complex), while $z_k=\alpha_k+j\beta_k$ denominates vector of complex zeros. The order of the filter is n (equal to the number of poles), m is the order of the numerator's polynomial (the number of zeros); σ_i is the real part of the pole, ω_i is the imaginary part of the pole, α_k is a real part of the zero, β_k is the imaginary part of the zero.

The response to the step-function is obtained when the set of poles is extended with one pole at zero. The new function whose inverse Laplace transform is to be determined would be:

$$T_0 \frac{1}{s} \frac{\prod_{k=1}^{m}(s-z_k)}{\prod_{j=1}^{n}(s-p_j)}. \qquad (14)$$

In (13) the sum is increased by one, and set of poles is extended with: $p_{n+1}=0+j0$.

III. The Development of the Python Application

Python is a programming language that was developed in compliance with the OSI (Open Systems Interconnection) standards. It supports various programming styles. License for this programming language is opened (Open source) and it is free to use and distribute, even for commercial purposes. There are hundreds of software modules and packages for Python, available as standard libraries, or as those created by specialized groups and communities [7], [8].

For the development of our Python application, the Eclipse programming environment [9] was used. The main reasons for this choice are its availability and openness, and easy development and debugging of programs. For the development of GUI applications, we used the standard Python's Tkinter GUI package [10]. Other Python's GUIs are also available, like WX [11], or QT [12]. We have chosen Tkinter, because it is a Python's default GUI toolkit. It is also stabile, mature, well documented, and has strong binding mechanism.

For the development of application for the analysis of analog filters, it was necessary to use tree Python's package. The first allowed the use of certain functions and operations of the numerical mathematics - Numpy (Numerical Python) [13], second enables symbolical calculations - Sympy (Symbolical Python) [14], while the third allows displaying of calculation results using versatile graphs – matplotlib [15].

A brief description of the application and instructions for its use will be given next.

IV. Analog Filter Analysis with the Python Application

When the application **Transfer function properties** is launched, first the user is asked to enter two integers: the number of zeros and poles of the filter's transfer function. It is also possible to enter the transfer function of the filter in the alternative way, representing it with coefficients of the numerator's and denominators polynomials. This is shown in Fig. 1. After pressing the "Fetch" button, the application window is expanded, in order to enter simple zeros and poles of the function, as well as the initial gain of the filter. This is shown in Fig. 2.

Figure 1. Openning the application; input parameters

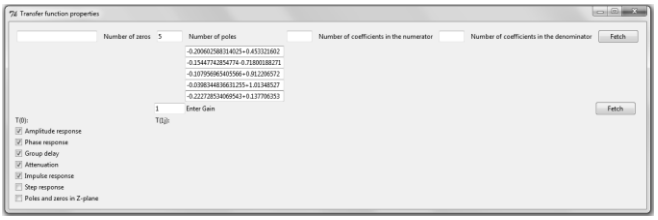


Figure 2. Extended application window for entering complex zeros and poles, and selecting commands

When represented with its polynomials' coefficients, the application automatically calculates their roots, considering them afterwards as sets of poles and zeros.

The coefficients for this filter example are taken from [16]. This is an example of one H-type critical-monotonic all-poles low-pass filter, of 10^{th} order. Table 1. shows those coefficients.

After pressing the "Fetch" button from the extended application window, the program first calculates typical gains of the filter. This is shown in Fig. 3.

Figure 3. Calculation of the filter unity gains

Figs. 4, 5, 6, and 7 give the results of the filter analysis in the frequency domain. Those characteristics are: the amplitude characteristic (in dB), phase characteristic (in rad), group delay (in sec), and the attenuation (in dB), respectively. At Fig. 7, a detailed view at the most important part of the attenuation characteristic is zoomed. The application offers that option as well as saving the graphs into corresponding .png files.

The responses of the filter in the time domain are shown in Fig. 8.

Finally, Fig. 9 shows the position of poles and zeros in the s-plane. Poles are marked with crosses, while zeros are marked with small circles. It can be noticed that this graph is symmetrical along the real axes.

V. Conclusion

This paper presents the development of application for the analysis of analog filters described with zeros, poles and gain, or with the coefficients of the numerator's and the denominator's polynomials of the transfer function. The developed applications successfully perform these analyzes. We are aware that many commercial tools support the same calculation and analysis. However, the important advantage of our solution is the availability of Python programming environment, and therefore the openness and availability of developed application. Further research and application development will be oriented towards the implementation of calculation of additional features of filters, as well as the additional controls in the application.

TABLE I. THE COEFFICIENTS OF THE DENOMINATOR POLYNOMIAL OF THE H FILTER

s^{10}	s^9	s^8	s^7	s^6	s^5	s^4	s^3	s^2	s^1	s^0
1.0	1.4512	3.5253	3.6016	4.1933	2.9624	1.9859	0.8937456	0.3222619	0.0712828	0.0078896

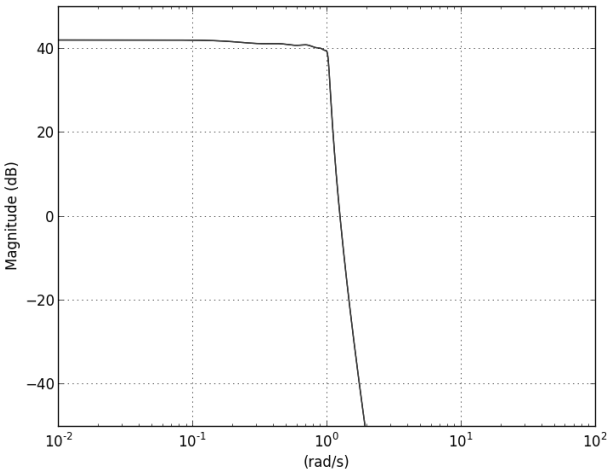


Figure 4. Amplitude characteristic of the filter

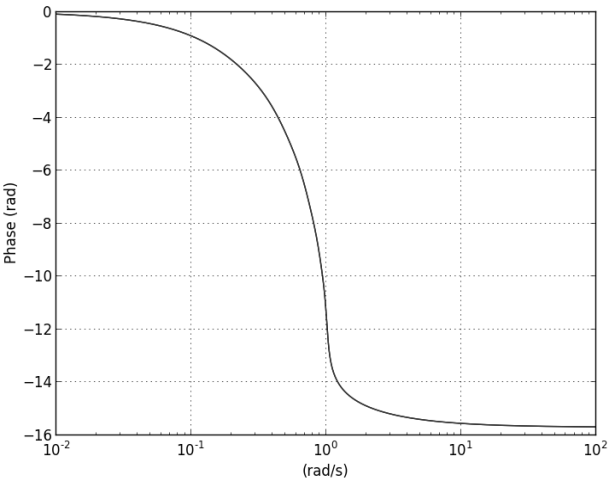


Figure 5. Phase characteristic of the filter

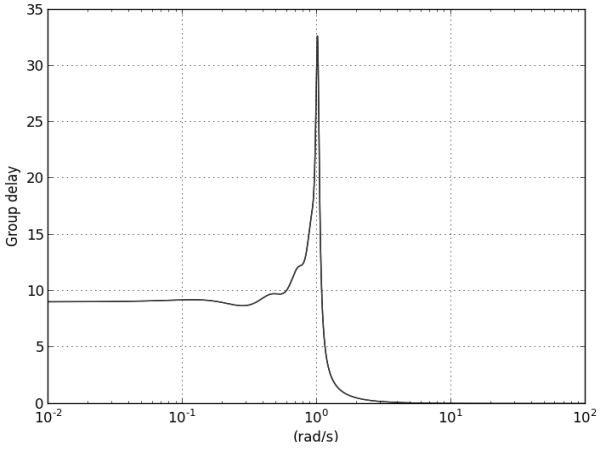


Figure 6. Group delay

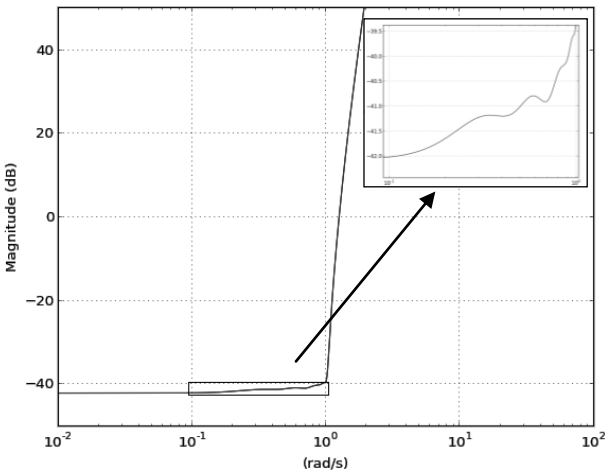


Figure 7. The attenuation

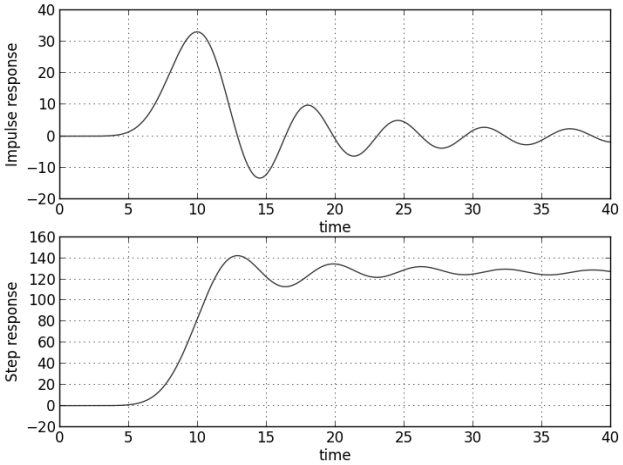


Figure 8. Impulse and Step response of the filter

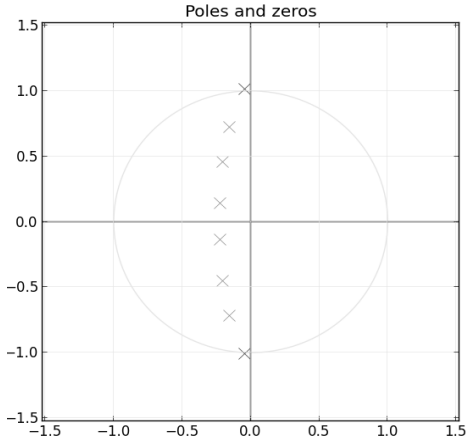


Figure 9. S-plane with all poles and zeros

ACKNOWLEDGMENT

This research was partly funded by The Ministry of Education, Science and and Technological Development of Republic of Serbia under the contract No. TR32004.

REFERENCES

[1] M. Milić and V. Litovski, "Analysis of analog filters' transfer function using Python programming language", Proc. of the LVIII ETRAN conf. Vrnjačka Banja, Serbia, 2014, EL.2.4.

[2] https://pypi.python.org/pypi

[3] http://gnuradio.org/

[4] M. Lutovac, D, Tošić, and B. Evans, Filter Design for Signal Processing Using MATLAB and Mathematica, Prentice Hall, NY, 2001.

[5] DeVerl Humpherys, The Analysis, design, and synthesis of electrical filters, Englewood Cliffs, Prentice-Hall, NY, 1970.

[6] V. Litovski and M. Zwolinski, VLSI Circuit Simulation and Optimization, Chapman and Hall, London, 1997.

[7] www.python.org

[8] www.pydev.org

[9] www.eclipse.org

[10] https://wikipython.org/moin/Tkinter

[11] http://wiki.wxpython.org/

[12] https://qt-project.org

[13] www.numpy.org

[14] www.sympy.org

[15] www.matplotlib.org

[16] D. Topisirović, V. Litovski, and M. A. Stošović, "Unified theory and state-variable implementation of critical-monotonic all-pole filters," Int. J. Circ. Theor. Appl. Wiley, 22 Oct, 2013.

Faults Simulations in XOR/XNOR Cell Resistant to Side Channel Attacks

Milena Stanojlović Mirković
Innovation Centre of Advanced Technologies
ICAT
Niš, Serbia
milena.stanojlovic@icnt.rs

Vančo Litovski, Predrag Petković and Dragiša Milovanović
Faculty of Electronic Engineering, University of Niš
Niš, Serbia
{vanco,predrag,gilem}@elfak.ni.ac.rs

Abstract—**This paper describes simulation results of testing the No Short-circuit current Dynamic Differential Logic (NSDDL) XOR/XNOR cell which consist of two NSDDL AND cells and one NSDDL OR cell. The goal is to consider the impact of individual defects in such complex circuit. Fault dictionary will be created based on repetitive simulations preformed for defects inserted one by one. For a short circuit defects detection logical function and supply current will be exploited. All cells are designed in CMOS TSMC035 technology using Mentor Graphics design tools.**

Keywords- cell; defects; short circuits; testing.

I. INTRODUCTION

The dynamics consumption tracking of an electronic crypto-system, can provide more information about the system behavior and to make cracking the key easier. The most effective methods for attack on the crypto-system are Simple Power Analysis (SPA), Differential Power Analysis (DPA) and Electromagnetic Analysis (EMA) [1]. All of them relay on tracking the crypto-system activity by monitoring the changes in power consumption. Practically this means that measurements of biasing current will provide the additional information about circuit behavior. Therefore one talks about this current as a source of leaked information or as a Side Channel. Complex cryptographic algorithms are designed to discourage the attacker, or to impede the breaking the key by searching for all possible combinations in real time. Additional information about the behaviour of an electronic crypto-system can significantly reduce the number of combinations needed to explore a cipher. Collecting such information is known as the Side Channel Attack - SCA[2, 3].

We chose the NSDDL (No Short-circuit current Dynamic Differential Logic) [4] as a cryptographic method in hardware for data protection. The method is based on a modification TDPL (Three-Phase Dual-Rail Pre-Charge Logic) approach which introduces a third phase of work, during which all the capacitors in the circuit are empty [5]. An important novelty in NSDDL method is its immunity on unbalanced load of the true and false outputs. In addition, the method requires only one new cell that is combined with standard logic cells.

To our knowledge the subject of test sequence synthesis and generally, testing of NSDDL based circuits was not considered in the literature and in that sense these proceedings are kind of pioneering work. Namely, the NSDDL method being based on (anti-) symmetry of two circuits named TRUE and FALSE (as will be explained later on in this paper) is by nature susceptible to faults that disturb the symmetry. From that point of view testing such circuits, or better to say, test signal synthesis should be a relatively straightforward task. It is the goal of this paper to propose a procedure for test signal synthesis and to give the first answers as to how easy the testability of this kind of circuits is. This is to be considered as a continuation of our research in testing of NSDDL circuit since in [6] fault simulation of a sequential circuit was performed.

For demonstration of the procedure we usually implement in such situations [7], in this paper, we will consider testing of one of the NSDDL cells - the XOR/XNOR circuit. In fact, after insertion of short-circuit defects in the fault free circuit, the output signal and the proper NSD value (calculated using supply current) for each defect for certain combinations of input signals will be monitored by simulation. Namely, besides examining the logic function of the circuit, it is also very important to compare the supply currents of the faulty and fault free circuits. When defect is present in the circuit, it is very likely that it will be mapped in to change of mentioned supply current.

Simulation results were obtained using ELDO simulator of Mentor Graphics Design Architect environment. To get the proper circuit parameters for simulation, layout design was performed first and post-layout parameter extraction took place. To draw the layout IC studio Mentor Graphics tools was used.

II. NSDDL METHOD

Cells resistant to SCA are based on the idea that each combination of input signals results in the same power consumption. This is possible when every logic cell has a counterpart that will react complementary. Therefore every functional cell has two outputs denoted as true and false. The hardware is doubled, but the effect of masking the true function of the cell is gained.

NSDDL method is based on the three phase clocking. The first phase named pre-charge is aimed to drive all outputs (true and false) of all logic cells to go to high logic level. In the second phase, known as evaluation phase true output takes

desired value and false output takes the complementary value. The third phase is named discharged because all outputs go to the low logic level.

The advantage of this method compared to other popular solutions, like WDDL [8], is its immunity to imbalance loads at true and false output. This is achieved by using a dynamic NOR circuit (DNOR) which minimizes the impact of short-circuit currents in the CMOS circuit. It is integral part of the control logic and NSDDL cells. Figure 1 illustrates the circuitry of DNOR cell.

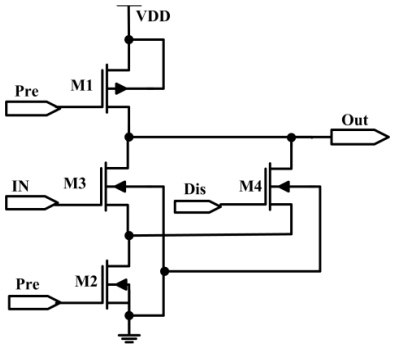


Figure 1. DNOR circuit

Figure 2 illustrates waveforms of control signals. During the pre-charge phase signals PRE=0 and DIS=0, transistorM1 is *on*, while the other transistors are *off*. The output goes to logic 1, regardless of the input signal IN. The *evaluation* phase begins when signal PRE=1 And DIS=0. Then M1 and M4 turn off, M2 is *on*, and the input signal IN controls the state of the transistor M3. If the signal IN=0, M3 is *off*, so that the output remains at logical 1. If IN=1, M3 and M2 are *on* and output switches to 0. It is obvious that the output becomes an inverting function of the input signal. Discharging phase occurs when PRE=1 and DIS=1. Therefore M3 is *off* and M4 is *on* and output goes to low logic level regardless to input signal.

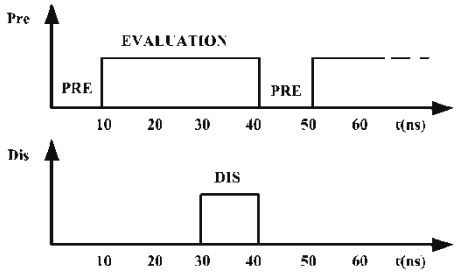


Figure 2. Time waveforms of control signals for DNOR cell

III. NSDDL XOR/XNOR CELL

Figure 3 illustrates block diagram of XOR/XNOR, SCA resistant cell. This cell consist of three cells, two NSDDL AND and one OR [9-10]. As all other NSDDL cells it has true and false inputs and output. It is clear that the same structure provides the XOR function at the true output (OT) and XNOR function at the false output (OF). Therefore it is referred to as NSDDL XOR/XNOR cell. NSDDL AND and NSDDL OR cells explore mutually complementary function, it is obvious that they can be realized using the same hardware. The only difference makes the meaning of the true and the false output.

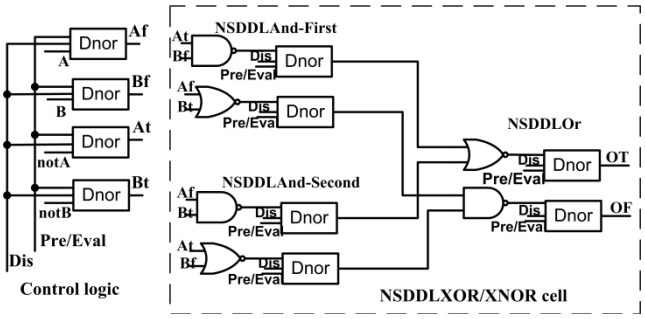


Figure 3. Block diagram of NSDDL XOR/XNOR SCA resistance cell

It is important to note that all functions are implemented using native logic circuits with negative logic (NAND i NOR) which can be easily implemented in CMOS technology. DNOR circuit represents basic element for all SCA resistant cells in NSDDL technique. It provides inverting function when transforming from standard to NSDDL logic.

Figure 4 shows layout of SCA resistant NSDDL XOR/XNOR cell. Layout of XOR and XNOR cells differs only in respect to the order of output ports which form desired functions.

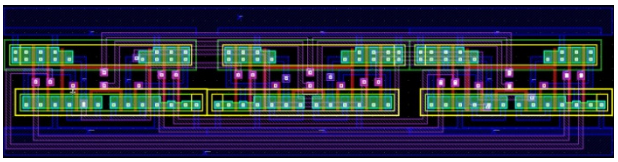


Figure 4. Layout of SCA resistant XOR/XNOR cell

For standard cells one can expect strong correlation between energy required for particular transition and combination of input signals. In particular any neutral event requires minimal energy while rise transition at the output needs more current to charge the output capacitance. NSDDL cells are designed with intention to mask cell operation regarding I_{DD}. Therefore they should provide minimal correlation between stimulus signals and I_{DD}.

Energy consumption is expressed as integral in time of power ($I_{DD} \cdot V_{DD}$) during one cycle of input signal change. For standard XOR and XNOR this cycle lasts as all three operational phases needed for NSDDL cell.

As a measure of SCA resistance we consider normalized standard deviation *NSD* according to (1).

$$NSD = 100 \cdot \frac{\sigma}{E_{avg}} \quad (1)$$

For each combination of input signals, energy is calculated. E_{avg} denotes average energy and is caluculated as the average value of energy for all input combinations. All energies for all input combinations calculated are further used to calculate the standard deviation, σ. Table I summarizes results of the comparison. Columns 1 and 2 indicate input combinations. Symbols "↑" and "↓" denote the rise and fall transition, respectively. Columns 3 and 4 present results obtained for standard XOR and XNOR cells, respectively, while column 5 refers to NSDDL cell.

TABLE I. CHARACTERISTICS COMPARISON OF CLASSIC AND NSDDL CELLS

1	2	3	4	5
A	**B**	**E_{XOR}[pJ]**	**E_{XNOR}[pJ]**	**E_{NSDDL}[pJ]**
0	↑	-0.35	-0.48	-6.38
0	↓	-0.51	-0.30	-6.21
↑	0	-0.34	-0.47	-6.22
↓	0	-0.48	-0.33	-6.22
↑	↑	-0.28	-0.05	-6.27
↓	1	-0.35	-0.47	-6.16
↑	1	-0.48	-0.31	-6.27
1	↓	-0.34	-0.47	-6.19
1	↑	-0.52	-0.32	-6.21
↓	↓	-0.27	-0.05	-6.23
E_{max} [pJ]		-0.27	-0.05	-6.16
E_{min} [pJ]		-0.52	-0.48	-6.38
E_{av} [pJ]		-0.39	-0.33	-6.24
δE [%]		63.64	131.76	3.53
σ [fJ]		91.77	154.18	56.58
NSD[%]		23.51	47.43	0.91

The simulation results for NSDDL XOR/XNOR logic cell, in the presence of a defect, are described in the next section.

IV. TESTING OF NSDDL XOR/XNOR CELL

To create a fault dictionary one is supposed to define the set of defects that are to be tested first. After that, the defect should be inserted into the circuit, one at a time, in order to analyze the effect of defect propagation. Two categories of defects are sought: catastrophic, that includes shorts and opens, and soft faults where the delay faults belong. Here only one sub-category will be considered the shorts between the transistor terminals. To get the response of the faulty circuit, namely to get the fault-effect, one has to perform electrical stimulation of the faulty circuit. Of course, a test signal is to be established beforehand that is supposed to be capable to expose the fault-effect if it is present into the response(s) of the faulty circuit.

True and *False* blocks are emphasized with dashed rectangles in Figure 5 and their outputs are denoted as *OT* and *OF*, respectively. Observing this figure, one can see that these blocks have complementary structure where *OT* depends on *At* and *Bt*, while *OF* is function of *Af* and *Bf*. Figure 6.a shows an SCA unprotected NAND cell as a generic block while Figure 6.b shows the schematic (taken from the TSMC035u library [11]) with marked defects. The same analogy is applied for NOR cell, which is presented to Figures 7.a and 7.b.

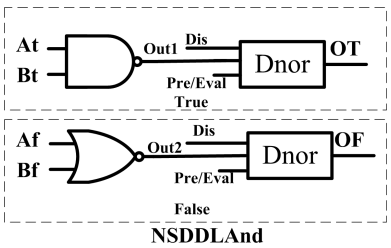


Figure 5. Block diagram of NSDDL AND cell

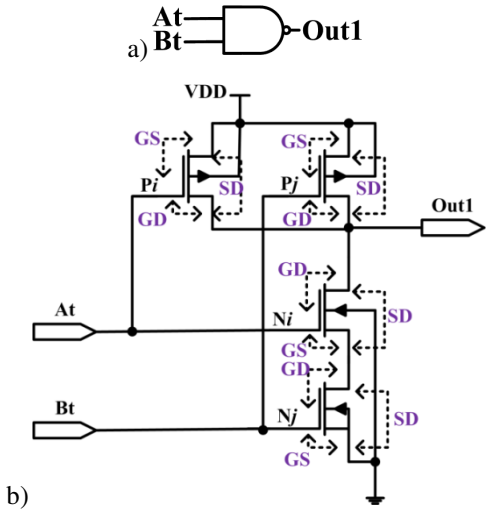


Figure 6. Standard NAND cell a) generic representation b) standard CMOS realization with marked defects

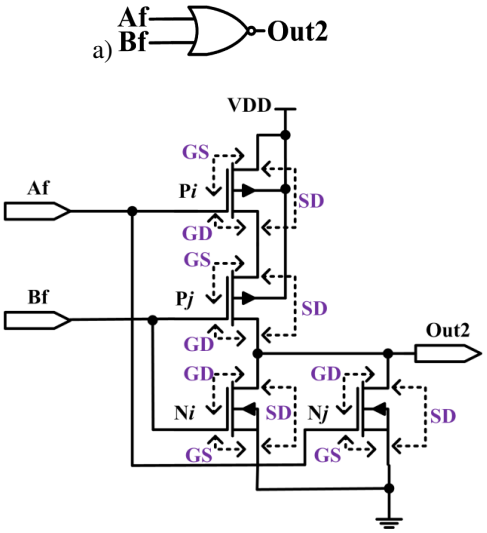


Figure 7. Standard NOR cell a) generic representation b) standard CMOS realization with marked defects

Transistors are denoted with $P_i\_F_{xy}$ or $N_j\_F_{xy}$, where P and N represent type of the transistor. Counters marked as *i* and *j* represents index of PMOS and NMOS transistor, respectively. These counters take next values: *i*=1-4, and *j*=1-4 for NSDDL AND-First cell, *i*=5-8, and *j*=5-8 for NSDDL AND-Second cell and *i*=9-12, and *j*=9-12 for NSDDL OR cell. F_{xy} denotes a short-circuit fault (hence F) between the *x* and *y* terminal of the proper transistor. Therefore *xy* can take values from the set {GD, GS, SD}, where GD stands for gate-drain, GS for gate-source, and SD source-drain.

Effect of every defect is firstly observed with respect to a logic function of the circuit. When logic function is violated it can be considered that defect is detected. Results summarized in Tables II, III and IV provides detail information about coverage of the defects for NSDDL XOR/XNOR cell in presence of short-circuit faults in NSDDL AND-First, AND-Second and OR cell, respectively. The symbol "↓" denotes the fall-transition.

TABLE II. COVERAGE OF DEFECTS FOR NSDDL XOR/XNOR CELL IN PRESENCE OF SHORT-CIRCUIT DEFECTS IN NSDDL AND-FIRST CELL

		Signal values of NSDDL XOR/XNOR cell												NSD
Fault free circuit														
At		**0**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**NA**
Bt		**1**	**0**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**NA**
Af		**1**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**1**	**NA**
Bf		**0**	**1**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**NA**
OT		**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0.91**
OF		**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
Value of OT and OF signals in presence of short-circuit defects in TRUE NSDDL AND-First sub-circuits														
P_1_F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	34.64
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_1_F_{GS}	**OT**	1	1	1	1	0	1	0	1	0	1	1	0	53.61
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_1_F_{SD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	47.69
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
P_2_F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	32.61
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_2_F_{GS}	**OT**	1	0	1	0	1	1	1	1	1	0	0	0	60.34
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_2_F_{SD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	47.69
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
N_1_ F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	34.64
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
N_1_ F_{GS}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	2.43
	OF	0	1	0	1	0	0	0	0	0	1	1	1	
N_1_ F_{SD}	**OT**	1	↓	1	↓	0	1	0	1	0	↓	↓	↓	115.82
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
N_2_ F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	33.79
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_2_ F_{GS}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	2.51
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_2_ F_{SD}	**OT**	1	0	1	0	↓	1	↓	1	↓	0	0	0	111.22
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
Value of OT and OF signals in presence of short-circuit defects in FALSE NSDDL AND-First sub-circuits														
P_3_F_{GD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	37.76
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_3_F_{GS}	**OT**	1	0	1	0	1	1	1	1	1	0	0	0	52.99
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_3_F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	110.29
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
P_4_F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	17.32
	OF	0	1	1	1	0	0	0	1	0	1	1	1	
P_4_F_{GS}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	34.12
	OF	0	1	1	1	0	0	0	1	0	1	1	1	
P_4_F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	104.12
	OF	0	1	0	1	0	0	0	0	1	1	1	1	
N_3_ F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	17.32
	OF	0	1	1	1	0	0	0	1	0	1	1	1	
N_3_ F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	2.41
	OF	0	1	0	1	0	0	0	0	0	1	1	1	
N_3_ F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	39.02
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
N_4_ F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	18.77
	OF	0	0	1	0	1	0	1	1	1	0	0	0	
N_4_ F_{GS}	**OT**	0	0	1	0	1	0	1	1	1	0	0	0	2.55
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_4_ F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	39.02
	OF	0	1	1	1	1	0	1	1	1	1	1	1	

TABLE III. COVERAGE OF DEFECTS FOR NSDDL XOR/XNOR CELL IN PRESENCE OF SHORT-CIRCUIT DEFECTS IN NSDDL AND-SECOND CELL

		Signal values of NSDDL XOR/XNOR cell												NSD
Fault free circuit														
At		**0**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**NA**
Bt		**1**	**0**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**NA**
Af		**1**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**1**	**NA**
Bf		**0**	**1**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**NA**
OT		**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0.91**
OF		**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
Value of OT and OF signals in presence of short-circuit defects in TRUE NSDDL AND–Second sub-circuits														
P_5_F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	34.04
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_5_F_{GS}	**OT**	1	0	1	0	1	1	1	1	1	0	0	0	52.99
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_5_F_{SD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	49.68
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
P_6_F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	32.77
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_6_F_{GS}	**OT**	1	1	1	1	1	1	1	1	1	0	0	0	53.52
	OF	0	1	1	1	1	0	1	0	1	1	1	1	
P_6_F_{SD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	49.68
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
N_5_ F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	34.04
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
N_5_ F_{GS}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	2.55
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_5_ F_{SD}	**OT**	1	0	1	0	↓	1	↓	1	↓	0	0	0	23.16
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
N_6_ F_{GD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	36.68
	OF	0	1	0	1	0	0	0	0	0	1	1	1	
N_6_ F_{GS}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	2.41
	OF	0	1	0	1	0	0	0	0	0	1	1	1	
N_6_ F_{SD}	**OT**	1	↓	1	↓	0	1	0	1	0	↓	↓	↓	115.72
	OF	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
Value of OT and OF signals in presence of short-circuit defects in FALSE NSDDL AND–Second sub-circuits														
P_7_F_{GD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	29.68
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_7_F_{GS}	**OT**	1	1	1	1	1	1	1	1	1	0	0	0	53.61
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_7_F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	105.32
	OF	0	1	0	1	0	0	0	0	0	1	1	1	
P_8_F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	18.73
	OF	1	0	0	0	1	1	1	0	1	0	0	0	
P_8_F_{GS}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	49.73
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_8_F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	109.25
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_7_ F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	16.83
	OF	1	0	0	0	1	1	1	0	1	0	0	0	
N_7_ F_{GS}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	2.51
	OF	0	0	0	0	1	0	1	0	1	0	0	0	
N_7_ F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	38.21
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
N_8_ F_{GD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	19.18
	OF	1	1	0	1	0	1	0	0	0	1	1	1	
N_8_ F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	2.44
	OF	0	1	0	0	0	0	0	0	0	1	1	1	
N_8_ F_{SD}	**OT**	**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	38.21
	OF	1	1	0	1	1	1	1	0	1	1	1	1	

TABLE IV. COVERAGE OF DEFECTS FOR NSDDL XOR/XNOR CELL IN PRESENCE OF SHORT-CIRCUIT DEFECTS IN NSDDL OR CELL

		Signal values of NSDDL XOR/XNOR cell												NSD
Fault free circuit														
At		**0**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**NA**
Bt		**1**	**0**	**0**	**0**	**1**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**NA**
Af		**1**	**1**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**1**	**NA**
Bf		**0**	**1**	**1**	**1**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	**NA**
OT		**1**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0.91**
OF		**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**1**	
Value of OT and OF signals in presence of short-circuit defects in TRUE NSDDL OR sub-circuits														
P_9_F_{GD}	**OT**	1	1	1	1	1	1	1	1	1	1	1	1	14.65
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
P_9_F_{GS}	**OT**	1	1	1	1	1	1	1	1	1	1	1	1	36.16
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
P_9_F_{SD}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	0.24
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
P_{10}_F_{GD}	**OT**	0	1	1	1	1	0	1	1	1	1	1	1	0.28
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_{10}_F_{GS}	**OT**	1	1	1	1	1	1	1	1	1	1	1	1	35.35
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
P_{10}_F_{SD}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	19.27
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
N_9_ F_{GD}	**OT**	0	1	1	1	1	0	1	1	1	1	1	1	14.65
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
N_9_ F_{GS}	**OT**	0	0	1	0	0	0	0	1	0	0	0	0	2.51
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
N_9_ F_{SD}	**OT**	1	1	1	1	1	1	1	1	1	1	1	1	130.19
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
N_{10} F_{GD}	**OT**	1	1	0	1	1	1	1	0	1	1	1	1	25.16
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
N_{10} F_{GS}	**OT**	1	0	0	0	0	1	0	0	0	0	0	0	2.43
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
N_{10} F_{SD}	**OT**	1	1	1	1	1	1	1	1	1	1	1	1	6.25
	OF	0	1	0	1	1	0	1	0	1	1	1	1	
Value of OT and OF signals in presence of short-circuit defects in FALSE NSDDL OR sub-circuits														
P_{11}_F_{GD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	16.41
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
P_{11}_F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	23.02
	OF	0	1	1	1	1	0	1	1	1	1	1	1	
P_{11}_F_{SD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	19.58
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
P_{12}_F_{GD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	15.31
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
P_{12}_F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	35.95
	OF	1	1	0	1	1	1	1	0	1	1	1	1	
P_{12}_F_{SD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	27.05
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
N_{11} F_{GD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	15.31
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
N_{11} F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	2.51
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
N_{11} F_{SD}	**OT**	1	0	1	0`	0	1	0	1	0	0	0	0	8.51
	OF	0	1	↓	1	1	0	1	↓	1	1	1	1	
N_{12} F_{GD}	**OT**	1	0	1	0`	0	1	0	1	0	0	0	0	16.56
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
N_{12} F_{GS}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	2.54
	OF	0	0	0	0	0	0	0	0	0	0	0	0	
N_{12} F_{SD}	**OT**	1	0	1	0	0	1	0	1	0	0	0	0	33.78
	OF	↓	1	0	1	1	↓	1	0	1	1	1	1	

Observing results given in Tables II and III one can see that defects, in the TRUE sub-circuit, can be mapped to both outputs, OT and OF. The same applies for FALSE sub-circuit. Defects which do not affect both outputs OT or OF are marked with a bold. This is the case with NSDDL OR cell. These results are given in Table IV. Here, e.g. if the defect is in TRUE sub-circuits, its effect will be demonstrated only at OT output. Similarly, effect of the defect in the FALSE sub-circuit will be visible only at OF output. Since operation of the circuit is very specific, logic function is observed during EVALUATION phase for fault free and faulty circuits under the same input conditions.

Testing based on the supply current is an excellent supplement to the testing of logic functions of a circuit. As discussed above, *NSD* parameter directly depends on the I_{DD} and for that reason, this parameter is used as a second indicator. It can be seen form Table II and Table III that both criteria indicated the presence of a defect in a circuit for any simulated case. This means that defect coverage is 100% by the test signal given in the first two rows of the tables. This confirms that rough destruction (catastrophic fault presence) of the NSDDL's circuit symmetry has apparent influence to its response. That is important for testing but also for evaluating its main function. Namely, in the presence of a fault the circuit is not so effective in data protection.

V. CONCLUSION

In this paper, we showed how individual defects in lower level circuits (OR, AND) affect behaviour of cells at higher level (XOR/NXOR). For testing NSDDL XOR/XNOR cell two criteria were examined: logic function verification and IDDQ testing performed by calculating the *NSD* parameter. Seventy two simulations were performed in order to make the appropriate fault dictionary for defects of short-circuit type. After completing the test synthesis procedure for a XOR/XNOR gate one may conclude that expected results were obtained. Namely, both criteria give excellent coverage of defects. All seventy two defects were detected in either case. In fact the symmetry being violated by insertion of a fault, the fault effect is immediately visible at the output. It is important to mention that the individual defects inserted in the input NSDDL-And circuits of the NSDDL XOR/XNOR cell can be demonstrated on both outputs. This is not the case for defects in output NSDD-OR circuitry where effect of the defect is visible only at the one of the outputs (true or false). The number of defects visible on both outputs for is twenty eight.

ACKNOWLEDGMENT

This research was funded by The Ministry of Education, Science and Technological Development of Republic of Serbia under contract No. TR32004.

REFERENCES

[1] Koc, Cetin Kaya (Ed.) *Cryptographic Engineering*, Springer, 2009.

[2] Petković P., Stanojlović M. and Litovski V. "*Design* of side-channel-attack resistive criptographic ASICS", Forum BISEC 2010, Zbornik radova druge konferencija o bezbednosti informacionih sistema, Beograd, Srbija, Maj 2010, pp 22-27.

[3] Stanojlović M. and Petković P., *"Hardware based* strategies against side-channel-attack implemented in *WDDL",* Electronics, Vol. 14, No. 1, Banja Luka, June, 2010, pp. 117-122

[4] J. Quan and G. Bai, "A new method to reduce the side-channel leakage caused by unbalanced capacitances of differential interconnections in dualrail logic styles", 2009 Sixth International Conference on Information Technology: New Generations, DOI 10.1109/ITNG. 2009.185, pp. 58-63.

[5] M. Bucci, L. Giancane, R. Luzzi, A. Trifiletti: *"Three-Phase Dual-Rail Pre-Charge Logic"*. In: Goubin, L., Matsui, M. (eds.) CHES 2006. LNCS, vol. 4249, pp. 232–241. Springer, Heidelberg (2006).

[6] Stanojlović, M. and Litovski, V., "*Simulation of defects in sequential NSDDL Master/Slave D flip flop circuit*", Proceedings of Small Systems Simulation Symposium 2012, Niš, Serbia, 12th-14th February 2012

[7] Milovanović, D. B., and Litovski, V. M., "*Fault models of CMOS Circuits*", Microelectrinics Reliability, Vol. 34, No. 5, pp. 883-896, 1994

[8] Danger, J.-L. Guilley, S. Bhasin, S. Nassar, M., "Overview of Dual Rail with Precharge Logic Styles to*Thwart Implementation-Level Attacks on Hardware Cryptoprocessors",* Proc. of International Conference on Signals, Circuits and Systems SCS'2009, Djerba, Tunisia, November 5-8 2009, pp. 1-8

[9] Stanojlović, M., Petković, P.: „*An ASIC cryptoosystem resistant to side channel attacks based on standard cells*", VIII Simposium on Industrial Electronics INDEL 2010, Banja Luka, Bosnia and Herzegovina, 4-6 November, 2010, pp. 110-114, ISBN 978-99955-46-03-8, In Serbian

[10] Petković, P., Stanojlović, M.: „*Hardware protection from side channel attacks based on masking the consumption information"*, Zbornik LV konferencije ETRAN, Banja Vrućica, Teslić, B&H, 2011, ISBN 978-86-80509-66-2.

[11] ASIC Design Kit, http://www.mentor.com/company/higher_ed/ic-asic

Bandgap Voltage Reference in 130nm: Design and Schematic Level Simulation

Aleksandar Pajkanovic
Faculty of Technical Sciences
University of Novi Sad
apaj@uns.ac.rs

Veljko Malbasa
Faculty of Technical Sciences
University of Novi Sad
malbasa@uns.ac.rs

***Abstract*—A bandgap voltage reference design based on a simple topology utilizing a current mirror and a complementary to absolute temperature voltage source is presented in the paper. Simulation results for implementation in 130 nm CMOS standard process show temperature variation of the output voltage within 5 mV over a temperature range of 165 K (from -40°C up to 125°C) in the nominal case, i.e. the absolute reference voltage temperature variation is 0.19 %, i.e. 11.5 ppm/°C. The influences of supply voltage and process variations (including corner and Monte Carlo analysis) on the output reference voltage are presented and discussed. Time domain analysis shows that the circuit is fully operational within 112 ns. The results are obtained through the schematic level simulations using Spectre Simulator from Cadence Design System.**

Keywords-CMOS, integrated circuits, bandgap, voltage reference, temperature stability

I. Introduction

The design of DC current and voltage sources internal to the chip is an essential step in monolithic integrated circuits design. The DC voltage and current references represent one of the basic blocks of analog electronics. Namely, the performance of nonlinear analog components depends strongly on the current flowing through them, i.e. biasing conditions. The task that these references should accomplish is to properly bias analog components of integrated circuits. Among a variety of solutions, efficient, simple and easy-to-design structures are the priority. These reference circuits should be autonomous when the supply voltage is switched on. Further, they should be based on intrinsic physical properties or on reproducible technology parameters [1], [2].

A process, voltage and temperature invariant reference voltage within an integrated circuits means stable biasing conditions - obtaining such a circuit for utilization in more complex systems is the main motivation for this work.

In this paper a simple topology is employed to design and simulate a bandgap voltage reference circuit intended to provide a temperature stable output voltage of around 1.2 V over a range from -40°C to 125°C. The technology process used is 130 nm, while the nominal supply voltage is 3.3 V and the design model is ACM model presented in [1]. The temperature coefficient achieved is 11.5 ppm/°C and the sensitivity of the reference voltage to the supply is 134 ppm/°C. Besides these, process variations are analyzed and treated through corner and Monte Carlo (statistical) analysis. Time response is also given to show that the circuit operates in a stable operating point, while the start-up time is 112 ns. The results listed are obtained through schematic level simulations using Spectre Simulator from Cadence Design System. In the end, the characteristics of this work are compared to several other such circuits found in literature and the possibilities of improvements are discussed.

II. Model Used

The design methodology is based on the ACM model [1], [3]. It is a current-based MOSFET model that uses the concept of inversion level. According to the ACM model, the drain current can be split into the forward (I_F) and reverse (I_R) currents:

$$I_D = I_F - I_R = I_S(i_f - i_r), \tag{1}$$

where I_S is the normalization specific current, and i_f and i_r are inversion levels, forward and reverse, respectively. The forward and reverse currents depend on the gate to source and gate to drain voltages, respectively. If the transistor operates in saturation, we have:

$$I_F \gg I_R, \text{ thus: } I_D \approx I_F = I_S \cdot i_f. \tag{2}$$

The normalization current is a function of the technology:

$$I_S = \frac{\mu C_{ox} \phi_T^2 n}{2} \frac{W}{L}, \tag{3}$$

where μ represents the charge mobility, C_{ox} gate-oxide capacitance per area, ϕ_T thermal voltage, n slope factor and W/L transistor aspect ratio.

The inversion level value signifies the transistor operation region in the following way: if $i_f < 1$, the transistor operates in weak inversion and if $i_f > 100$ the transistor operates in strong inversion region. If $1 < i_f < 100$, the transistor operates in moderate inversion region.

According to this model, the voltage and current are related in the following manner [1]:

$$\frac{V_P - V_{S(D)}}{\phi_T} = \sqrt{1 + i_{f(r)}} - 3 + \ln\left(\sqrt{1 + i_{f(r)}} - 1\right), \tag{4}$$

where V_P is the pinch-off voltage, given as the difference of the gate potential and the zero bias threshold voltage (modified by the slope factor):

$$V_P \approx \frac{V_G - V_{T0}}{n}, \text{ if } V_G \approx V_{T0}. \tag{5}$$

Further detailed description of the model can be found in [1] and [3].

III. Bandgap Voltage Reference Theory of Operation

A bandgap voltage reference operation is based on adding two voltages having equal and opposite temperature coefficients, i.e. a voltage proportional to absolute temperature (PTAT) and a voltage complementary to absolute temperature (CTAT). The logic behind this idea is to cancel out the temperature dependence, so that the output voltage, i.e. reference voltage is independent of temperature variation. Therefore, the reference voltage, V_{REF} we can write as:

$$V_{REF} = \alpha \cdot V_{CTAT} + \beta \cdot V_{PTAT}. \tag{6}$$

where α and β are coefficients of proportionality.

The temperature coefficient of a voltage is numerically expressed by the first derivative of that voltage as a function of temperature. Thus, in order to accomplish that V_{REF} is independent of temperature, its derivative needs to be zero:

$$\frac{\partial V_{REF}}{\partial T} = \alpha \frac{\partial V_{CTAT}}{\partial T} + \beta \frac{\partial V_{PTAT}}{\partial T} = 0. \tag{7}$$

To achieve this, a CTAT and a PTAT voltages need to be designed and, then, the coefficients α and β need to be set in such a way that slopes of these two voltages are equal.

The easiest way to obtain a CTAT voltage is the forward-biased pn junction - a diode. The diode current is given by [1]:

$$I_D = I_S \cdot e^{\frac{V_D}{\phi_T}}, \tag{8}$$

where I_S represents a diode saturation current and V_D is the diode voltage. From this relation, we extract the diode voltage as:

$$V_D = \phi_T \ln\left(\frac{I_D}{I_S}\right). \tag{9}$$

The two members comprising the expression for the diode voltage are of opposite temperature coefficient, but the PTAT (the thermal voltage) is negligible when compared to CTAT (the logarithm expression). Namely, the diode voltage temperature coefficient is given as [4]:

$$\frac{\partial V_D}{\partial T} = \frac{V_D - (4+m)\,\phi_T - \frac{E_g}{q}}{T}, \tag{10}$$

where $m \approx -3/2$ is the exponential coefficient of the mobility dependence on temperature (expressed as $\mu \propto T^m$) and $E_g \approx 1.12$ eV represents the bandgap energy of silicon [4].

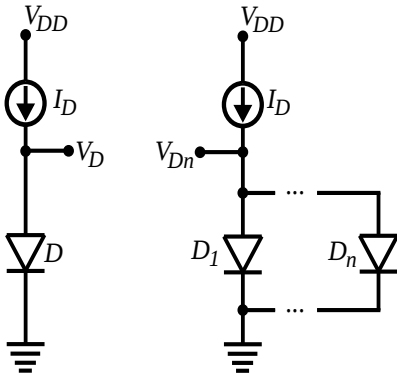


Figure 1. A PTAT voltage can be obtained as the difference of V_D and V_{Dn}

Therefore, the diode voltage is a CTAT in its nature. Assuming a standard value of $V_D = 0.75$ V and measuring at room temperature, (10) yields a temperature coefficient of around -1.5 mV/K.

To obtain a PTAT voltage, a pn junction is again used, i.e. several pn junctions are used. Namely, the idea is to subtract two V_D voltages, thus canceling the CTAT members and extracting the thermal voltage, i.e. the PTAT member. Such an effect can be obtained if there are different currents flowing through physically identical pn junctions, Fig. 1. Assuming an array of n identical diodes (Fig. 1), according to (9), the voltage of each one of them is given as:

$$V_{D1} = V_{D2} = ... = V_{Dn} = \phi_T \cdot \ln\left(\frac{I_D}{n \cdot I_S}\right). \tag{11}$$

Observing the difference of V_D and V_{Dn}, yields:

$$\begin{aligned} V_D - V_{Dn} &= \phi_T \cdot \ln\left(\frac{I_D}{I_S}\right) - \phi_T \cdot \ln\left(\frac{I_D}{n \cdot I_S}\right) \\ &= \phi_T \cdot \ln(n), \end{aligned} \tag{12}$$

which means that this voltage difference is PTAT in its nature.

The subtraction can be performed by inserting a resistor, as shown in Fig. 2. The voltage over R_1 is given as:

$$V_{R1} - V_{Dn} = I_D \cdot R_1. \tag{13}$$

If V_D and V_{R1} are equal, we can write:

$$V_{R1} - V_{Dn} = I_D \cdot R_1 = \phi_T \cdot \ln(n). \tag{14}$$

Therefore, in order for this idea to work, the current generators in Fig. 2 must be replaced by a circuit which will make sure that $V_D = V_{R1}$. In this circuit, the current I_D is completely determined by the resistor R_1.

There are different ways to provide equal voltage levels at points V_D and V_{R1}, including the application of a current mirror and an operational amplifier [1], [4]. Within this paper, the current mirror is used because of its simpler realization and a lower number of transistors. Thus, the ideal current generators in Fig. 2 are replaced by a current mirror, as shown in Fig. 3.

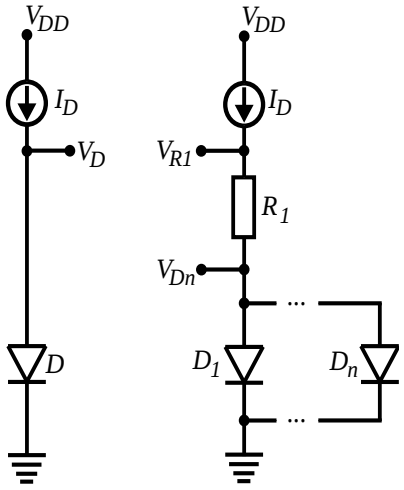


Figure 2. To obtain a PTAT, voltages V_D and V_{R1} must be equal

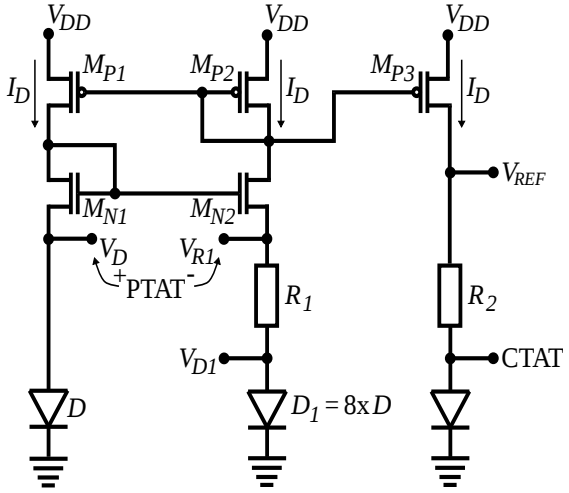


Figure 3. Proposed circuit schematic

The temperature coefficient of $V_{R1} - V_{Dn}$ is that of the thermal voltage:

$$\frac{\partial \phi_T}{\partial T} = \frac{k}{q} = 86.25 \frac{\mu \mathrm{V}}{\mathrm{K}}. \tag{15}$$

Now, both CTAT and PTAT voltages are provided, even though of unequal temperature coefficients. This issue is solved by setting the coefficients of proportionality, α and β, in the correct way. The last step is to sum these two voltages so that the whole circuit yields an output voltage constant to temperature variation. The easiest way to sum voltages is to simply connect them in series. This is what the third branch in Fig. 3 actually does. Namely, the output transistor, M_{P3}, copies the current from the current mirror. Thus, the voltage V_{REF} is given as:

$$V_{REF} = V_D + V_{R2} = V_D + I_D \cdot R_2. \tag{16}$$

From (14) we can write:

$$I_D = \frac{\phi_T \cdot \ln(n)}{R_1}, \tag{17}$$

which, used in (16), yields:

$$V_{REF} = V_D + \frac{R_2}{R_1} \cdot \ln(n) \cdot \phi_T. \tag{18}$$

Comparing this expression with (6), we realize the following identities: $V_{CTAT} = V_D$, $V_{PTAT} = \phi_T$ and $\alpha = 1$, $\beta = (R_2/R_1) \cdot \ln(n)$. Substituting according to these relations into (7), we obtain:

$$\frac{\partial V_D}{\partial T} + \frac{R_2}{R_1} \cdot \ln(n) \cdot \frac{\partial \phi_T}{\partial T} = 0. \tag{19}$$

Temperature coefficients of pn junction and thermal voltages are known from (10) and (15), respectively, whereas R_1, R_2 and n remain to be determined during circuit design.

IV. Circuit Design Using ACM Model

Some of the parameter values must be assumed. For example, since this is a DC circuit and it is of primary importance that all of the transistors be as equal as possible, we are allowed to choose a wide channel transistor - therefore, $W = 10$ μm. In this way, using larger transistors, matching is improved [4]. For the same reason, the number of parallel diodes in the second branch is chosen as $n = 8$. Namely, this means that there will be 9 diodes in the PTAT circuit, that can, thus, be positioned in a matrix of 3x3. In that way, all of the matching conditions will be fulfilled, and the probability of system mismatch highly lowered [4]. Further, a somewhat larger resistor is chosen, $R_1 = 1$ kΩ. Further, for the proposed circuit to function properly, all transistors must operate in saturation. Obviously, transistors M_{P1}, M_{P2} and M_{N1} will always operate in saturation, but for M_{N2}, this condition must be secured through component design. According to the ACM model, that means: $i_f \gg i_r$. To make sure that they will all remain saturated and in strong inversion in all cases, we assume $i_{f1} = 200$.

Using the transistor model approximate parameters, it is possible to calculate the sheet normalization current:

$$I_{SH} = \frac{I_S}{W/L} = \frac{\mu C_{ox} \phi_T^2 n}{2} = 61 \text{ nA}. \tag{20}$$

According to (17), the drain current, i.e. the current of each of the branches, is determined by the resistor R_1. Therefore, it can be calculated as $I_D \approx 53\mu$A. Now it is possible to calculate the transistor ratio:

$$\frac{W}{L} = \frac{53 \cdot 10^{-6}}{200 \cdot 61 \cdot 10^{-9}} = 4.4 \approx 5, \tag{21}$$

which means that $L = 2$ μm.

The final unknown value of this circuit is the resistivity of R_2, which is determined from the temperature invariance condition, (19). Even though an approximate value for $\partial V_D / \partial T$ is already mentioned in section III of this paper, it varies significantly for each technology. Therefore, the actual value of this coefficient for this particular technology process is obtained through an experiment: a simple DC simulation yields $\partial V_D / \partial T \approx -1.115$ mV/K, for the range of temperatures from -40 up to 125°C.

Substituting now the obtained temperature coefficients and the rest of the known values, we write an equation with only one unknown:

$$-1.115 \cdot 10^{-3} + \frac{R_2}{10^3} \cdot \ln 8 \cdot 86.25 \cdot 10^{-6} = 0, \qquad (22)$$

which provides a solution to the unknown value of the final circuit component, the resistor $R_2 \approx 6500\ \Omega$.

V. Simulation Results

Since we used approximate values for the calculations performed in section IV of this paper, some fine tuning is expected. After several iterations of simulations, these are the final parameter values of all the components used in circuit in Fig. 3 all transistor channel widths are 10 μm and all lengths are 2 μm, n =8 $R_1 = 1$ kΩ and $R_2 = 6.86$ kΩ.

In several simulations, the characteristics of the designed bandgap reference voltage circuit are demonstrated. First, the output of the circuit, i.e. the reference voltage, is shown as a function of temperature and of supply voltage. Then, process variations are simulated, observing the worst case scenarios through corner analysis and the statistical mismatch through Monte Carlo analysis. In the end, transient analysis is performed to show that the circuit's operating point is stable.

From Fig. 4 we see that the output voltage varies only $\pm$2.370 mV around the mean value of 1245.341 mV over a temperature range 160 K wide, i.e. from -40 to 125 °C. This is an error of 0.19 % and of 0.0012 % per °C, i.e. 11.5 ppm/°C.

Supply voltage variations are demonstrated in Fig. 5. The nominal value of V_{DD} is 3.3 V and here the circuit behavior at $\pm$10 % is shown. The simulation results show that, with the supply voltage change from 2.97 V to 3.63 V, the reference voltage changes its value within the border of $\pm$25.725 mV around the mean value of 1203.494 mV. Relative error is 2.14 % and 0.0134 % per °C, i.e. 134 ppm/°C.

In Figs. 6-9 process variations of the circuit in question are analyzed. Corner and Monte Carlo analysis are performed on top of temperature (Figs. 6 and 8) and on top of supply voltage (Figs. 7 and 9) variation analysis. Through corner analysis, worst case scenarios are treated, showing that the maximum possible change of the output voltage taking into account both temperature and process variations is around 30 mV, which is $\pm$1.24 %. Also, the maximum possible variation of the output voltage as a consequence of both supply and process variations is around 53 mV, i.e. $\pm$ 2.20 %.

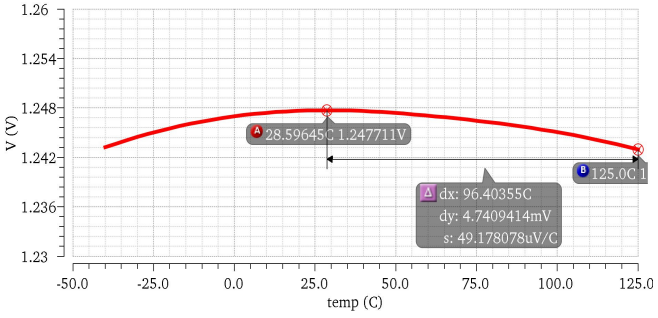


Figure 4. Temperature dependence of the output voltage at nominal supply voltage

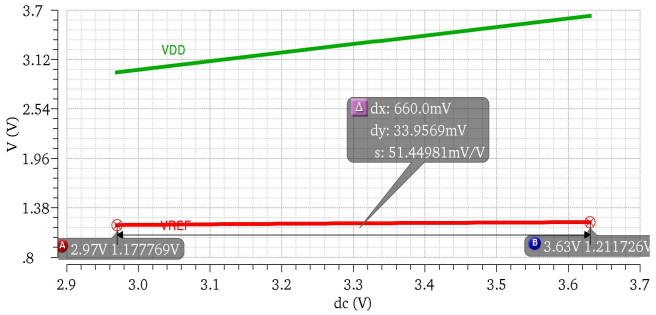


Figure 5. Output voltage dependence on the supply voltage

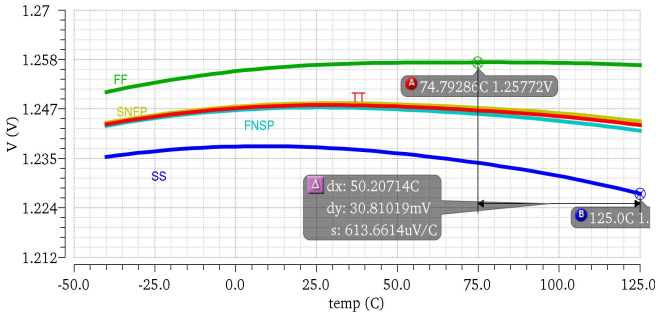


Figure 6. Corner analysis performed on top of the output voltage temperature dependence analysis

Statistical variations of the process are analyzed through Monte Carlo simulations, each with 30 iterations. According to the technology process documentation, statistical significance of 30 iterations is quite high. Namely, if the circuit operates correctly for all 30 runs, there is 99 % probability that over 80 % of all possible component values operate correctly. Statistical analysis performed on top of temperature analysis shows that the output voltage change is within $\pm$32.96 mV, which is $\pm$2.65 %, Fig 8. The same analysis performed on top of supply voltage analysis shows variance of the reference voltage within $\pm$49.34 mV, which is $\pm$4.10 % of the mean value, Fig. 9.

In Fig. 10 time response of the circuit with start-up circuitry added is shown. The addition does not influence DC operation of the circuit and it contains two PMOS transistors and a capacitor. Its purpose is to enable the operation of the bandgap reference voltage in its stable operating point [1]. The time domain analysis shows that the circuit provides nominal output voltage within 112 ns.

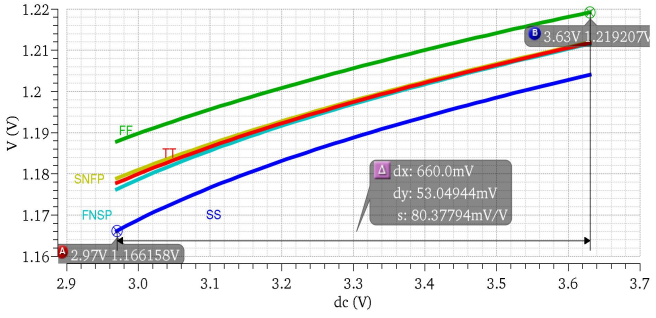


Figure 7. Corner analysis performed on top of a supply voltage analysis of the output voltage

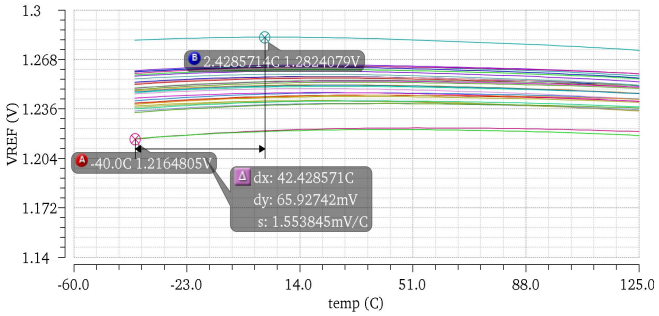


Figure 8. Statistical process variation observed through the Monte Carlo analysis of the temperature dependence of the output voltage

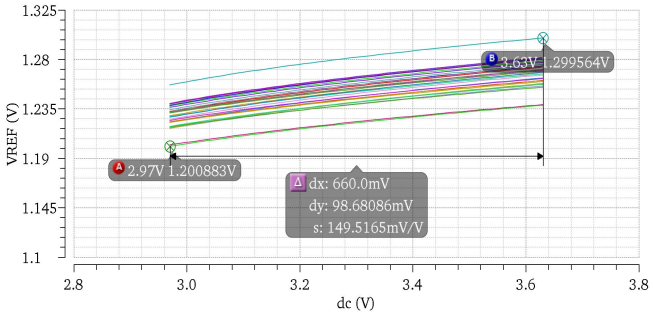


Figure 9. Statistical process variation observed through the Monte Carlo analysis of the output voltage dependence on the supply voltage

VI. Discussion

In Tab. I the characteristics of several circuits of the same purpose are compared. In [5] a bandgap voltage reference is implemented through a topology employing a current-voltage mirror instead of an amplifier, while in [6] a two-stage amplifier is used. These two are characterized by a somewhat larger temperature coefficient and lower power consumption, but at a narrower temperature range. Circuits presented in [7] and [8] show very weak variance of the reference voltage as a function of temperature over a wider temperature range. Also, these are characterized by higher power consumption, up to 1.3 mW. The topologies implemented use operational amplifier, with the addition of a 4-bit trimming circuit in [8]. The purpose of this additional circuit is to reduce the influence of the process variations.

The circuit presented in this paper shows temperature coefficient in the higher end of the shown scale, but of the same order of magnitude. In the context of power consumption, our work belongs in the middle with $P_D = 600\mu$W. Concerning topology, this is the simplest of those investigated, since it employs only a current mirror and a start-up circuit.

The circuit presented in this paper shows satisfying behavior concerning temperature variations. On the other hand, its supply voltage variation is larger. This is a consequence of a very simple circuit employed to force equal voltages V_D and V_{R2}. This current mirror's responsibility is to prevent the influence of supply voltage's disturbances of the output voltage, V_{REF}. A way to improve this feature of the circuit is to use an operational amplifier with extremely high gain in order to stabilize the noise coming from V_{DD}. In future work, such an attempt is to be made.

The purpose of this comparison in no way is a claim which of the circuits is better, since each one of them is designed with

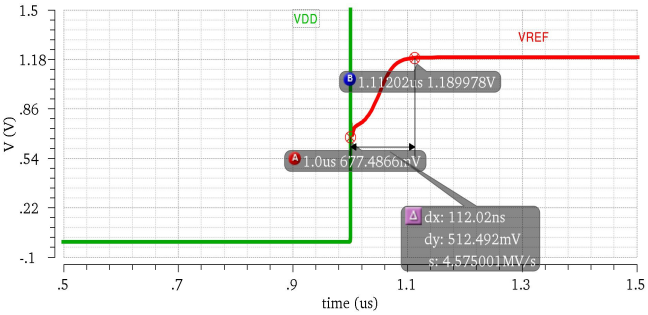


Figure 10. Time domain analysis, showing that the circuit starts-up in 112 ns

TABLE I. Characteristics comparison

Circuits	Technology process [nm]	Supply voltage [V]	Temperature coefficient [ppm/°C]	Temperature range [°C]	Power [μW]
[5]	350	2.5	12.1	5:95C	276
[6]	180	1.8	6.5	-20:90	150
[7]	130	3.3	5.5	-40:105	1300
[8]	180	1.8	2	-20:140	620
This work	130	3.3	11.5	-40:125	600

the author having a different goal in mind, e.g. low power, higher stability or smaller chip area. The Tab I is given in order to point out that the results obtained within this paper are of the same order like those found in current literature.

VII. Conclusion

A temperature stable bandgap reference voltage circuit is presented in this paper. Process variation analysis performed (both corner and Monte Carlo) show that the output is influenced by the fabrication process, and this characteristic remains to be worked upon during the layout phase, which follows. Supply voltage sensitivity is somewhat higher, but a way to improve this feature is given and is to be implemented in future work.

Acknowledgment

This work is funded from the European Union's Seventh Framework Program for research, technological development and demonstration under grant agreement no. 289481. - project SENSEIVER.

References

[1] M. C. Schneider and C. Galup-Montoro, *CMOS Analog Design Using All-Region MOSFET Modeling*. Cambridge University Press, 2010.

[2] T. Lee, *The RF CMOS Integrated Circuit Design*. Prentice Hall, 1998.

[3] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, "An MOS Transistor Model for Analog Circuit Design," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 10, pp. 1510–1519, 1998.

[4] B. Razavi, *Design of Analog CMOS Integrated Circuit*. McGraw-Hill, 2001.

[5] Y.-H. Lam and W.-H. Ki, "CMOS Bandgap References with Self-biased Symmetrically Matched Current-Voltage Mirror and Extension of Sub-1-V Design," *Very Large Scale Integrateion (VLSI) Systems, IEEE Transactions on*, vol. 18, no. 6, pp. 857–865, 2010.

[6] Y. Yusoff, H. C. Lah, N. Razali, S. N. Harun, and T. Yew, "Design and Characterization of Bandgap Voltage Reference," in *Semiconductor Electronics (ICSE), 10th IEEE International Conference on*, 2012, pp. 686–689.

[7] M. Lazic, D. Mirkovic, and P. Petkovic, "Integrated Voltage Reference Design in 130 nm CMOS Technology Process (in Serbian)," in *Proceedings of 57th ETRAN Conference, Zlatibor, Serbia*, 2013, pp. EL2.1.1–6.

[8] A. Martinez-Nieto, M. Sanz-Pascual, P. Rosales-Quintero, and S. Celma, "A Bandgap Voltage Reference in 0.18μm CMOS Technology," in *Circuits and Systems (MWSCAS), IEEE 56th International Midwest Symposium on*, 2013, pp. 97–100.

Operational Transconductance Amplifier in 350nm CMOS technology

Dejan Mirković, Predrag Petković
Department of Electronics, LEDA laboratory
University of Nis, Faculty of Electronic Engineering
Niš, Serbia
{dejan.mirkovic, predrag.petkovic}@elfak.ni.ac.rs

Ilija Dimitrijević, Igor Mirčić
Department of Electronics
University of Nis, Faculty of Electronic Engineering
Niš, Serbia
idimitrijevic@elfak.rs

***Abstract*— This paper presents transistor level design of operational transconductance amplifier in CMOS technology. Custom designed, circuit is to be built-in into the mixed-signal, switched capacitor circuit. Amplifier targets relatively high slew-rate and moderate open loop gain with megahertz order gain-bandwidth. Adopted architecture is discussed appreciating application in switched capacitor circuits. Circuit behavior is examined through set of simulations. Obtained results confirmed desired behavior. Target technology process is TSMC 350nm.**

Keywords-Integrated circuit; Amplifier; Switched capacitor circuits; CMOS technology

I. INTRODUCTION

Operational Transconductance Amplifier (OTA) is considered to be the very fundamental part of analog integrated circuits (IC). Almost every on-chip active block for analog signal conditioning is built on top of it. Concerning this fact switched capacitor (SC) circuits are not exception. Design covered in this work is meant to be embedded into analog part of the second order ΔΣ analog-to-digital converter (ADC) discussed in [1]. Since ADC consists of SC circuits there is inherent request for relatively high slew-rate and gain-bandwidth. As shown in [2], open loop i.e. DC, gain has the smallest influence on SC circuit characteristics therefore moderate open loop gain is sufficient. Table I summarizes main OTA design parameters set by the higher order circuit requirements.

TABLE I. TARGET OTA PARAMETERS

Parameter	*Description*	*Value*
A_0	DC, open loop, gain	> 50 dB
f_{gbw}	Gain-bandwidth	> 120 MHz
SLR	Slew rate	> 120 V/μs

Parameters like, input/output dynamic range (DR), common mode (CMRR) and power supply (PSRR) rejection ratios should be as large as possible. Since TSMC 350nm technology process supports relatively high, 3.3V, power supply voltage this requirements are expected to be fulfilled.

Circuit supposed to be fully differential which implies utilizing some form of common mode feedback (CMFB) circuitry. Also OTA has to have its own bias point generator in order to provide appropriate transistor operation. Since on-chip capacitors are considered, 2pF differential load capacitance is adopted. This value is also set by higher order circuit requirements concerning kT/C noise of ΔΣ structure explained in [2]. It should be mentioned that target technology process offers Poly-insulator-Poly (PiP) capacitors with 864 aF/μm^2 capacitance per unit area. Hence the value of 2 pF for load capacitance gives reasonably high capacitor area of 2314.81 μm^2 (48.11μm x 48.11μm).

Paper is organized as follows. In second section adopted OTA architecture will be briefly discussed and appropriate subsections will cover circuitry in more details. Third section presents simulation results. Finally, in the fourth section, educative conclusions are drawn and possible improvements are discussed.

II. OTA ARCHITECTURE

Although folded cascode (FC) architecture is commonly adopted for building SC circuits; telescopic architecture is chosen for OTA design in this case. Some related work supporting this idea is published in [3],[4]. It is well known that FC provides wider input common mode range, better input-output common mode relation and high input/output swing [5]. All those advantages imply higher power consumption, lower gain, higher noise and, most importantly in this case, lower speed (i.e. slew-rate and gain-bandwidth). Choosing telescopic architecture means stricter constraint on input/output common-mode voltage choice. Transistor level schematic of OTA with bias and CMFB circuitry is depicted in Fig. 1. Dimensions of all transistors are summarized in Table II. Design can be partitioned in three sub-blocks namely: Core, CMFB and Bias.

A. Core

Transistors M0-M8 are the core of the design. Analyzing structure utilizing simplified model open loop gain is:

$$A_0 \approx g_{m1,2}(g_{m3,4}r_{03,4}r_{01,2} \,\|\, g_{m5,6}r_{05,6}r_{07,8}). \qquad (1)$$

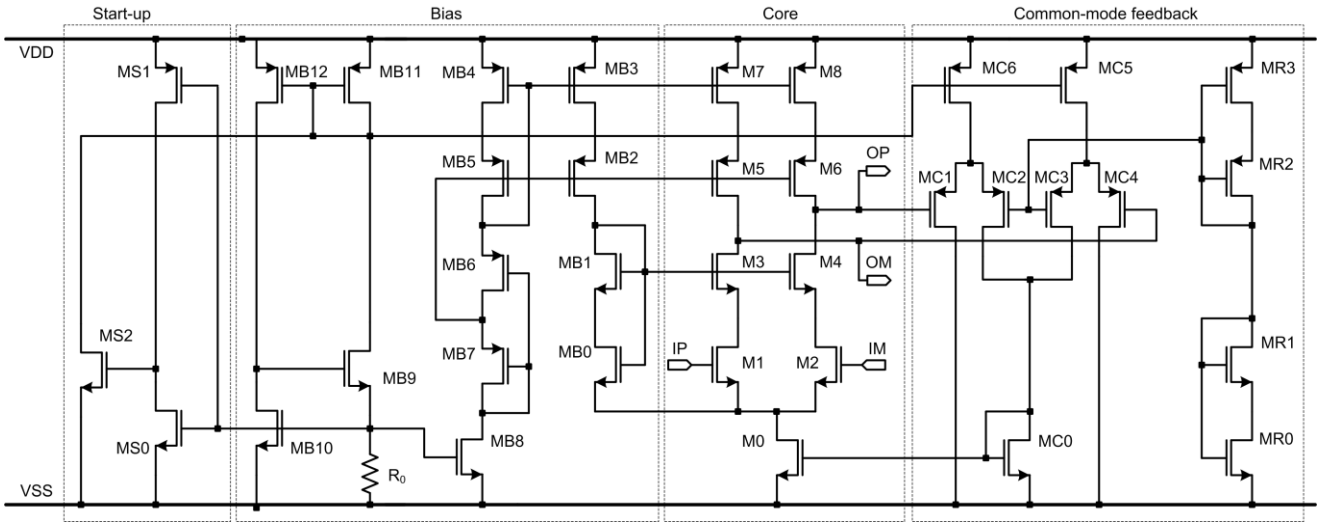


Figure 1. Telscopic OTA with bias and CMFB circuitry

TABLE II. TRANSISTOR DIMENSIONS

Transistors	width/length [µm/µm]
(M1, M2, M3, M4),(M0, M5, M6, M7, M8)	(128/0.8), (256/0.8)
(MB1, MB8), (MB2, MB3, MB4, MB5, MB7, MB10, MB9),(MB11, MB12), MB0,MB6, MS0, (MS1, MS2)	(128/0.8), (256/0.8), (512/0.8), 24/0.8, 88/0.8, 400/0.8, (4/0.8)
(MC0, MC1, MC2, MC3, MC4), (MC5, MC6)	(6.4/0.8), (12.8/0.8)
(MR1, MR3), MR2, MR0	(2.4/0.8), 7.2/0.8, 0.8/0.8

Cascode configuration by itself provides sufficient DC gain and (1) is expected to meet the DC gain requirements. Being single stage, there is no need for frequency compensation. Stability is also guaranteed by relatively large, 2pF differential load capacitance, C_L. Therefore gain-bandwidth is mainly determined by transconducatance of amplifying devices, $g_{m1,2}$, and load capacitance ratio. Design procedure is as follows. Transconductance of M1 and M2 devices should satisfy the following equality:

$$g_{m1,2} = 2\pi f_{gbw} C_L . \qquad (2)$$

For given gain-bandwidth and load capacitance, g_m equals to about 1.5mS. Taking into account fully differential case this value is doubled. In order to properly size amplifying devices, characteristics shown in Fig. 2, 3 and 4 are addressed. All those curves are extracted using SPICE [6]. Fig. 2 shows composite figure of merit i.e. unity current gain frequency, f_t, and transistor efficiency, g_m/I_D, product versus overdrive voltage, $V_{ov}=V_{GS} - V_{TH}$. Observing these curves for different channel lengths one can find optimum bias point which compromises between speed and efficiency. This value is about 200mV.

Knowing this, the channel current can be extracted. Namely, for overdrive voltage of 200mV Fig. 3 indicates the efficiency of about $10V^{-1}$ which gives the channel current I_D= 300µA.

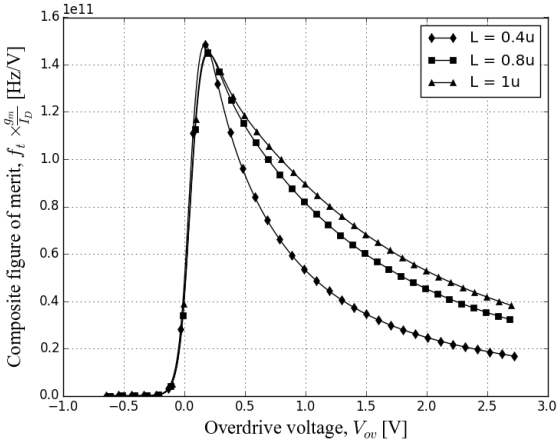


Figure 2. Composite figure of merit versus overdrive voltage

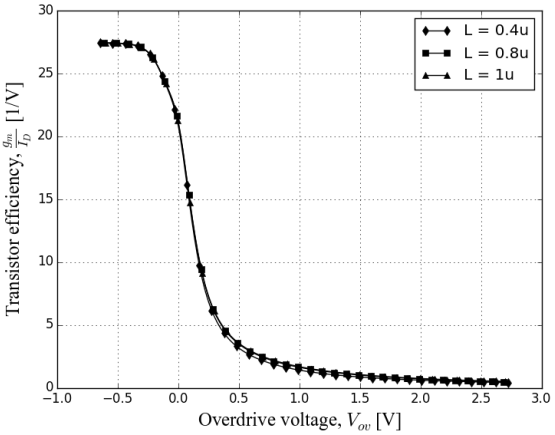


Figure 3. Transistor efficiency versus overdrive voltage

To pick suitable channel length one should observe Fig. 4 which shows small signal gain, a_{v0}, versus drain-source voltage of the MOS device in target technology for different channel lengths. It is notable that shorter channel lengths give relatively constant a_{v0} in wide dynamic range. On the other hand a_{v0} reduces significantly as length decrees. It is obvious that there is a tradeoff between dynamic range and gain.

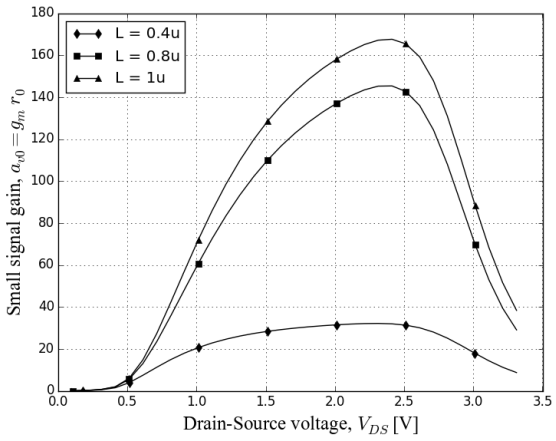


Figure 4. Samll signal gain versus drain-source voltage

If (1) is heavily approximated assuming equal transconductances/resistances, A_0 reduces to $(g_m r_0)^2 = a_{v0}^2$. Picking the L= 0.4μm gives a_{v0} no lower than 20 times in reasonably high dynamic range i.e. 1-3V. Therefore the total gain would be $A_0 \approx 400$ or roughly 52dB. After this value is adopted as good enough the following should be appreciated. Firstly, (1) is heavily approximated and secondly short channel effect is always present. Therefore to mitigate short channel effects, and to ensure gain higher than 50dBs, L= 0.8μm is adopted.

It is also estimated, again using SPICE, that there is a 3.75 μA drain current per 1μm of channel width for chosen channel length in target technology. Accordingly, for 300μA current the minimal width of amplifying devices is W= 80 μm resulting with width-length ratio of 100. This value sets the initial dimensions and the dimensions of all other transistors are drawn based on it. Eventually, final dimensions end up being larger in order to fully meet requirements given in Table I.

B. Common Mode Feedback

CMFB circuitry in Fig. 1 consists of transistors MC0-MC6. Circuit is realized as differential structure which serves as detector of output common-mode voltage change. Ideally circuit should stay inactive in presence of differential output signal. In combination with core transistor M0 negative feedback is formed.

Circuit operation can be qualitatively explained as follows. Suppose there is a positive increment in output common-mode voltage i.e. gate potential of transistors MC1 and MC4 is increased. This will cause reduction of MC1/4 drain currents. Consequently currents of transistors MC2 and MC3 will increase since constant current is forced through MC5/6 and gate of MC2/3 is tied to fixed potential. To ensure maximum output swing, this potential is set to $V_{DD}/2$ by appropriate sizing of MR0-MR3. Because sum of MC2 and MC3 currents flows into MC0 its gate-source voltage, will increase as well. Since core transistor M0 practically represents common-source stage in this signal path, the drain voltage of M0 will decrease. Finally this voltage drop propagates to the output through common-gate structure M1,2/3,4 opposing the initial common-mode voltage increase. Similarly, for differential change at the OTA output, current through MC0 remains unchanged hence, ideally, there is no feedback reaction. Amount of negative reaction is trimmed by MC0/M0 dimensions ratio. To avoid potential unwanted positive feedback reaction through MC5/6, M7/8/5/6 path, MC5/6 and M7/8 transistors are biased from different points.

It is obvious that CMFB loads OTA core degrading dynamic characteristics. Therefore dimensions of CMFB transistors are kept as small as possible. As OTA is to be built into SC circuit, it is likely that this part of the design will be replaced with active CMFB SC network [7]. This way power consumption will be significantly reduced.

C. Bias

Bias circuit is composed of transistors denoted as MB0-MB12 in Fig. 1. Reference current is generated using supply independent, self biased, V_{TH} reference. This reference uses the fact that sensitivity of the active device voltage to the power supply change is always less than unity. This is governed by square root relation between transistors overdrive voltage and drain current. When circuitry is arranged to generate current of active device by its own overdrive voltage, it results with independent voltage reference for power supply. In this case it is done with transistors MB9-MB12 and resistor, R_0. Practically current generated at the gate of MB10 transistor is mirrored back into its drain current thorough MB11/12 current mirror. Consequently it is important to determine appropriate value of R_0 resistance.

From one side current, I_0, in MB9/11 branch is limited by resistor R_0. On the other hand the very same current sets MB10 overdrive voltage. Therefore equation (3) holds. All values in (3) are referred to transistor MB10.

$$I_0 R_0 = V_{TH} + \sqrt{\frac{2I_0}{k'(W/L)}}. \quad (3)$$

Here k' is intrinsic MOSFET transconductance i.e. mobility and gate oxide capacitance pre unity area product, $\mu_0 C'_{ox}$. Solving (3) by R_0, and appreciating relation $g_m^2 = 2k'(W/L)I_0$ (4) arises.

$$R_0 = \frac{V_{TH}}{I_0} + \frac{2}{g_m}. \quad (4)$$

Choosing I_0= 600μA (tail source M0), g_m = 6mS (assuming relatively constant overdrive voltage of M0/MB10) and

knowing that V_{TH} for NMOS device in target technology is about 0.78V, it comes that the value for R_0 is 1.63kΩ. This value is reduced to 1.2kΩ trading power consumption for better dynamics.

Since the reference voltage is self-biased there is a need for start-up circuit to prevent zero current state. Start up circuit is designed with transistors MS0-MS2. If there is a zero current in the circuit the voltage at R_0 is low. This low state feeds the MS0/1 inverter which turns on MS2 and providing the low potential at the gates of PMOS MB11/12. This condition opens the path for the current to flow from power supply towards R_0. Consequently, voltage at the MB10 gate increases. Inverter triggers once again turning the MS2 off. It is important to emphasis that size of the MS0 should be much greater then size of MS1. This way the overdrive voltage of MS0 is quite small, allowing inverter to trigger with lower voltage then $V_{DD}/2$. This ensures that inverter turns of MS2 which normally should be cutoff.

The reset of the bias circuitry (MB0-MB8) serves to distribute generated reference to appropriate points. Transistors MB3-MB5 form high swing cascode current mirror biased with MB6/7 Sooch structure [8]. Transistors MB0/1 are used in similar manner to bias M3/4.

III. SIMULATION RESULTS

Circuit's behavior is examined through set of various simulations in SPICE. Results at room temperature are summarized in Table III.

TABLE III. SIMULATED OTA CHARACTERISTICS

Param.	Description	Condition	Value
A_0	DC, open loop gain	open loop/closed loop[a]	57.6 dB
Φ_M	Phase margin	open loop/closed loop	83 °
f_{GBW}	Gain-bandwidth	open loop	140 MHz
		closed loop	126 MHz
SLR	Slew rate	closed loop, excitation: pulse, ±(*ICMR*/2) V, 100kHz	190 V/μs
t_S	Settling time		16.6 ns
V_{OMAX}	Maximum output swing	closed loop, excitation: sine, ± 3.3 V, 1MHz	± 1.83 V
PSRR	Power supply rejection ratio	open loop, from V_{DD}	215 dB
		open loop, from V_{SS}	218 dB
CMRR	Common-mode rejection ratio	open loop, from V_{CM}	240 dB
ICMR	In. common-mode range	open loop	4 mV
		closed loop	2.54 V
OCMR	Out. common-mode range	open loop	1.92 V
		closed loop	2.74 V

a. Unity gain feedback configuration

As can be seen from Table III, target design requirements concerning open loop gain, gain-bandwidth and slew rate are met. It can be also noted that circuit is slightly overdesigned. This is to leave some margin for PVT (Process, Voltage, Temperature) variations and noise which will inevitable arise at layout/physical level. Good circuit dynamics are paid with burning extra power. Total power of the circuit is quite high and it is estimated to 9.83mW. Again, this value can be significantly reduced by changing the CMFB circuitry as explained in section II. Since fully differential power-supply and common-mode rejection ratios are quite high as expected. Usage of high swing cascodes bias resulted with satisfactory output swing.

Even open loop analysis confirms stability it is of curtail importance to check circuit's closed loop behavior. This is done by using famous Middlebrook method, where instead open loop, total loop gain is examined [9]. Results are graphically presented in Fig 5.

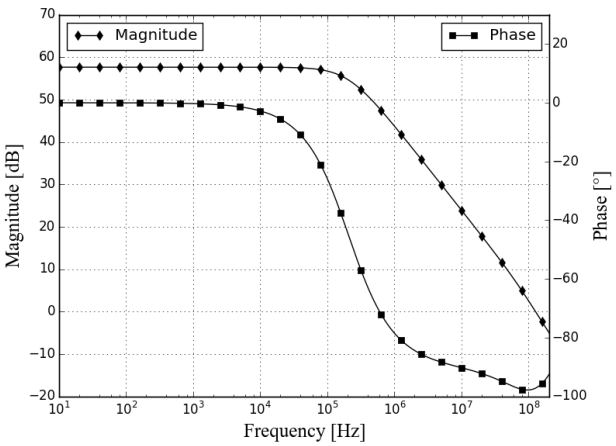


Figure 5. Total loop gain (Magnitude and Phase) versus frequency

This method is considered to be the most trustable when examining stability of feedback systems. It is also favorable because there is no need to break feedback loop hence bias points are not deteriorated. Usually leading CAD vendors, implement this method into its simulation software (e.g. Cadence® Spectre, `iprobe` component in conjunction with `stb` simulation directive). Nevertheless, diving into the [9] one can build its own SPICE deck for implementing the method.

IV. CONCLUSION

This paper presents one design example of OTA circuit considering CMOS 350nm technology process. Designed circuit is to be integral pat of ΔΣ ADC. Adopted architecture is discussed with emphasis on individual sub-blocks namely: Core, Common-Mode Feedback and Bias.

Design procedure of each sub-block is given, as well. For this purpose a set of useful curves is extracted using SPICE giving the insight into MOS device behavior in target technology process. Important design tradeoffs are drawn based on those curves. Transistor level simulation results are presented and discussed. Based on these results one can conclude that circuit meets severe dynamic requirements while preserving stability. Consequently power consumption is increased hence further work will be focused on optimizing design in this direction.

ACKNOWLEDGMENT

This research was funded by The Ministry of Education, Science and Technological Development of Republic of Serbia under contract No. TR32004.

REFERENCES

[1] D. Mirkovic, P. Petkovic, "Analog front end for multichannel Delta-Sigma ADC," in Proc. IX International Symposium of Industrial Electronics, INDEL, Banja Luka, 2012, pp. 55–59

[2] D. Mirkovic, P. Petkovic, "Design Automation of ΔΣ Switched Capacitor Modulators Using Spice and MATLAB," Serb. J. of Electronics, vol. 11, no. 1, pp. 47–59, February 2014. doi: 10.2298/SJEE131129004M

[3] L. Tianwang, Y. Bo, and J. Jinguang, "A Novel Fully Differential Telescopic Operational Transconductance Amplifier," J. of Semiconductors, vol. 30, no. 8, p. 085002, 2009. [Online]. Available: http://stacks.iop.org/1674-4926/30/i=8/a=085002

[4] S. V. Nageshwarrao, D. Chalam and V. M. Rao, "Gain Boosted Telescopic OTA With 110db Gain and 1.8ghz. UGF," Int. J. of Electronics Engineering Research, vol. 2, no. 2, pp. 159–166, 2010.

[5] B. Razavi, "Operational Amplifiers," in Design of Analog CMOS Integrated Circuits, New York, NY: McGraw-Hill, 2001, pp. 291-340

[6] HSPICE® Reference Manual: MOSFET Models (MOSFET Output Templates), Ver. A-2008.03, Synopsys Inc., USA, 2008, pp. 19–31

[7] Choksi, O., Carley, L.R., "Analysis of switched-capacitor common-mode feedback circuit," Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on , vol.50, no.12, pp. 906–917, December 2003. doi: 10.1109/TCSII.2003.820253,

[8] B. Minch, "A low-voltage MOS cascode bias circuit for all current levels," Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on, vol.3, pp.619–622, May 2002
doi: 10.1109/ISCAS.2002.1010300

[9] D. Middlebrook, "Measurement of Loop Gain in Feedback Systems," Intl. J. of Electronics, vol. 38, no. 4, pp. 485–512, April. 1975.

The Design of MCU's Communication Interface

Borisav Jovanović, Dejan Mirković and Milunka Damnjanović
University of Niš, The Faculty of Electronic Engineering,
Niš, Serbia
{borisav.jovanovic, dejan.mirkovic, milunka.damnjanovic}@elfak.ni.ac.rs

***Abstract*— In this paper, the communication between a microcontroller IP block and external Base band microprocessor is examined. The microcontroller is a part of a complex integrated System-on-chip and uses standard 8051 instruction set. The paper describes the operation of embedded circuits that allow programming, software debugging and communication with external microprocessor. The communication is based on SPI interface.**

Keywords - 8051 microcontroller, SPI interface, communication

I. INTRODUCTION

The 8051 microcontrollers are used in many electronic circuits and systems wherever some signal processing or a process control is required. Depending on the user requirements, many microcontrollers' characteristics have to be chosen, such as clock frequency and set of peripheral units.

There have been different implementations of the 8051 microcontroller [1], [2]. Also, there have been many examples of chips in which the 8051 microcontroller is implemented together with some other IP blocks [3]. In our case, the designed 8051 microcontroller IP block is part of an integrated power meter (IPM) System-on-chip [4]. The designed IPM incorporates all the required functional blocks for 3-phase metering, including a precision energy measurement front-end consisting of Sigma Delta AD converters, digital filters and digital signal processing block; 8051 microcontroller, real-time clock, LCD driver and programmable multi-purpose inputs/outputs. The IC requires a minimum of external components, inherently improving meter reliability, simplifying manufacturing process and providing a fast time-to-market metering solution. [4]

This paper describes one practically implemented system-on-chip, with communication between the integrated 8051 microcontroller block and the external Base band microprocessor. This paper thoroughly explains one efficient method for serial communication that allows microcontroller integrated block programming, software debugging and data transfer implementation.

In the Section 2 the global characteristics of the integrated microcontroller are given. Then, the microcontroller's interconnections to surrounding chips are explained. The serial communication interface which is used for basic control of the microcontroller is given. The operations of embedded circuits are explained that allow programming, software debugging and communication. After, in the Section 3, the microcontroller's implementation is described, which also explains the basic steps of verification phases within the integrated circuit design flow. Then, the suggested chip testing setup is described, which uses the serial communication routines for chip programming and testing.

II. THE MICROCONTROLLER

A. *The microcontroller's description*

The proposed microcontroller (MCU) IP block executes the industry standard 8051 instruction set. The instruction set is complex since it contains exactly 255 different assembler instructions. Moreover, the presence of six different addressing modes classifies this type of microcontroller as Complex instruction set computer (CISC). The microprocessors with complex instruction set are difficult for design, especially when tight design requirements have to be met, such as low power operation or high clock frequency.

Although the MCU has only 8 bits, because of the fact that the instruction set is supported by many software development tools, it is still very popular and widely used [5]. Also, the microcontroller fulfils the design requirements of System-on Chip (SoC) in which the MCU is incorporated, including the speed and benefits of rich instruction set.

The global architecture of proposed microcontroller can be divided into following blocks:

- the MCU core,
- the memory blocks,
- the block for programming and initialization
- the peripheral units.

The MCU core fetches, decodes and executes the instructions. It has the speed of one byte instruction executed in only two clock cycles. The maximum clock frequency is equal to 60MHz. The MCU offers the low power options, since it can operate at several clock frequencies. The MCU integrates the clock divider circuit which can reduce the clock frequency by a factor of 32.

The peripheral units include three 8-bit programmable digital parallel input/output ports, Inter-integrated circuit (I2C), Universal serial receiver/transmitter block (UART) and Serial protocol interface (SPI). The communication modules are used for the communication with external chips such as EEPROM memories and Base band processors. Furthermore, three counter/timer circuits are included.

All memory blocks, specific to 8051 microcontrollers, are physically implemented in the chip, including the 8kB SRAM used for program code storing. The 8kB memory block is volatile and program code is lost when chip is powered down. Instead of using on-chip non-volatile memory block, an

external I2C serial memory is utilized. Namely, every time after the chip is powered up, the I2C memory content is read though I2C communication lines and loaded into 8kB SRAM memory block. Beside program code memory the chip incorporates two more SRAM memory blocks: Dual-port 256B internal RAM (IRAM) and 2kB external RAM (XRAM) blocks.

B. The MCU's interconnections to surrounding blocks

The microcontroller IP block is a part of complex SoC and it is used for autonomous control of other IP blocks which belong to the same chip - the digital filters, digital signal processing (DSP) block and real-time clock (RTC) circuit. Namely, the MCU is connected via local SPI lines (Fig. 1) to on-chip mixed signal blocks, which all have embedded SPI communication sub-modules with distinctive SPI addresses. In this communication the MCU is considered as a master unit, which initiates data transfer with other (controlled) IP blocks. The microcontroller's software simply read the status registers and writes the control registers of IP blocks via SPI.

The microcontroller itself is not fully autonomous. Instead, it is controlled by external Base band processor (Fig.1). Base band processor can reset the MCU or select one of the available options for MCU program code initialization. Namely, the MCU's program code can be loaded either from external I2C EEPROM memory or it can be loaded from Base band microcontroller via external SPI communication data lines (Fig.1). Beside the reset and programming options, the Base band can access the IRAM memory locations and SFR registers when MCU's software verification and debugging is performed. The communication with Base band controller will be further examined in detail.

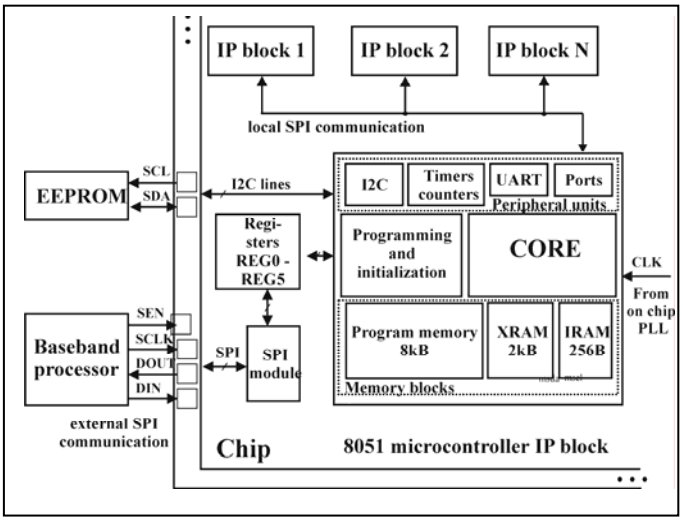


Figure 1. Global diagram of 8051 microcontroller, embedded into SoC

The MCU programming and control options are enabled by SPI module (Fig.1), which represents the main interface in the communication between Base Band processor and MCU. The SPI module utilizes the standard SPI communication protocol and following digital lines: SCLK – for the clock signal; DIN - input data line; DOUT – output data line and SEN – enable signal. The SCLK, DIN and SEN are module's input signals and DOUT is the output signal.

In the external SPI communication the role of MCU is changed compared to the internal chip SPI communication between MCU and other on-chip IP blocks. Now, the Base band processor is a master and the MCU can be considered as a slave. The main challenge during the chip design and verification is related to the Base band SPI clock (signal SCLK in Fig.1) which is not in synchronization with the main MCU clock signal. These two clock signals can have different frequencies and phases. Therefore, the synchronization between two clock domains had to be implemented too.

The SPI module reads and writes the 8-bit registers named with REG0 to REG5. These registers are used for MCU initialization, control and data transfer. For example, the registers REG0 and REG1 are connected to the parallel data input/output ports P0 and P1 which are used for the communication between the Base band and MCU. Namely, two MCU peripheral ports P0 and P1 are 8-bit wide and they are used to expand communication possibility between Base Band processor and MCU. The data at port P0 input is changed by writing data into the REG0 register; the Base band can read the port P1 output by reading the REG1 register.

Through REG2 the Base band controls the MCU's control input pins such as the main reset pin and programming mode selection input pins. Over REG3, the MCU provides the Base band its status signals. The REG4 and REG5 are used for data transfer between the MCU and Baseband which is needed during MCU programming, verification and testing procedures. The names, corresponding register addresses and short descriptions are given in the Table 1.

TABLE I. SHORT DESCRIPTION OF REGISTERS REG0-REG5

mSPI's register	register address	Description
REG0	0000000	Writes to port P0 inputs
REG1	0000001	Reads from port P1 outputs
REG2	0000010	Controls MCU input control pins
REG3	0000011	Reads MCU status signals
REG4	0000100	The data byte to be written into MCU
REG5	0000101	Reads data byte from MCU

C. The SPI registers used for data transfer between the MCU and Base band processor

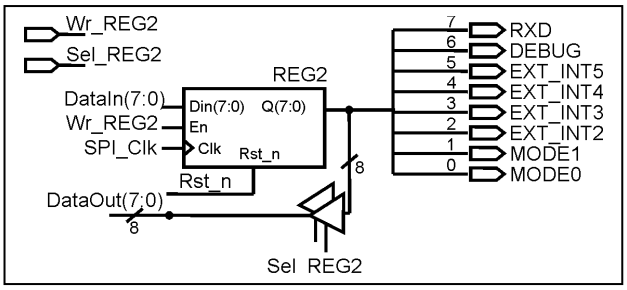


Figure 2. The register REG2 content controlling MCU input pins

The brief description of most important bits of registers REG2 and REG3 which are necessary for 8051 MCU communication with the Base band processor is given as follows.

The register REG2 (Fig. 2) is connected to MCU's input pins. The bits MODE1 and MODE0 of REG2 (bit positions 1 and 0) are used to control the chip boot-up. The detailed descriptions of chip programming modes will be given in the next Section.

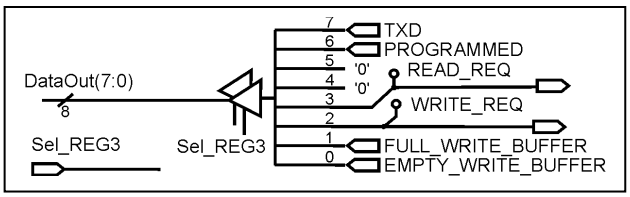


Figure 3. The register REG3 reading the status signals

The REG3 (Fig. 3) is read-only register. The bit 3 – the READ_REQ indicates that 8-bit data, available in the register REG5, is ready for transfer from MCU to the Base band. The bit 2 - WRITE_REQ is status signal holding the information that a new data byte in REG4 is waiting to be read by MCU. The FULL_WRITE_BUFF at bit position 1 indicates that 32-byte microcontroller's input FIFO buffer is full, so, Base band processor has to wait. The signal at bit position 0 - the EMPTY_WRITE_BUFF tells that 32-byte input FIFO is empty.

D. The read and write operations and the synchronization between the two clock domins

The REG4 write operations are performed during program code transfer from Base band to MCU and MCU's software debugging procedures.

During a write operation a data byte is written into the REG4. The following control signals are used: WRITE_REQ, EMPTY_WRITE_BUFF and FULL_WRITE_BUFF.

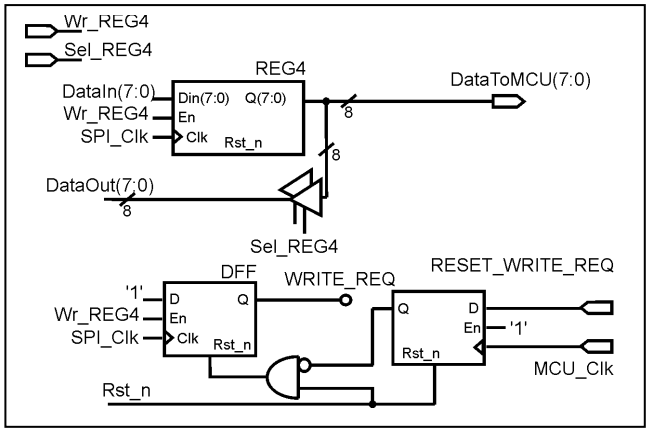


Figure 4. The data transfer to MCU through the register REG4

Since Base band's SPI clock and MCU's clock are not synchronized, the handshaking is implemented by signal WRITE_REQ (Fig. 4). The synchronization between two clock domains is achieved at the flip-flop producing the signal WRITE_REQ. The WRITE_REQ is set when new data byte is written into the register REG4. Namely, the WRITE_REQ is set to logic 1 synchronously with Base band's clock signal (the signal SPI_Clk given in Fig.4). After the MCU takes the data byte, it automatically resets the WRITE_REQ. The MCU's control logic which operates at clock frequency MCU_Clk, now asynchronously resets the flip-flop producing the signal WRITE_REQ.

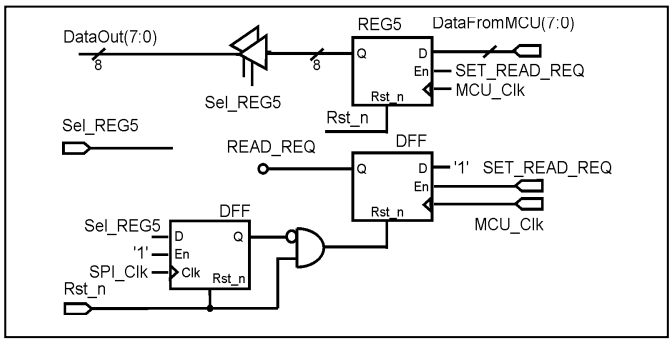


Figure 5. The data transfer to the Base band through the register REG5

The read operations happen during MCU's debugging. Then, the data is read by Base band through register REG5. Since Base band SPI clock and MCU's clock do not have to be synchronized the handshaking is implemented. The circuit implementing the read operation is given in Fig. 5. The synchronization between two clock domains is achieved at flip-flop producing the signal READ_REQ. When data byte is loaded into the REG5, the MCU sets the READ_REQ signal. Note that the signal READ_REQ is set synchronously with main clock signal of the microcontroller (the MCU_Clk given in Fig. 5). When REG5 is read through SPI module, the READ_REQ is automatically reset by SPI interface control logic. The control logic loading the register REG5 operates at Base band's clock frequency (signal SPI_Clk given in Fig. 5) and generates the short pulse which asynchronously resets the flip-flop producing the signal READ_REQ.

Synchronizing between clock domains is accomplished by registering the signals through a flip-flops that are clocked by the source clock domain, thus holding the signal long enough to be detected by the higher frequency clocked destination domain. To avoid issues with meta-stability in the destination domain, a minimum of 2 stages of re-synchronization flip-flops are included in the destination domain.

III. THE IMPLEMENTATION OF THE MICROCONTROLLER

A. Design and verification

The microcontroller was implemented in TSMC 65nm technology [6]. It operates at voltage supply of 1.2V. The following Cadence tools [7] have been used during the chip design:

- RTL Compiler for logical synthesis
- SoC Encounter for implementation.
- NCSim for logical verification

The MCU has the following advantages:

• new architecture provides speed of one 8-bit instruction executed in only two clock cycles

• operating frequency maximum is 60MHz

• low power operation; at frequency of 60MHz the power consumption is only 3mW.

The logical verification procedures focus on the MCU programming options which are achieved through SPI interface.

The MCU logical verification process begins by writing the C programs, which are compiled by Keil [8] or SDCC [9] into 8051 .hex file. After, the special C program converts the .hex file into VHDL code. This code provides the program memory content, which is instantiated into the main test bench.

The test bench program simulates the MCU's surroundings. The same test bench includes the instance of MCU which is being verified, the SPI communication module of Base band processor and external I2C EEPROM memory. The VHDL description of MCU also includes the VHDL code of internal SPI module. The control logic of test bench simulates the operation of Base band processor and describes the program code transfer to the MCU and EEPROM memory. The data bytes are sent first via SPI interface and after to the EEPROM.

During the logical verification process, the SPI module of Base band reads the program memory content and byte-after-byte transfers it to the SPI module of MCU. The logical verification results, obtained by simulation waveforms, present the SFR register's content. The MCU operation is simply verified by comparing the expected and obtained results.

The absence of synchronization which exists between the Base band SPI clock signal and main MCU's clock signal made the design of communication modules more difficult. To overcome this problem, additional circuits for synchronization are added to the communication blocks which synchronize two clock domains. The verification process considered different combinations of clock signal frequencies. The MCU's clock frequency maximum is 60MHz and Base band SPI communication is 100MHz. The MCU's clock frequency value is decreased by clock divider circuit. The SPI clock of Base band is changed from 100MHz down to 1MHz.

Two options are supported here, one using external (off chip) EEPROM and another without external EEPROM.

Option A: Using external EEPROM

- Base band sets the MCU programming mode MODE(1:0)="01"
- Base band uploads 8KB into SRAM Program memory.
- While receiving 8KB, MCU flushes Program memory into the EEPROM.
- When transfer is finished, the MCU starts executing the code.

With this option, also, the following is enabled:

- If Baseband sets MCU programming mode MODE(1:0)="11", the EEPROM content is read and uploaded into 8kB SRAM. The SPI is not used for program code transfer
- When transfer is finished, the MCU starts executing the code.

Option B: No external EEPROM

- Base band processor sets MODE(1:0)="10"
- Base band processor uploads 8KB into Program memory via SPI registers.
- After receiving 8KB, the MCU starts executing the code.

Option A loads the on-chip SRAM and external EEPROM. The Base band controller sets the bits Mode(1:0)="01" of REG2. Then, the bytes are sent in 32B packets over SPI. After the chip has been programmed, the MCU sets the status bit Programmed, which is located in REG3.

The on-chip program code memory is loaded from external EEPROM when Base band sets the REG2 bits Mode(1:0)="11". After chip is reset, the program code bytes are automatically read from I2C EEPROM and loaded into the SRAM. The control logic sent the control codes to the MCU which performs the program code loading procedure via I2C lines. The programming is finished when status bit Programmed is set.

B. The 8051 MCU testing setup

The general test setup is depicted in Fig. 6. It is comprised of control application, installed on computer and Printed Circuit Board (PCB), specially designed for testing purposes.

The computer is connected to PCB over USB port. PCB includes also the Baseband processor and I2C EEPROM, dedicated to MCU program code storage (Fig. 6).

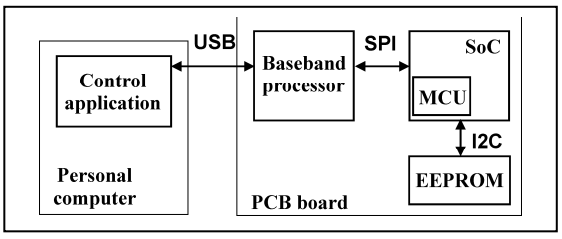


Figure 6. Global schematic of MCU testing setup

The control application implemented by [10], [11] sets the parameters of all subsystems within the chip. The part of control application is dedicated to 8051 MCU testing. It enables:

- The MCU reset
- HEX file loading, dedicated to MCU program code
- Selecting the MCU programming options A and B, as described below.

- Checking the MCU operation using Debug mode
- providing the IRAM memory and SFR registers content, which is valuable for MCU testing, and also, for MCU program code debugging.

The Base band firmware includes the software routines to read and write the MCU registers. These functions directly control the SPI communication lines between Base band processor and MCU. The algorithms which is used for MCU programming is described as follows.

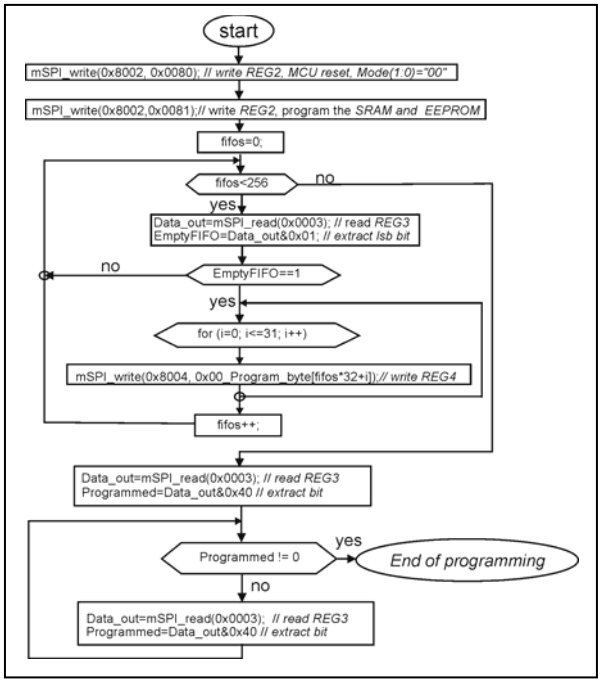


Figure 7. The MCU programming operations

At the beginning, the MCU is reset by writing the command into register REG2 (bits MODE(1:0)=”00”). After, the bits are changed with MODE(1:0)=”01” or MODE(1:0)=”10”, dependent of the programming mode wanted. When programming mode 1 is chosen, the program code is written into both SRAM and EEPROM. When mode 2 is selected, the code enters only the SRAM. The program code is sent in data packets consisting of 32 bytes. The EmptyFIFO status signal gives the information when MCU’s FIFO is ready to take incoming bytes. The EmptyFIFO is checked by reading the REG3 register. The bytes are written into REG4. The operation is repeated 256 times to load 8kB MCU SRAM. At the end, the bit Programmed is checked.

IV. CONCLUSION

The 8051 MCU IP block is a part of a complex System-on chip and it is verified by systematic and thorough simulations. The results are confirmed after the layout was designed.

The microcontroller is connected to external Base band controller via SPI interface. In the serial communication, the MCU is a slave and the Baseband is a master. The MCU offers several programming options. In one of them, the program code memory is loaded from Baseband via SPI into both on-chip SRAM memory and external I2C EEPROM memory. Moreover, the Base band has control over MCU. The serial communication is particularly verified by simulations, the methods are described in paper.

The MCU’s main clock signal does not have the same frequency as SPI clock signal, which is produced by Base band controller. Therefore, the synchronization between different clock domains is done for reliable data transfer between the MCU and Base band.

ACKNOWLEDGMENT

This paper was funded by the Ministry of Education, Science and Technological Development, the project number TR32004, entitled with "Advanced technologies for measurement, control, and communications in the electricity grid."

REFERENCES

[1] A. J. Martin, M. Nystrom, K. Papadantonakis, P. I. Penzes, P. Prakash, C. G. Wong, J. Chang, K. S. Ko, B. Lee, E. Ou, J. Pugh, E. V. Talvala, J. T. Tong and A. Tura, “The Lutonium: A Sub-Nanojoule Asynchronous 8051 Microcontroller”, in Proc. of ASYNC, 2003, pp. 14–23.

[2] J. H. Lee, Y. H. Kimand K. R. Cho, “A Low-Power Implementation of Asynchronous 8051 Employing Adaptive Pipeline Structure”, IEEE Transactions on Circuits and Systems II: Express Briefs, Vol. 55 , Issue 7, 2008, pp. 673 – 677

[3] K. L. Chang and B. H.Gwee, “A low-energy low-voltage asynchronous 8051 microcontroller cores,” in Proc. ISCAS, 2006, pp. 3181–3184

[4] B. Jovanović, M. Zwolinski, M. Damnjanović, “Low power digital design in Integrated Power Meter IC,” in Proc. of the Small Systems Simulation Symposium 2010, Niš, Serbia

[5] K. Arnold, Embedded controller hardware design, LLH Technology Publishing, Eagle Rock, USA, 2000.

[6] TSMC 65nm standard cell library, http://www.europractice-ic.com/technologies_TSMC.php?tech_id=

[7] Cadence, EDA software and verification tools http://www.cadence.com/us/pages/default.aspx

[8] KEIL C compiler and development tool for 8051 microcontrollers, http://www.keil.com/c51/ devproc.asp

[9] Open source SDCC C compiler for 8051 microcontrollers http://sdcc.sourceforge.net/.

[10] Code::Blocks IDE, http://www.codeblocks.org/

[11] GCC, the GNU Compiler Collection http://gcc.gnu.org

Session TO-4
ELECTRICAL MACHINES AND DRIVES

Experimental Identification of the Mechanical Parameters of an Induction Motor Drive

Dejan D. Reljić, Dejan G. Jerkan
Department for Power, Electronic and Telecommunication Engineering
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
reljic@uns.ac.rs

Abstract **– In order to obtain fast dynamic response performance of an induction motor drive, the identification of mechanical parameters such as the drive inertia and the coefficients of friction, with a good accuracy, is highly desirable. They are essential for the design of the high-performance induction motor drive speed, as well as position controllers and speed observers, since a drive response is influenced not only by load disturbances but also by these mechanical parameters. Moreover, they are of great importance for the accurate dynamic modeling and simulation of various high-performance induction motor control strategies. In this paper an experimental off-line method for the mechanical parameters identification is presented. The method uses speed-time curve, obtained during the retardation test on the drive, with an appropriate mechanical losses model of the drive, and the mean squared error performance function based on a genetic algorithm (GA) approach, to obtain unknown mechanical parameters of the tested drive. The proposed method is verified by experiments.**

Keywords – genetic algorithm; induction motor; mechanical parameters identification; retardation test

I. INTRODUCTION

Induction motors, as a part of electrical drives, are by far the most widely used rotating electrical machines in many industrial applications [1]. The vast majorities of induction motors are usually used in a low-performance drives [2], such as pump and fan applications. However, induction motors are also used in a high-performance drives [2], such as machine tools, extruders, propulsion systems for the electrical vehicles, etc. In order to obtain fast dynamic response performance of an induction motor drive, the knowledge of electrical parameters is not enough [3]. The mechanical parameters are also of great importance.

The identification of the mechanical parameters, such as the drive inertia and the coefficients of friction, with good accuracy, is highly desirable. These parameters are crucial for the design of high-performance speed and position controllers. This is because a drive response is influenced not only by load disturbances but also by drive inertia and friction [4]. On the other hand, the identification of the mechanical parameters is also important for the accurate dynamic modeling and testing of various high-performance motor control strategies.

Many researchers have done a great number of researches on parameters identification of the induction motor. However, most of them are mainly focused on electrical parameters determination, while only a few of them are related to the mechanical parameters identification. In [4] an observer-based auto-tuning scheme with two adaptive controllers is used to separately adjust the drive inertia and friction torque to their correct value. In [5] a genetic algorithm (GA) is employed with the aim of determining the mechanical and electrical parameters of an induction motor. However, the most common and simple method of determining the moment of inertia is by performing a retardation test on the induction motor drive [6]-[11]. The retardation test is also suitable for the viscous friction coefficient identification [10], while the Coulomb friction is mainly neglected [9]. Nevertheless, the Coulomb friction may greatly affect the behavior of high precision drive systems [12]-[13]. In [7] an example of the dry torque determination using a progressive start-up experiment is presented. However, the proposed experiment in [7] is more related to static friction determination, but not the kinetic, i.e. Coulomb friction. The static friction is usually larger than the Coulomb friction [13].

The retardation test method is also used in this paper for the off-line mechanical parameters identification. The estimation of the drive inertia is similar to the procedure described in [6]-[11] with the exception that before the retardation test is done the flux density level in the motor is to be reduced. Thus, the speed-time curve is less influenced by electromagnetic transients. This results in more accurate drive inertia estimation. The Coulomb friction torque is estimated to be equal to the torque developed by the motor at the drive speed where the absolute value of the deceleration of the retardation curve begins to increase. Finally, a GA approach is employed to fit the proposed mechanical losses model of the drive (with previously estimated drive inertia and the Coulomb friction torque) to deceleration curve obtained by the retardation test. As a result, the drive inertia, the Coulomb friction torque and the unknown friction coefficients of the drive are obtained.

II. ESTIMATION OF THE DRIVE INERTIA

The moment of inertia is a measure of the body's resistance against a change of its rotational motion. It depends on the distribution of its mass relative to the axis of rotation [14]. The moment of inertia of a machine can be determined by an analytical calculation and by an experiment. Analytical computation is suitable during the design stage of the machine [11] but in the absence of the drive geometry details and used materials, it is ineffective for the whole drive inertia determination. Therefore, the drive inertia is usually determined by a direct experiment. The most common method of determining the drive inertia is by performing a test known as the retardation test on the drive [6]. The same method is used in this paper. However, in order to obtain a more accurate value of the drive inertia, the method is slightly modified in comparison to the ones described in [6]-[11].

The drive is run up to the speed above rated at no load. The supply voltage of the motor should be reduced as much as possible in order to reduce the flux density level in the motor to avoid electromagnetic transients (braking effect) after switching the power off, particularly for delta-connected windings. Then the motor's supply voltage is switched off. The drive slows down and comes to rest while speed-time curve is recorded. At any rotational speed ω, power P consumed in overcoming the mechanical losses is given by [6]-[11]:

$$P = \frac{d}{dt}\left(\frac{1}{2}\cdot J\cdot \omega^2\right) = J\cdot\omega\cdot\frac{d\omega}{dt}, \tag{1}$$

where J is the drive inertia. From the retardation test the slope of the deceleration curve $d\omega/dt$ is obtained at the speed ω_1 (usually ω_1 is the rated speed). Then the drive is reconnected to the power supply and run at the speed ω_1 by controlling the motor's supply voltage, and the electrical power input P_{01} to the motor is measured. As an approximation, the mechanical power P_1 at the speed ω_1 may be taken as:

$$P_1 = \left(P_{01} - P_{JS1} - P_{Fe1}\right)\cdot\left(1 - s_1\right), \tag{2}$$

where P_{JS1} is the stator Joule losses obtained from the measurement of the stator current and resistance, P_{Fe1} denotes the core losses, and s_1 is the slip at the speed ω_1 (the synchronous speed is known). The stray load losses in (2) are neglected, while the core losses, i.e. the stator core losses can be obtained by separation of no-load losses of the motor according to [15]. Now, using (1) and (2), the drive inertia can be calculated as:

$$J = \frac{P_1}{\omega_1\cdot\left|\left(\frac{d\omega}{dt}\right)_{\omega=\omega_1}\right|}. \tag{3}$$

Main problem in this method is that the mechanical losses cannot be estimated accurately because the rotor core losses in (2) are neglected. Besides, in the case of motors operated by PWM inverters, the core losses and the mechanical losses cannot be accurately separated because additional core losses appear due to higher order harmonics in the supply voltage. The core losses due to PWM harmonic voltages can be significant compared to a sinusoidal supply. These additional core loss components depend on the modulation index and on the inverter switching frequency [16]-[17]. The results in [16]-[17] show that core losses increase with decreasing modulation index, while increasing the switching frequency (above 5 kHz) has no significant effect on additional core losses, i.e. they remain almost constant independently of switching frequency. For those reasons it is rather difficult to achieve a good accuracy of the mechanical losses determination described by (2). However, the difficulties associated with additional core losses can be overcome if the motor in this test is supplied by the sinusoidal voltage instead by PWM inverter. This will significantly improve the accuracy of the drive inertia determination.

On the other hand, in order to provide higher level of accuracy, the drive inertia obtained from (3) will be tuned to more accurate value using the mechanical losses model of the drive and the GA approach. This will be discussed in the next sections.

III. THE MECHANICAL LOSSES MODEL

The mechanical losses in an induction motor drive are a result of mechanical losses in the induction motor and power losses in various parts of the transmission systems. Before the mechanical losses model of the tested drive system is established, the paper gives a short overview of each loss component.

A. Modeling the Induction Motor's Mechanical Losses

Mechanical losses in the induction motor occur due to friction in motor bearings and air movement in the motor. They are usually referred as bearings friction losses and windage and ventilation losses. The amount of these losses can be considerably large during a high-speed motor operation, and for totally enclosed fan-cooled motors.

The bearings are used in electrical machines to support the rotor and to keep the rotor centered in the stator [15], [18]-[19]. There are numerous different types of bearings that are used for electrical machines. The right selection depends on design requirements and operating conditions of the application [19]. In small machines ball bearings are commonly used, while in larger or heavily loaded machines roller bearings are typically used [19]. Regardless the type of bearings, they always contribute to the machine's overall friction power losses. The total friction in a bearing is the result of the rolling and sliding friction in the contact areas, between the rolling elements and raceways, between the rolling elements and a cage, and between the rolling elements and other guiding surfaces [20]. Friction is also generated by the lubricant drag and contact seals, if applicable [20]. The amount of friction depends on the loads, bearing type and size, operating speed, and properties and quantity of the lubricant [20]. The theoretical background of each of these sources of friction, including the mathematical formulation, is briefly explained in [20]-[21].

Friction is an extremely complex nonlinear phenomenon. Different models for the estimation of the bearings friction torque are available in literature. Five different models are described in [21] including the more advanced SKF bearing friction model [20]. Each of the models from [21] estimates the friction torque by grouping the individual sources together and quantifying them by empirical formulas. However, in this paper, the rotary Coulomb friction model [21] is adopted due to its simplicity. According to this model, the bearings friction torque T_b can be approximated as [20]-[21]:

$$T_b = 0.5 \cdot \mu \cdot F \cdot d \cdot \mathrm{sgn}(\omega), \tag{4}$$

where μ is the kinetic friction coefficient, F is the resultant bearing load, and d is the bearing bore diameter. Equation (4) is valid under certain conditions, such as: equivalent bearing load is 10% of the load rating, good lubrication, normal operating conditions (speed range 30-70% of the kinematically permissible speed) and no additional stress [21]. The kinetic friction coefficient μ is an experimentally determined constant. Its value depends on the bearing type and size, the bearing load, the load angle, and the rotational speed [21]. In [20] some typical values of the kinetic friction coefficient can be found.

It should be noted that at low-speed operation there is a considerable amount of mixed friction in bearings because rolling contacts are not yet separated by a lubricant film [20]. This is not included in (4). Low-speed drive operation is not considered in this paper.

A more complete evaluation of the bearings friction, including the influence of each contact areas and friction components, can be done using more advanced friction models and specialized software tools, but the manufacturer's specifications on the bearings and the lubricant are required. Usually, these data are difficult to be provided. For that reason, the model defined by (4) is more convenient because it only requires the knowledge of a few constants.

However, the bearings friction torque T_b can be estimated in a different way than it is proposed by (4). In [7] an example of the dry torque determination using a progressive start-up test experiment is presented. Nevertheless, the proposed experiment in [7] is more related to the static friction determination, but not the kinetic friction. This paper suggests that the bearings friction torque can be approximated as the torque developed by the unloaded induction motor at the speed ω_b at which the absolute value of the deceleration of the retardation curve begins to increase. Here it is assumed that from the speed ω_b the bearing friction torque starts to increase as speed decreases since a lubricant film is lost. Accordingly, the stator supply voltage at the rated frequency is reduced until the motor speed decreases to the value that is approximately equal to ω_b. Then, the bearings friction torque can be approximated as follows:

$$T_b = \frac{(P_b - P_{JSb}) \cdot (1 - s_b)}{\omega_b} = p \cdot \frac{P_b - P_{JSb}}{\omega_s}, \tag{5}$$

where P_b is the power absorbed by the motor when the motor speed is approximately equal to ω_b, P_{JSb} denotes the stator Joule losses obtained from the measurement of the stator current and resistance, s_b is the slip at speed ω_b, p is the number of pole pairs, and ω_s is the synchronous speed at the rated frequency. Note that the core losses in (5) are neglected since the stator voltage is small. To determine the speed ω_b speed-time curve, obtained during the retardation test, is needed.

Windage losses are friction losses associated with the friction between the rotor surface and the surrounding fluid, usually air within the air-gap [22]. This term is generally used to denote the loss due to fluid drag on a rotating body [23]. Windage losses are influenced by the rotor peripheral velocity, stator and rotor geometry and their surfaces roughness, the properties of the rotor's surrounding air and the length of the air-gap (the smaller the air-gap, the bigger the windage losses) [24]. Under the assumption that the rotor can be modeled as a rotating cylinder in an enclosure, the windage torque T_{w1} can be expressed as follows [22]:

$$T_{w1} = \frac{1}{32} \cdot k \cdot C_{M1} \cdot \pi \cdot \rho \cdot \omega^2 \cdot D_r^4 \cdot l_r, \tag{6}$$

where k is a roughness coefficient, C_{M1} is the torque coefficient, ρ is the density of air, D_r is the rotor diameter, and l_r is the rotor length. The torque coefficient C_{M1} depends on the Couette Reynolds number and it is determined by measurements [22].

The end surfaces of the rotor also create windage losses. Under the assumption that these parts can be modeled as discs rotating in free space, the windage torque T_{w2} can be expressed as follows [22]:

$$T_{w2} = \frac{1}{64} \cdot C_{M2} \cdot \rho \cdot \omega^2 \cdot (D_r^5 - D_{ri}^5), \tag{7}$$

where D_r is the outer diameter of the rotor, D_{ri} is the shaft diameter, and C_{M2} is the torque coefficient which, in contrast to (6), depends on the Reynolds number. However, the Couette Reynolds number and Reynolds number depend on the speed, which means that torque coefficients C_{M1} and C_{M2} are speed dependent.

The windage torque caused by the rotating parts of the machine is now the sum of (6) and (7):

$$T_w = T_{w1} + T_{w2}. \tag{8}$$

It follows from (6)-(8) that the windage torque increases with the square of the rotor speed, i.e. the windage power losses increases with the cubic of the rotor speed. This is of particular interest for high-speed motors.

Besides bearings friction losses and windage losses, ventilation losses also belong to mechanical losses. Induction motors are normally provided with a cooling system with the main objective to remove the heat generated by power losses in the motor. The cooling system consists of an internal shaft

mounted fan, or it can be driven by another motor (usually in variable-speed drives) [22]. The blades attached to the rotor short-circuiting ring have also a ventilating effect [25]. Windage losses and ventilation losses occur together. In [22] an experimental equation for the sum of windage and ventilation losses is given. These losses, expressed via the windage and ventilation torque T_{wv}, are:

$$T_{wv} = \frac{1}{4} \cdot k_{\rho} \cdot D_r^3 \cdot (l_r + 0.6 \cdot \tau_p) \cdot \omega \,, \tag{9}$$

where k_{ρ} is an experimental factor (typical values can be found in [22]), D_r is the rotor diameter, l_r is the rotor length and τ_p is the pole pitch. Equation (9) is valid for normal-speed motors, while for high-speed motors (7) and (8) have to be used.

B. Modeling the Drive's Mechanical Losses

The induction motor drive which is considered in this paper is shown in Fig. 1. The system consists of the three-phase cage induction motor (supplied by the power inverter) coupled to the wound-field synchronous machine. Besides, there is a flywheel system coupled directly to the induction motor shaft.

The drive from Fig. 1 is used to provide a variable frequency sinusoidal voltage source used for AC machines tests, as well as to test wound-field synchronous machine and flywheel energy storage system in renewable and distributed energy applications.

The total mechanical losses of the induction motor drive from Fig. 1 are a superposition of the mechanical losses from the both machines, and from flywheel and mechanical coupling. On the basis of the results obtained for the induction motor's mechanical losses formulation, the mechanical losses model for the synchronous machine will be established. These losses results from bearings and brush friction losses, and windage and ventilation losses.

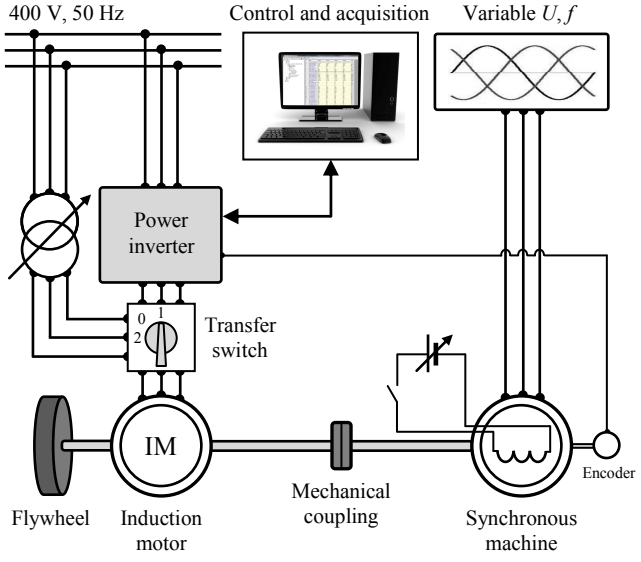


Figure 1. Induction motor drive system configuration.

Considering that a similar bearings type are used in both machines, and that they operate under similar working conditions, the bearings friction torque in the synchronous machine can be expressed as in (4). Brush friction usually has the Coulomb sliding friction characteristic. The torque to overcome the brush friction depends on the brush pressure, brush grade, contact area, and the condition of the slip rings [26]. Brush friction torque is a constant. It has been suggested in [26] to perform two retardation tests, one with brushes lifted and one without brushes lifted, to experimentally obtain brush friction torque. However, such a test has not been performed in this paper since the brush friction torque will be incorporated in the bearings friction torque of the tested drive. Besides, the sum of the bearings and brush friction torques can be approximately estimated as it is previously proposed for the bearings friction torque estimation for the induction motor.

The wound-field synchronous machine from Fig. 1 has a salient-pole rotor with a squirrel-cage winding distributed over the rotor. For that reason, this machine can be considered to have cylindrical-rotor with respect to windage and ventilation losses. Knowing the machine geometry, windage and ventilation torque can be calculated from (9).

The torque T_f of a thin flywheel disk caused by air resistance is described by [27]:

$$T_f = C_m \cdot \rho \cdot \omega^2 \cdot D_f^5 \,, \tag{10}$$

where C_m is the torque coefficient associated with the type of air flow (depends on the Reynolds number), ρ is the density of air, and D_f is the is the radius of the flywheel.

The air resistance torque of mechanical coupling losses can also be approximated by (10). However, it should be noted that the mechanical losses due to flywheel and mechanical coupling air resistance to motion are much lower than the windage and ventilation losses of the drive.

In order to obtain a more convenient mechanical losses model of the whole drive from Fig. 1, and to avoid the need for knowledge of the machines geometry, some modifications of the presented models will be carried out. The windage and ventilation losses model of the drive from Fig. 1 can be similarly expressed as (9), while for high-speed (6) and (7) have to be used. The flywheel and mechanical coupling air resistance to rotation can be modeled as (10). Taking (6), (7), (8), (9) and (10) into consideration, it is possible to replace these equations with their simpler and more convenient form. Namely, the paper proposes the corresponding torque T_{ρ} (sum of the mechanical losses due to windage and ventilation, and flywheel and mechanical coupling air resistance to rotation) to be modeled by the following expression:

$$T_{\rho} = b \cdot \omega^{(1+a \cdot \omega)} \,, \tag{11}$$

where a and b are coefficients of the drive. This model is consistent with the previous ones, except that the loss influence factors are expressed in a much more convenient way, through

the coefficients a and b. Substituting $a = 0$ in (11) the model (9) is obtained, while substituting $a \cdot \omega = 1$ (with different b values) in (11) the models (6), (7) and (10) are obtained. The term $a \cdot \omega$ in (11) is speed dependent and it is used as the replacement for coefficients associated with the type of air flow which depend on the Couette Reynolds and Reynolds numbers.

Using the estimated drive inertia J from (3), bearings and brush friction torque T_{bb} from (5), and the speed-time curve recorded during the retardation test, the unknown coefficients from (11) can be obtained on the basis of the Newton's second law of rotational motion for unloaded drive:

$$\frac{d\omega}{dt} = -\frac{T_{bb} + b \cdot \omega^{(1+a \cdot \omega)}}{J}. \tag{12}$$

The deceleration curve, obtained by performing the retardation test on the drive, is fitted to model (12) using the mean squared error performance function based on the GA approach. This will be discussed below.

IV. Genetic Approach to Mechanical Parameters Identification

In this section the model defined by (12) uses a GA approach to identify unknown coefficients a and b after a number of iterations that satisfy the fitness function and all constraints. Moreover, the GA is also used to tune the drive inertia J and the bearing and brush friction torque T_{bb} to more accurate values. This is done because the methods defined by (3) and (5) are influenced by the modeling and measurement errors. A brief introduction to the GA is presented.

A GA was introduced by J. H. Holland [28]. It is a model of machine learning which is based on the natural processes of selection and evolution. The GA uses a stochastic approach which operates on individuals of a population, applying the principle of survival of the fittest which then evolves toward the solution of the selected problem [28]. The basic principles of the GA are quite simple. The first step is the selection of individuals from an initial population. The individuals carry chromosomes which are potential solutions to the selected problem. The algorithm usually selects individuals that have better fitness values. The second step is genetic manipulation of the selected individuals by crossover and mutation methods, including elitist selection. This results in new population which is better than the previous generation. At each next step, the GA uses the current population to create the next generation. The algorithm stops when the stopping criteria are met.

In this paper the main goal of the GA is to find the mechanical parameters of the drive from the model defined by (12). The fitness function is defined in such a way that the fitness values are minimized:

$$fit = mse\left(\frac{\frac{d\omega_m}{dt}(i) - \frac{d\omega}{dt}(i)}{\frac{d\omega_m}{dt}(i)}\right), \tag{13}$$

where *fit* denotes the fitness function, *mse* stands for the mean squared error, $d\omega_m/dt(i)$ are the elements of the vector which is composed of the angular deceleration points obtained by the retardation test on the drive, and $d\omega/dt(i)$ are the vector's elements of the angular deceleration points estimated by (12).

Equation (13) is an optimization problem which has been solved in this paper by the GA solver using Matlab's Optimization Toolbox [29]. The GA is a relatively new optimization technique, but it has been proved as an effective method for the process optimization. Yet, options for the GA solver need to be carefully selected. Otherwise, the GA method would get stuck in local minima. The knowledge about the proper selection of the GA parameters has a rather empirical background and some recommendations can be found in [30]. Table I provides a quick overview of the most important selected parameters and methods provided by Matlab's Optimization Toolbox that has been used in this paper for mechanical parameters identification.

The size of the population depends on the problem complexity. The problem of optimal population size has been studied in literature, but a general rule cannot be applied to every problem. In this paper population size has been tuning until reaching reasonable value of the fitness function (13).

Furthermore, it is particularly important to specify an initial population of the GA. If not, Matlab creates a random initial population using a creation function which can lead to bad performance of the algorithm. The initial values of the drive inertia J and the bearings and brush friction torque T_{bb} can be calculated according to (3) and (5), respectively, while the initial value of the coefficient a can be selected to be zero. The initial value of the coefficient b can be roughly selected according to the mechanical losses data of the motor (drive) obtained by no-load test [15]. This ensures diversity in the GA and provides the GA to converge quickly for finding near-optimal solution to the problem.

The fitness scaling option is Rank, which scales the raw scores based on the rank of each individual.

Selection function selects individuals, i.e. the parents from the population, based on their scaled values from the fitness scaling function. In this GA, the Stochastic uniform function has been selected to perform the selection.

TABLE I. Parameterization of the GA

Options	Values, methods
Population size	50
Fitness scaling	Rank
Selection function	Stochastic uniform
Reproduction	Elite count: 2; Crossover fraction: 0.8
Mutation function	Adaptive feasible
Crossover function	Heuristic, ratio: 2
Stopping criteria	Generations: 100; Stall generations: 50

Reproduction determines method which is used for each new generation creation. Elite count specifies the number of individuals that will be moved to the next generation. Crossover fraction defines the fraction of the next generation produced by a crossover manipulation. The remaining individuals in the next generation will be produced by mutation. Elite count has been selected to be 2 and Crossover fraction 0.8. This provides good results.

Mutation function provides genetic diversity by preventing the population of chromosomes from becoming similar to each other. The Adaptive feasible function has been specified to perform the mutation.

Crossover function combines two individuals in order to generate a new individual for the next generation. Several crossover functions have been tested. Heuristic function with ratio 2 gives satisfactory results.

The GA is iterated until a termination condition has been reached. Terminating conditions are the maximum number of iterations (100) and stall generations (50).

V. Experimental Results

The mechanical parameters were determined on the drive from Fig. 1. The cage induction motor has following rated data: 20 kW, 380 V, 40 A, 50 Hz, 1455 rpm, delta-connected windings. For this experiment the motor was connected to the sinusoidal three-phase supply 400 V, 50 Hz (transfer switch in Fig. 1 was in position 2) though the variac which controls the supply voltage. An optical encoder was attached to the shaft for speed measurement. During tests, the drive was unloaded, i.e. the synchronous machine (Fig. 1) was not excited.

Before the retardation test was done, the no-load test (at the motor's rated frequency) [15] of the drive was performed to separate the mechanical losses from the core losses. The result of this experiment is shown in Fig. 2.

For the retardation test the motor was supplied with the reduced voltage at the rated frequency and was run under the steady-state on the unloaded drive. The motor (drive) speed had to be slightly over rated speed and was set to 1490 rpm (156 rad/s), while the supply voltage was 113 V. It was necessary to wait some time to stabilize the temperature of the drive's bearings and then the motor was disconnected from the power supply and the speed-time curve was recorded. The result of the retardation test is shown in Fig. 3. From Fig. 3 a time derivative of the rotating speed can be obtained (Fig. 4). The specific part of the angular deceleration curve at low speeds is enlarged in Fig. 4. After the retardation test was done, the motor was reconnected to the power supply and ran at the near of its rated speed by controlling the motor's supply voltage. The motor (drive) speed was set to 1450 rpm (ω_1=151.8 rad/s), while the supply voltage was 57 V. The electrical input power to the motor P_{01}, the stator current I_{01} and the stator winding resistance R_s were measured. The core losses P_{Fe1} were read on the curve of Fig. 2, while the angular deceleration $d\omega/dt$ at the speed ω_1 was read on the curve of Fig. 4. The drive inertia was calculated according to (3). Table II describes the results obtained on the drive. The value of the drive inertia will be used in initial population of the GA.

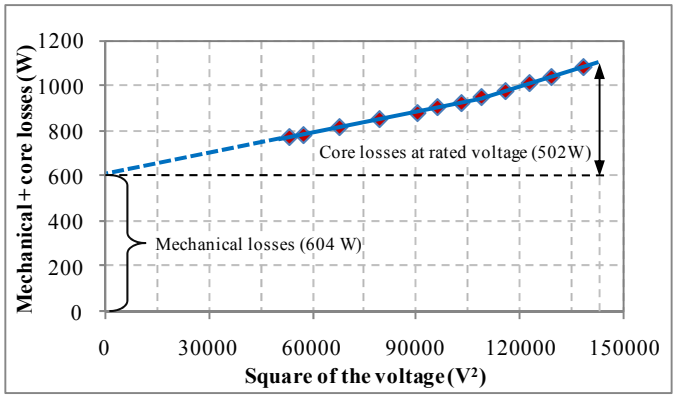


Figure 2. Separation of the no-load losses.

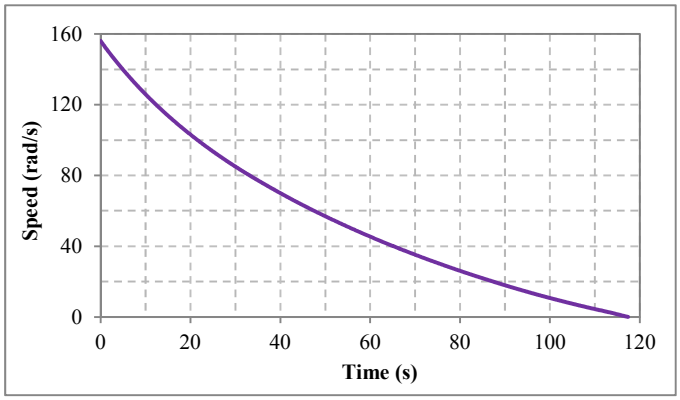


Figure 3. Retardation test on the drive.

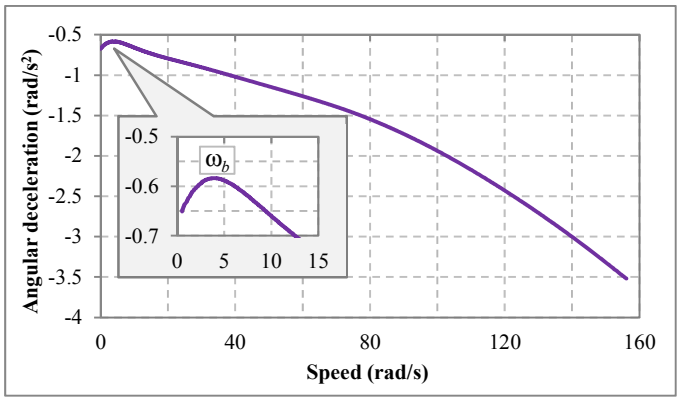


Figure 4. Angular deceleration curve.

TABLE II. The Results Obtained from the Retardation Test

ω_1 (rad/s)	$(d\omega/dt)_{\omega=\omega_1}$ (rad/s²)	P_{01} (W)	I_{01} (A)	R_s (mΩ)	P_{Fe1} (W)	P_1 (W)	J (kgm²)
151.8	− 3.379	602	7.37	311	10	556	1.084

The bearings and brush friction torque T_{bb} can be easily approximated as described in (5) for the bearings friction torque T_b determination. However, it is previously necessary to identify the speed ω_b at which the absolute value of the angular deceleration of the retardation curve begins to increase. This

characteristic speed was readily found on angular deceleration curve and its value is $\omega_b = 3.87$ rad/s (enlarged part of the curve in Fig. 4). Thereafter, the stator supply voltage was being reduced until the motor speed was decreased to the value of ω_b. Nevertheless, the proposed procedure was associated with the static stability problem of the drive system and exact value of ω_b could not be achieved. For that reason electrical input power to the motor P_b, the stator voltage U_b, current I_b and the stator winding resistance R_s were measured for a few operating points slightly above the speed ω_b, and then the bearings and brush friction torque T_{bb} was estimated for the speed ω_b by an extrapolation procedure. This procedure involved the use of (5) for each achieved operating point in order to obtain the torque T developed by the motor. Afterwards, the bearings and brush friction torque T_{bb} was extrapolated. It should be noted that (5) is valid for low-speed drive operation, as was the case here. The results are summarized in Table III.

Eventually, the mechanical parameters identification was undertaken using the proposed GA and the fitness function (13), including angular deceleration curve from Fig. 4, whereby the estimated drive inertia J and the bearings and brush friction torque T_{bb} from Table II and Table III, were used as initial values. Since the rotor core losses in (2) and (5) were neglected, as well as the windage and ventilation losses in the bearings and brush friction torque estimation, it was expected that the actual values of the drive inertia and the bearings and brush friction torque were lower than the previously estimated ones. With regard to this, previously estimated values of the drive inertia J and the bearings and brush friction torque T_{bb} were specified as upper bounds on these variables. The initial value of the coefficient a from (11) was selected to be zero, while the initial value of the coefficient b was roughly selected according to the mechanical losses data of the motor (drive) obtained by no-load test (Fig. 2), e.g. 0.024 (Table IV). The initial value of the coefficient a was specified as lower bound on this variable, while the coefficient b was specified as upper bound on this variable. The GA was run several times until the mechanical parameters set was obtained (Table IV). The objective function value of the best solution is $1.9809 \cdot 10^{-4}$.

TABLE III. THE RESULTS OF THE BEARING AND BRUSH FRICTION TORQUE ESTIMATION

ω (rad/s)	P_b (W)	U_b (V)	I_b (A)	R_s (mΩ)	T (Nm)	T_{bb} (Nm)
9.74	184.2	25.17	11.67	311	0.903	
7.64	169.4	24.17	11.20	311	0.830	-
5.12	153.3	23.00	10.70	311	0.749	
3.87	-					0.709

TABLE IV. THE RESULTS OBTAINED BY THE GA APPROACH

Mechanical parameters	J (kgm^2)	T_{bb} (Nm)	a	b
Initial values	1.084	0.709	0	0.024
Final values	1.078	0.6544	0.00098	0.0093

The mechanical losses components of the tested drive, obtained by the proposed methods, are shown in Fig. 5. It should be noted that low-speed drive operation was not considered in this paper, i.e. it is not covered by the proposed mechanical losses model. It is assumed that below the speed ω_b the bearings friction torque starts to increase as speed decreases since a lubricant film is lost. This is not included in the proposed bearings and brush friction torque model. For that reason the mechanical parameters set from Table IV are not valid below speed ω_b (Fig. 4). However, it is interesting to calculate the mechanical losses of the drive (P_m) from Fig. 1 at no-load (ω=157 rad/s) according to the proposed model and the loss coefficients obtained by the GA approach (Table IV):

$$P_m = \left(T_{bb} + b \cdot \omega^{(1+a\cdot\omega)}\right) \cdot \omega = 602 \text{ W} . \qquad (14)$$

The result of (14) is in good agreement with the experimental result obtained by no-load test of the drive (Fig. 2).

In order to validate the proposed mechanical losses model and methods for the estimation of the mechanical parameters of the drive, additional experiments were carried out. The motor was connected to a 60 Hz source at reduced voltage (transfer switch in Fig. 1 was in position 1 and the motor was connected to power inverter) and was run under steady-state on the unloaded drive. Then the retardation test was performed. The result of this experiment is shown in Fig. 6, where it is also shown the estimated retardation curve obtained according to (12) with the parameters from Table IV.

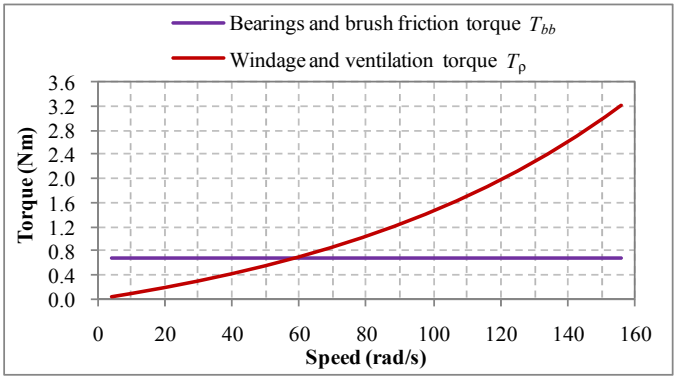


Figure 5. The mechanical losses components of the tested drive.

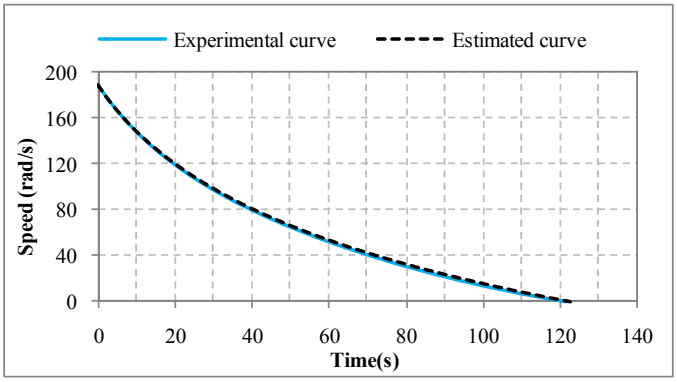


Figure 6. Retardation test on the drive at 60 Hz.

As shown in Fig. 6 the estimated retardation curve is in a good agreement with the experimental curve.

Another experiment was performed on the drive from Fig. 1 with the flywheel removed. The flywheel is a solid, uniform disk which has inertia of 0.1454 kgm^2. Since the flywheel was directly coupled to the motor shaft, after it was removed, the drive inertia referred to the motor shaft is 0.9336 kgm^2. In this experiment the motor was supplied with the reduced voltage at the rated frequency and was run under the steady-state on the unloaded drive. At that moment the retardation test was performed. The result of this experiment is shown in Fig. 7. The same figure shows the estimated retardation curve obtained according to (12) with the parameters from Table IV, but with the drive inertia of 0.9336 kgm^2. As can be seen in Fig. 7, there is a certain discrepancy between the experimental and the estimated retardation curves. This disagreement is noticeable at lower speeds regions. The reason for discrepancy between results can be found in the resultant bearings load. Namely, when the flywheel was removed, according to (4) the resultant bearings load was decreased. Besides, disagreement also exists because there were no the mechanical losses due to flywheel air resistance to motion. To correct the estimated retardation curve, the bearings and brush friction torque was estimated again employing the proposed GA method, while the other mechanical parameters (a and b from Table IV) were not changed. The results are shown in Fig. 8. The estimated retardation curve, obtained according to (12) with new estimated value of the bearings and brush friction torque (0.62 Nm), shows good agreement with the experimental curve, validating the effectiveness of the proposed method.

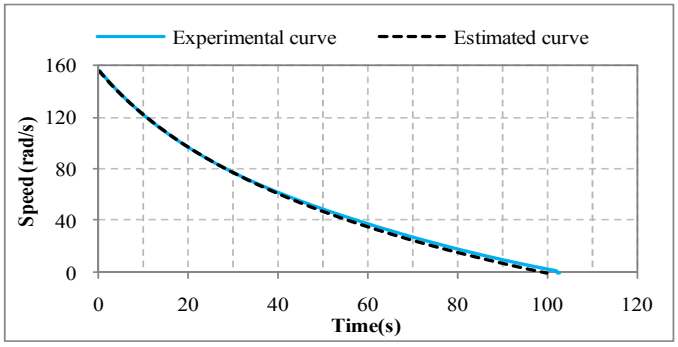


Figure 7. Retardation test on the drive at 50 Hz without flywheel.

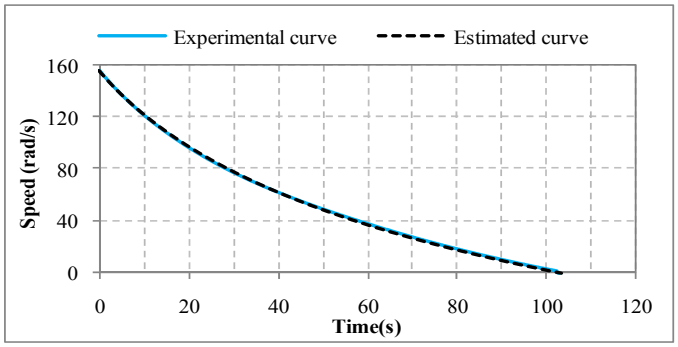


Figure 8. Retardation test on the drive at 50 Hz without flywheel - Corrected estimated curve.

VI. Conclusion

In this paper, the experimental method for the identification of the mechanical parameters of the induction motor drive was presented. The method is based on the improved retardation test with the help of the GA.

In addition, the mechanical losses model of the tested drive was established. Novel modeling approach of the windage and ventilation losses was introduced. The model is much more suitable for practical use, compared to the ones that can be found in literature, since it does not require a detailed knowledge of the machines geometry.

The effectiveness of the proposed methods was demonstrated by experiments. Good agreements were found between the estimated and the experimental results.

Future researches should follow up on improving the mechanical losses model of the drive system in low-speed regions and at standstill.

Acknowledgment

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under project III42004.

References

[1] S K Bhattacharya, Electrical machines, 3rd ed., New Delhi, India: Tata McGraw-Hill Education, 2009.

[2] B. Blanusa and B. Knezevic, "Simple Hybrid Model for Efficiency Optimization of Induction Motor Drives with Its Experimental Validation," Advances in Power Electronics, vol. 2013, Article ID 371842, pp. 1-8, February 2013.

[3] H. Sediki, A. Bechouche, D. O. Abdeslam, S. Haddad, "ADALINE approach for induction motor mechanical parameters identification," Mathematics and Computers in Simulation, vol 90, pp. 86-97, April 2013.

[4] S. M. Yang, Y. J. Deng, "Observer-Based Inertial Identification for Auto-Tuning Servo Motor Drives," IEEE Industry Applications Conference, vol. 2, pp. 968-972, Oct. 2005.

[5] F. Alonge, F. D'Ippolito, G. Ferrante, F.M. Raimondi, "Parameter identification of induction motor model using genetic algorithms," Control Theory and Applications, IEE Proceedings, vol. 145(6), pp. 587–593, Nov. 1998.

[6] S.K. Pillai, A first course on electrical drives, 2nd ed., New York: Wiley, 1989.

[7] Fouad Giri, AC Electric Motors Control : Advanced Design Techniques and Applications, New York: Wiley, 2013.

[8] M.V. Deshpande, Electric Motors: Applications and Control, New Delhi, India: Phi Learning, 2010.

[9] G. K. Dubey, Fundamentals of Electrical Drives, 2nd ed., India: Alpha Science International, 2001.

[10] W. Saekok, P. Lumyong, "Characteristics Evaluation of 3 Phase Induction Motors Based on an Acceleration Method with Increasing Moment of Inertia Technique," SDEMPED 2003, pp. 93-98, August 2003.

[11] I. D. Ilina, "Experimental determination of moment to inertia and mechanical losses vs. speed, in electrical machines," 7th International Symposium on Advanced Topics in Electrical Engineering, pp. 1-4, May 2011.

[12] L. Borello, M. D. L. Dalla Vedova, "Dry Friction Discontinuous Computational Algorithms," International Journal of Engineering and Innovative Technology (IJEIT), vol. 3(8), pp. 1-8, February 2014.

[13] I. Virgala, P. Frankovský, M. Kenderová, "Friction Effect Analysis of a DC Motor," American Journal of Mechanical Engineering, vol 1(1), pp. 1-5, 2013.

[14] (2014) LD DIDACTIC Group webpage. [Online]. Available: http://www.ld-didactic.de/literatur/hb/e/p1/p1452_e.pdf

[15] IEEE Standard Test Procedure for Polyphase Induction Motors and Generators, IEEE Std 112-2004

[16] T. L. Mthombeni, P. Pillay, "Lamination Core Losses in Motors With Nonsinusoidal Excitation With Particular Reference to PWM and SRM Excitation Waveforms," IEEE Transactions on Energy Conversion, vol. 20, No. 4, pp. 836-843, December 2005.

[17] S. Khomfoi, V. Kinnares, P. Viriya, "Influence of PWM characteristics on the core losses due to harmonic voltages in PWM fed induction motors," Power Engineering Society Winter Meeting, vol. 1, pp. 365-369, Jan. 2000.

[18] G. P. Shultz, Transformers and Motors, 1st ed., USA: Elsevier Science, 1997.

[19] (2013) SKF Rolling bearings and seals in electric motors and generators.

[20] (2012) SKF rolling bearings catalogue.

[21] J. Croes and S. Iqbal, D2.1 Document 3: Literature Survey: bearing losses, Energy Software Tools for Sustainable Machine Design EC - 7th Framework Programme, K.U.Leuven, 2009.

[22] J. Pyrhonen, T. Jokinen and V Hrabovcova, Design of Rotating Electrical Machines, 1st ed., Chippenham, Great Britain: John Wiley & Sons, 2008.

[23] W. Tong, "Numerical Analysis of Flow Field in Generator End-Winding Region," International Journal of Rotating Machinery, vol. 2008, Article ID 692748, pp. 1-10, 2008.

[24] Improving the Efficiency of Squirrel Cage Induction Motors: Technical and Economical Consideration, The Swedish Energy Agency publication. [Online]. Available: http://www.energimyndigheten.se/PageFiles/18106/Elmotorer%20Beakt ande%20803970.pdf

[25] N. Voicu, L. Dumitriu, M. Iordache, D. Niculae, "Modeling of the Induction Motor Ventilation by Electric Circuits," The Scientific Bulletin, Series C, vol. 71 (4), pp. 175-182, 2009.

[26] H. A. Toliyat, G. B. Kliman, Handbook of Electric Motors, USA: CRC Press, 2004.

[27] K. Ludlu, "Optimizing Flywheel Design for use as a Kinetic Energy Recovery System for a Bicycle," Senior Theses, Pomona College, Claremont, California, 2013.

[28] J. H. Holland, Adaptation in Natural and Artificial Systems, 2nd ed., Cambridge, MA: MIT Press, 1992.

[29] MATLAB and Optimization Toolbox, The MathWorks, Inc., Natick, Massachusetts, United States.

[30] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, 3rd ed., Berlin, Germany: Springer-Verlag, 1996.

Design of Tesla's Two-Phase Inductor

Đorđe M. Lekić, Petar R. Matić
University of Banja Luka
Faculty of Electrical Engineering
Banja Luka, Republic of Srpska, Bosnia and Herzegovina
djordje.lekic@etfbl.net, pero@etfbl.net

***Abstract*—This paper describes a new method for designing Tesla's two-phase inductor for demonstration of rotating magnetic field, based on the famous Tesla's Egg of Columbus experiment. The design is based on electromagnetic and thermal analytical models of the two-phase inductor and verified by computer simulation and FEM analysis.**

Keywords-FEM analysis; rotating magnetic field; single-phase induction motor; Tesla's Egg of Columbus;

NOMENCLATURE

D, d	Outer and inner iron core diameter [mm],
a, h	Iron core width and height [mm],
$\mu_0 = 4\pi \cdot 10^{-7}$	Permeability of free space [H/m],
ω, f	Angular frequency [rad/s] and frequency [Hz],
α, β	Phase index,
N_α, N_β, N	Number of turns per phase,
$m = N_\beta / N_\alpha$	Transfer coefficient,
$N_{1/2} = N/2$	Number of turns per half-winding,
$\mathfrak{R}_g$	Equivalent reluctance of air gap [H^{-1}],
X_α, X_β, X	Reactance of phase winding [Ω],
$X_{1/2} = X/2$	Reactance of one half-winding [Ω],
R_α, R_β, R	Resistance of phase winding [Ω],
Z_α, Z_β, Z	Impedance of phase winding [Ω],
X_C, R_C	Capacitor reactance and resistance [Ω],
U_C, C	Capacitor voltage [V] and capacitance [μF],
$\rho_{Cu0} = 1/57$	Specific resistance of copper at 20°C [Ωmm^2/m],
l_{Cu}, $L_{Cu} = N \cdot l_{Cu}$	Average length of turn and phase winding [m],
$S_{Cu\alpha}$, $S_{Cu\beta}$, S_{Cu}	Copper wire cross section [mm^2],
$d_{Cu\alpha}$, $d_{Cu\beta}$, d_{Cu}	Copper wire diameter [mm],
n_α, n_β, n	Number of conductor layers,
I_α, I_β, I	RMS value of phase and total current [A],
U_α, U_β, U	RMS values of phase and supply voltage [V],
F_α, F_β, F	Amplitude of magneto-motive force [Aturns],
$\Phi_{g\alpha}$, $\Phi_{g\beta}$, Φ_g	Amplitude of air gap magnetic flux [Wb],
$B_{g\alpha}$, $B_{g\beta}$, B_g	Amplitude of air gap magnetic flux density [T],
Δ, P_{Cu}	Current density [A/mm^2] and copper losses [W],
m_{Cu}, V_{Cu}	Total mass [kg] and volume of copper [m^3],
$\gamma_{Cu} = 8900$	Specific mass of copper [kg/m^3],
θ, θ_a	Operating and ambient temperature [°C],
$\theta_f = (\theta + \theta_a)/2$	Average (film) temperature [°C]
$\Delta\theta = \theta - \theta_a$	Temperature increase [°C]
α_c, α_r, α_t	Heat transfer coefficients [W/°C/m^2],
S_{hc}, S_{hr}	Heat transfer surface areas [mm^2],
$\sigma = 5.674 \cdot 10^{-8}$	Stefan-Boltzmann constant [W/m^2/K^4],
ε	Emissivity coefficient,
Nu, Gr, Pr	Nusselt, Grashoff and Prandtl number,
λ_f	Thermal conductivity of air [W/°C/m],
v_f	Kinematic viscosity of air [m^2/s],
$g = 9.81$	Acceleration of free fall [m/s^2],
$\beta = (235 + \theta_f)^{-1}$	Thermal expansion coefficient [°C^{-1}],
$\gamma_f = 1.2928$	Specific mass of air [kg/m^3],
$c_f = 1005$	Specific thermal capacity of air [J/kg/°C].

I. INTRODUCTION

In February 1882, the Serbian inventor and scientist Nikola Tesla came up with the design of the rotating magnetic field and in 1888 he constructed the first induction motor. Tesla's first induction motor had a toroidal stator with a two-phase inductor, unlike a cylindrical stator with a three phase inductor which is nowadays a standard design. In order to demonstrate the effects of the rotating magnetic field at the World fair exhibition in Chicago in 1893, Tesla used his two-phase inductor to lift and spin a metal egg, thus "performing the feat of Columbus without cracking the egg"[1]. Tesla's original design of the two-phase inductor is shown in Fig. 1.

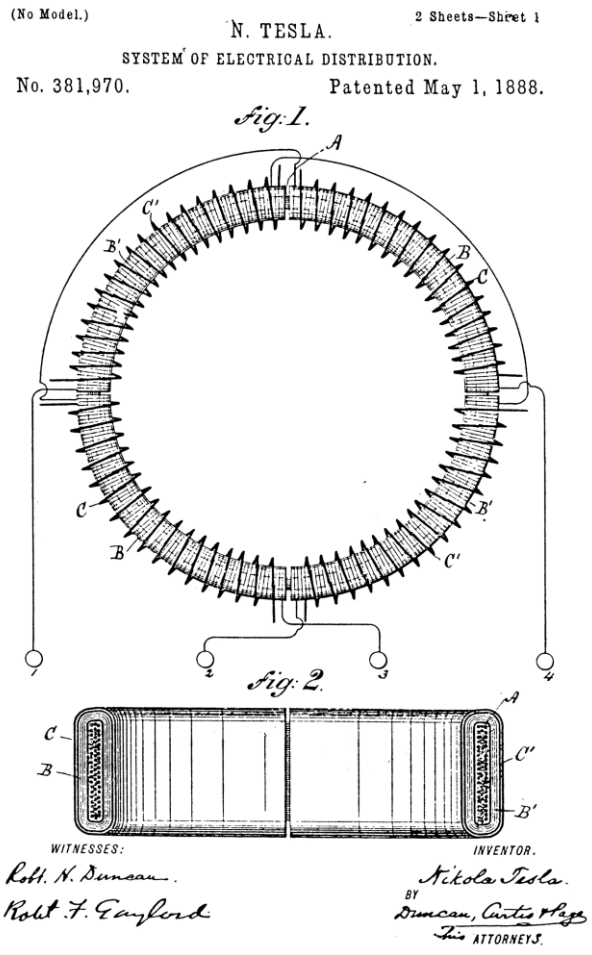


Figure 1. The original design of the two-phase Tesla inductor [2]

Two coils are wound on an iron core of toroidal shape. The two phase coils are displaced by 90 (geometrical) degrees and are supplied with two alternating currents phase shifted by 90 (electrical) degrees. In this way a rotating magnetic field is generated in the air gap. If a metal object, e.g. a copper or aluminum egg, is placed on a platform in the centre, it will start to rotate, and eventually stand on its main axis, due to gyroscopic effect. Tesla's two-phase inductor is in fact the stator of a two-phase induction motor, with the metal egg being the rotor.

Replicas of "Tesla's Egg of Columbus" can be found in museums dedicated to Tesla and almost every Faculty of electrical engineering has its own unique copy. There are even companies that create authentic replicas for museums and private collectors [3]. These replicas are not difficult to reproduce, however in most cases they deviate from Tesla's original design. For example in [4] using the laws of electrical machine resemblance, a stator of a conventional four-pole, three-phase induction motor was redesigned to demonstrate the "Egg of Columbus" experiment, and in [5] an electronically controlled rotating magnetic field inductor, using power switches, oscillator and counter circuits is presented. All these replicas don't properly illustrate the brilliance and simplicity of the rotating magnetic field concept and they are not a practical educational tool which, in a simple way, conveys basic electro-technical engineering principles to young students.

This paper gives the design and construction principles of Tesla's two phase inductor. The paper is organized as follows. In Section II the analytical model for electromagnetic design is presented and proper level of magnetic flux density in air gap is selected. Equations for calculating the phase reactance and resistance, required capacitance in one phase and number of turns in each phase are derived. Section III gives a method for determining adequate current density, based on analytical thermal analysis. Using the derived equations, all the necessary parameters are calculated in Section IV and the results of the design are verified via computer simulation and Finite Element Method (FEM) analysis. Conclusion is given in Section V.

II. ELECTROMAGNETIC DESIGN

The design of electrical machines is an iterative procedure based on calculations and assumptions on the first step, and on necessary corrections and experimental verification on the second step. Assuming fixed dimensions of used iron core, the electromagnetic design starts with the selection of magnetic flux and current density and ends by calculating number of turns in each phase.

A. Magnetic Circuit

The magnetic circuit for one phase winding of the Tesla inductor is shown in Fig. 2. The phase winding is divided in two halves which are connected in series, but wound in opposite directions and placed on opposing sides of the iron core. The resulting magnetic flux Φ_g passes through the air gap, while only half of the resulting flux $\Phi_{Fe} = \Phi_g/2$ passes through each side of the iron core. For design analysis, the magnetic circuit in Fig. 2 can be represented with an equivalent circuit, shown in Fig. 3. The magnetic circuit equivalent takes only the equivalent magnetic resistance (reluctance) of the air gap into account, while the reluctance of the iron core is neglected due to high values of iron core permeability.

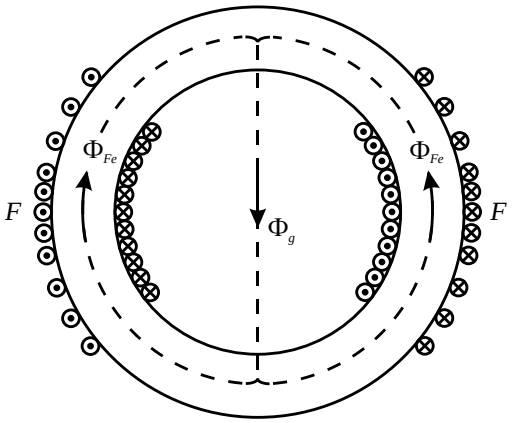


Figure 2. Magnetic circuit

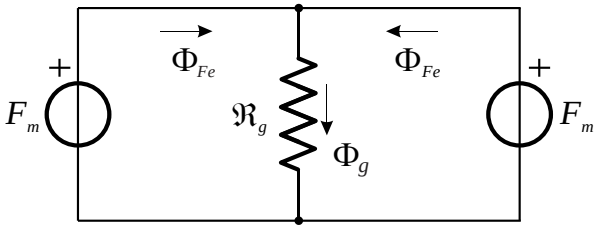


Figure 3. Magnetic circuit equivalent

B. Phase Impedance

Assuming a homogeneous magnetic field in the air gap, the phase reactance can be calculated by using similar approach to slot leakage reactance calculation for standard induction motors [6]. Later in the text, by using computer simulation and FEM analysis, it will be shown that this is a very good approximation. The reactance of one half-winding of each phase, is:

$$X_{1/2} = \omega \cdot \frac{N_{1/2}{}^2}{\Re_g} = \frac{1}{3} \cdot \pi \cdot \mu_0 \cdot h \cdot \omega \cdot N_{1/2}{}^2 \qquad (1)$$

where the equivalent air gap reluctance is:

$$\Re_g = \frac{3}{\pi \cdot \mu_0 \cdot h} \qquad (2)$$

Since two half-windings are connected in series to form one phase winding, after introducing frequency and number of turns per phase in (1), the following equation for reactance of each phase winding is obtained:

$$X = 2 \cdot X_{1/2} = \frac{1}{3} \cdot \pi^2 \cdot \mu_0 \cdot h \cdot f \cdot N^2 \qquad (3)$$

The phase resistance at θ_0 = 20°C can be calculated as:

$$R_0 = \rho_{Cu0} \cdot \frac{L_{Cu}}{S_{Cu}} = \rho_{Cu0} \cdot \frac{N \cdot l_{Cu}}{S_{Cu}} \tag{4}$$

The cross section of copper wire in (4) is:

$$S_{Cu} = \frac{\pi}{4} \cdot {d_{Cu}}^2 \tag{5}$$

The average length of copper wire is:

$$L_{Cu} = 2 \cdot N_{1/2} \cdot l_{Cu} = 2 \cdot N \cdot \left(a + h + 2 \cdot n \cdot d_{Cu}\right) \tag{6}$$

The coil resistance at operating temperature θ is:

$$R = R_0 \cdot \frac{235 + \theta}{235 + \theta_0} \tag{7}$$

The impedance of each phase winding is then:

$$Z = \sqrt{R^2 + X^2} \tag{8}$$

In equations (1)-(8) phase indexes have been left out, meaning that the equations are valid for both phases. In order to calculate the parameters for a particular phase, one must use the number of turns, number of conductor layers and cross section of that particular phase winding.

C. Permenantly Connected Capacitor

In order to get a symmetrical rotating magnetic field, it is needed to have equal amplitudes of Magneto-Motive Forces (MMFs) for both phases, and an electrical phase delay of 90º between the phase curents. In standard single-phase induction motors, this is achieved by using permenantly connected capacitor in one of the phases, e.g. phase β in Fig. 4. For calculating the capacitance in phase β a standard procedure described in [7] will be used.

The transfer ratio is calculated using [7]:

$$m = \frac{X_\alpha + \sqrt{{X_\alpha}^2 - 4 \cdot R_C \cdot R_\alpha}}{2 \cdot R_\alpha} \tag{9}$$

while capacitor reactance and capacitance are [7]:

$$X_C = m^2 \cdot X_\alpha + m \cdot R_\alpha \tag{10}$$

$$C = \frac{1}{\omega \cdot X_C} = \frac{1}{2 \cdot \pi \cdot f \cdot X_C} \tag{11}$$

While calculating the transfer ratio, one must assume the value of capacitor resistance, but for rough calculations, the capacitor resistance can be neglected.

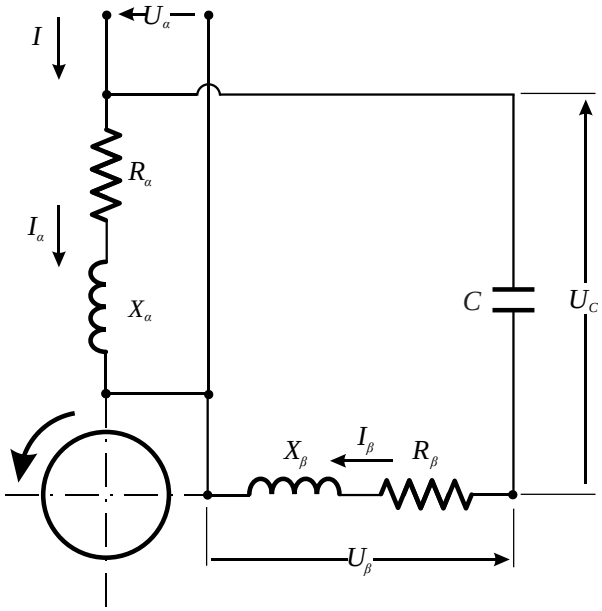


Figure 4. Electric circuit equivalent

D. Number of Turns per Phase

The magnetic flux density in the centre of the air gap due to excitation from one phase can be calculated from Fig. 3. For that purpose, the Kapp-Hopkins law for magnetic circuits is used:

$$\Phi_g = \frac{F}{\Re_g} \tag{12}$$

The Magneto-Motive Force (MMF) generated from both phases is:

$$F = \frac{\sqrt{2}}{2} \cdot N_\alpha \cdot I_\alpha = \frac{\sqrt{2}}{2} \cdot N_\beta \cdot I_\beta \tag{13}$$

With the assumption of a homogeneous magnetic field in the air gap, the amplitude of the magnetic flux is given by:

$$\Phi_g = B_g \cdot S_g = B_g \cdot h \cdot d \tag{14}$$

Substituting (2), (13) and (14) in (12) the following equations for calculating the number of turns in each phase are derived:

$$N_\alpha = \frac{3}{\pi} \cdot \sqrt{2} \cdot \frac{B_g}{\mu_0} \cdot \frac{d}{I_\alpha} \tag{15}$$

$$N_\beta = \frac{3}{\pi} \cdot \sqrt{2} \cdot \frac{B_g}{\mu_0} \cdot \frac{d}{I_\beta} \tag{16}$$

From (15) and (16) it can be seen that required level of magnetic field density in the air gap can be achieved by increasing the number of turns, or by increasing the current.

The required level of magnetic field density is determined from the condition that it only has to spin the egg, and therefore can be much lower than as in conventional electric machines. This level is determined by analyzing a prototype. The results of the experiments show that for rotating magnetic field demonstration, a magnetic flux density of 5-10 mT is large enough to lift and spin a smaller metal egg. Heaving this range of values in mind, the design problem is to select the appropriate number of turns for each phase winding in order to obtain required magnetic flux density without overheating the coils due to large currents needed to generate this flux density in the large air gap.

If a standard value of copper wire diameter for phase α is selected, with a proper value for current density without overheating, the phase current and the number of turns in phase α, using (15) can be calculated. After the number of turns is found, the reactance and resistance of phase α using (3) and (4) and the required transfer ratio using (9) can be calculated. In that case, phase current and number of turns in phase β is further calculated as:

$$I_\beta = \frac{I_\alpha}{m} \tag{17}$$

$$N_\beta = m \cdot N_\alpha \tag{18}$$

From (17) the diameter and cross section of copper wire in phase β should be adopted to the closest standard values. Since the core has fixed dimensions, it must be must ensured that there is enough space to place both phase windings on the core. The windings will fit on the core if they are wounded in n layers, where n is calculated for each phase independently using:

$$n \geq \frac{2 \cdot N \cdot d_{Cu}}{\pi \cdot d} \tag{19}$$

where the first larger integer value should be adopted. At the end the number of turns for phase β is calculated by using the same equations as for phase α.

III. THERMAL DESIGN

Current density of conventional induction motors is typically in the range 3-8 A/mm^2 [6]. Current density determines the cross section of copper wire, or in other words, the total mass of copper used to wound the inductor. Since the losses depend on the value of current density, the steady state temperature of the windings also depends on current density. In order to prevent overheating, a thermal model is used to select the required value of current density.

A. Coper Losses

The design is based on requirement that the steady state temperature of windings does not exceed 80ºC. Since the magnetic flux density is in the range of 5-10 mT, the iron losses can be neglected, and only the copper losses are considered.

The total copper losses in both phases at operating temperature θ are:

$$P_{Cu} = R_\alpha \cdot {I_\alpha}^2 + R_\beta \cdot {I_\beta}^2 = 2 \cdot R_\alpha \cdot {I_\alpha}^2 \tag{20}$$

If it is assumed that operating temperature should be θ = 80ºC, the resistance should be increased using (7) and the following equation for copper losses in common electrical machines is obtained [6]:

$$P_{Cu} = \rho_{Cu} \cdot \frac{m_{Cu}}{\gamma_{Cu}} \cdot \Delta^2 \approx 2.42 \cdot m_{Cu} \cdot \Delta^2 \tag{21}$$

where Δ is given in A/mm^2 and total mass of copper is:

$$m_{Cu} = \gamma_{Cu} \cdot V_{Cu} = \gamma_{Cu} \cdot \left(L_{Cu\alpha} \cdot S_{Cu\alpha} + L_{Cu\beta} \cdot S_{Cu\beta} \right) \tag{22}$$

From (21) the losses per cubic meter of copper depend only on the current density:

$$\frac{P_{Cu}}{V_{Cu}} \approx 2.42 \cdot \gamma_{Cu} \cdot \Delta^2 = 21538 \cdot \Delta^2 \tag{23}$$

If the copper losses from (21) are combined with (4) and (15) or (16), the following equation is derived:

$$P_{Cu} = \frac{6\sqrt{2}}{\pi} \cdot \frac{\rho_{Cu} \cdot d \cdot l_{Cu}}{\mu_0} \cdot B_g \cdot \Delta = k \cdot B_g \cdot \Delta \tag{24}$$

in which k is a constructive constant which depends on characteristics of used materials and core dimensions. From (24) the characteristic product of magnetic flux and current density can be found which will provide that copper losses don't lead to overheating.

B. Current Density

The allowed copper loss can be calculated using Newton's law of cooling [6,8]:

$$P_{\max} = \alpha_t \cdot S_{hc} \cdot \Delta\theta \tag{25}$$

where convection and radiation heat transfer surface areas and total heat transfer coefficient are:

$$S_{hc} = S_{hr} + \pi \cdot h \cdot d = \frac{\pi}{4} \cdot \left[4 \cdot h \cdot \left(D + d \right) + D^2 - d^2 \right] \tag{26}$$

$$\alpha_t = \alpha_c + \alpha_r \cdot \frac{S_{hr}}{S_{hc}} \tag{27}$$

Using (25) the copper losses which will lead to steady state temperature rise of θ = 80ºC can be calculated, and after that, using (24) the maximum allowed current density for a given value of magnetic flux density is calculated.

The heat transfer coefficient due to radiation is calculated using Stefan–Boltzmann's law [8]:

$$\alpha_r = \varepsilon \cdot \sigma \cdot \frac{(273+\theta)^4 - (273+\theta_a)^4}{\theta - \theta_a} \tag{28}$$

where the emissivity is taken to be 0.2 for new, and close to 1 for old copper conductors. The heat transfer coefficient due to natural convection is calculated using dimensional thermal analysis, involving Prandtl, Grashoff and Nusselt numbers, and is given by following equation [8]:

$$\alpha_c = Nu \cdot \frac{\lambda_f}{h} \tag{29}$$

In (29) Nusselt number is a function of the product of Grashoff and Prandtl number and is calculated using [8]:

$$Nu = A \cdot (Gr \cdot Pr)^n \tag{30}$$

where values for coefficients *A* and *n* depend on product of Grashoff and Prandtl number and are given in Table I [8].

TABLE I. NUSSELT NUMBER COEFFICIENTS

Gr · Pr	***A***	***n***
10^{-10}- 10^{-4}	0.675	0.058
10^{-4}- 10^{-1}	0.889	0.088
10^{-1}- 10^{2}	1.020	0.148
10^{2}- 10^{4}	0.850	0.188
10^{4}- 10^{7}	0.480	0.250
10^{7}- 10^{12}	0.125	0.333

Grashoff number takes into account forces due to differences in density of heated air layers and is calculated using following equation [8]:

$$Gr = \frac{g \cdot \beta \cdot \Delta\theta \cdot h^3}{\nu_f^{\,2}} \tag{31}$$

Prandtl number depends on physical properties of air and can be calculated as [8]:

$$Pr = \frac{\gamma_f \cdot \nu_f \cdot c_f}{\lambda_f} \tag{32}$$

but is often taken to have a constant value of 0.7.

In (29), (30), (31) and (32) thermal conductivity of air, Prandtl number and viscosity of air are functions of average (film) temperature and are calculated as [8]:

$$\lambda_f = 2.42 \cdot 10^{-2} + 7.2 \cdot 10^{-5} \cdot \theta_f \tag{33}$$

$$Pr = 0.715 - 2.5 \cdot 10^{-4} \cdot \theta_f \tag{34}$$

$$\nu_f = 1.32 \cdot 10^{-5} + 9.5 \cdot 10^{-8} \cdot \theta_f \tag{35}$$

Assuming a value of steady state and ambient temperature and using equations (28)-(35), the total heat transfer coefficient is calculated, and value for allowed copper losses is obtained from (25). The allowed current density is then calculated from (24) and the final result of this calculation is showed on diagram in Fig. 5. The diagram shows allowed current density versus operating temperature, for several values of ambient temperature.

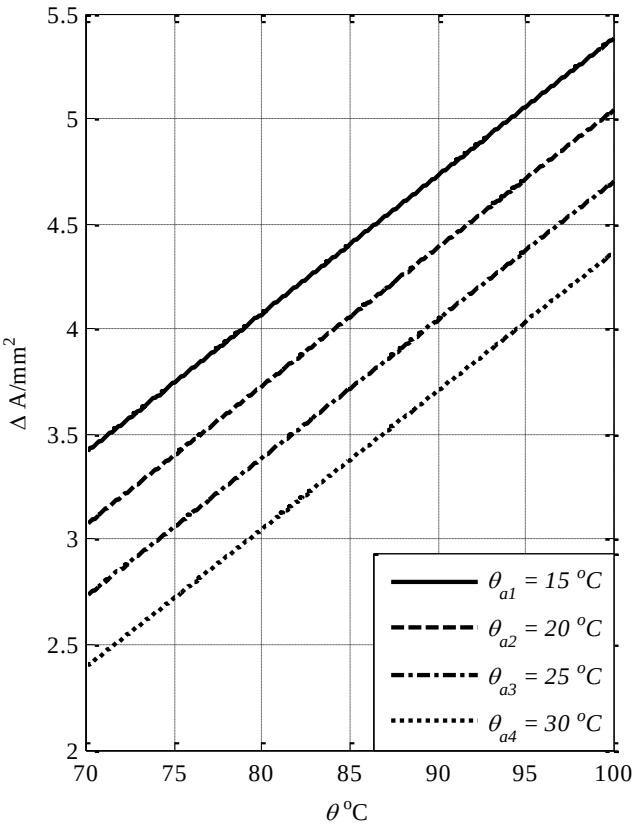


Figure 5. Current density versus operating temperature

It can be seen from diagram on Fig. 5 that for operating temperature of θ = 80°C and ambient temperatures of 15°C to 30°C, allowed current density is in the range of 3-4 A/mm^2. For most likely ambient temperatures of 20°C to 25°C current density is approximately 3.5 A/mm^2.

IV. RESULTS OF DESIGN

A. Calculation

For demonstration of the proposed calculation of two-phase Tesla inductor a toroidal iron core from an instrument transformer for voltage 258V and current 25A is selected. The dimensions of the iron core are given in Fig. 6 with the values in Table II.

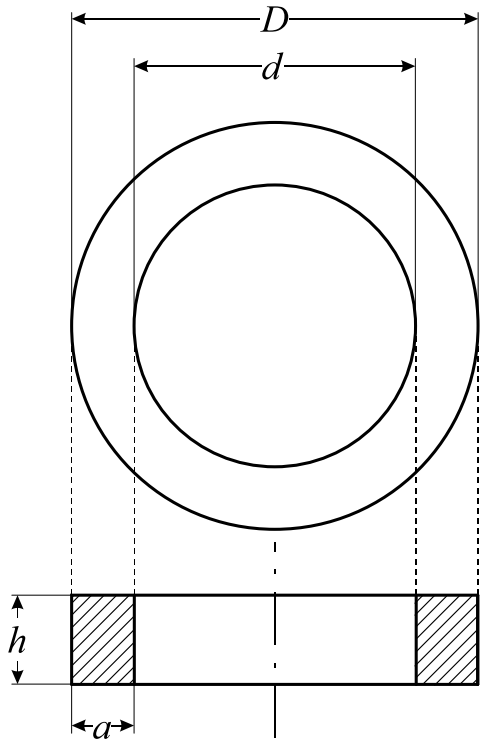


Figure 6. Dimensions of the used iron core

TABLE II. CORE DIMENSIONS

Parameter	Unit	Value
D	mm	260
d	mm	180
h	mm	60
a	mm	40

For selected standard copper wire with diameter of 1 mm for phase α, using equations derived in Section II and Section III, with the selected values of 10 mT for magnetic flux density and 3.5 A/mm^2 for current density, all the parameters for constructing a two-phase Tesla inductor are calculated and their values are given in Table III.

TABLE III. RESULTS OF CALCULATION

Parameter	Unit	Value
$d_{Cu\alpha}$	mm	1.0
$d_{Cu\beta}$	mm	0.8
$S_{Cu\alpha}$	mm^2	0.7854
$S_{Cu\beta}$	mm^2	0.5027
B_g	mT	10
Δ	A/mm^2	3.5
I_α	A	2.75
I_β	A	1.84
N_α	turns	704
N_β	turns	1050
m	/	1.5
n_α	layers	3
n_β	layers	4
R_α	Ω	4.11
R_β	Ω	9.49
X_α	Ω	6.15
X_β	Ω	13.67
Z_α	Ω	7.40
Z_β	Ω	16.64
X_C	Ω	19.84
R_C	Ω	/
C	μF	160
U_α	V	21
U_β	V	31
U_C	V	37
$L_{Cu\alpha}$	m	150
$L_{Cu\beta}$	m	220
m_{Cu}	kg	2
V_{Cu}	m^3	$2.28 \cdot 10^{-4}$
P_{Cu}	W	63.3
P_{Cu}/V_{Cu}	W/m^3	277840

It should be noted that total wire lengths in each phase should be increased for approximately 10% in order to take into account side connections and terminal connections. In this way, the total mass and volume of copper will be slightly increased.

B. *Electromagnetic Simulation*

For electromagnetic and thermal simulation, FEMM software is used. Fig. 7 shows magnetic flux density plot from phase α, at the moment when the current in phase α is at its maximum, while current in phase β is equal to zero. This justifies the assumption of homogeneous magnetic filed in the air gap. Fig. 8 shows the similar results for phase β. From Fig. 7 and Fig. 8 it can be concluded that the magnetic fields are orthogonal, thus creating a rotating magnetic field in the air gap.

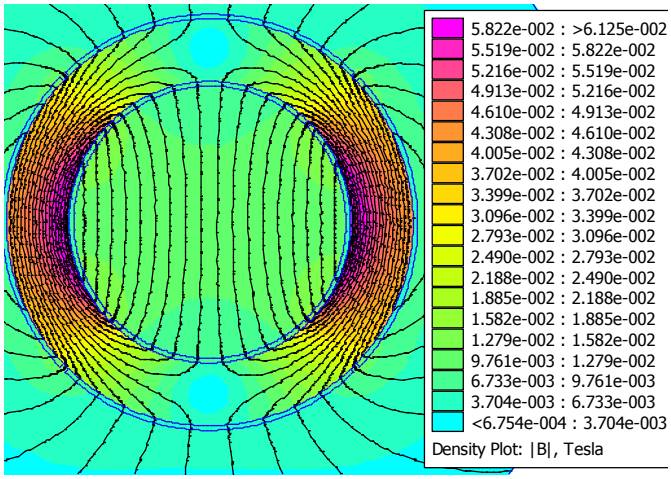


Figure 7. Magnetic flux density plot for phase α

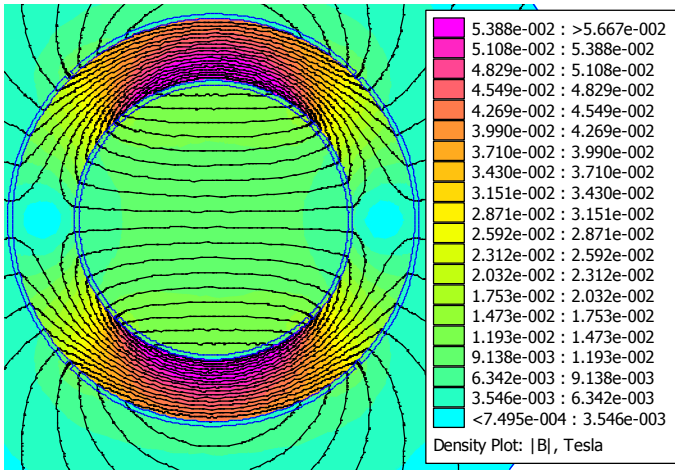


Figure 8. Magnetic flux density plot for phase β

Fig. 9 shows value of magnetic flux density along the air gap. It is clear from Fig. 9 that the magnetic flux density changes from 11 mT to 12.5 mT along the diameter of the air gap, which is more than anticipated 10 mT. The coil parameters are also calculated by FEM simulation (Table III), and compared with analytically calculated values shown in Table II. It can be seen that the calculation error is less than 3% in all cases.

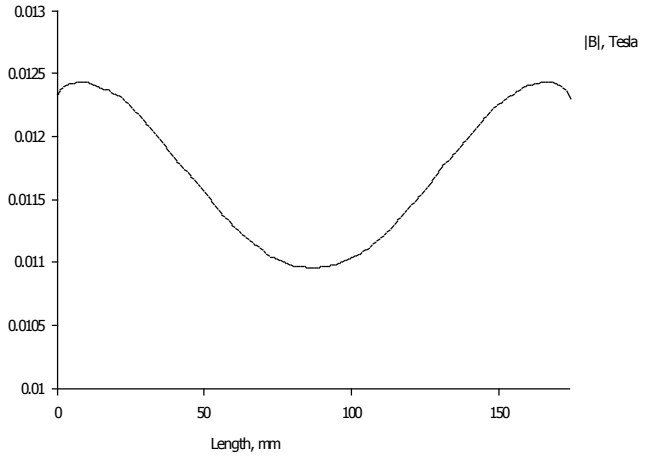


Figure 9. Magnetic flux density distribution along air gap

TABLE IV. SIMULATION RESULTS

Parameter	Unit	Value	Calculation error %
R_α	Ω	4.13	0.50
R_β	Ω	9.61	1.25
X_α	Ω	6.30	2.38
X_β	Ω	14.00	2.36

C. *Thermal Simulation*

The electromagnetic FEM simulation considered a planar problem, while the thermal analysis should consider a axisymmetric problem. The copper winding is modeled with a hollow torus, with the same volume as calculated in Table II, while copper losses per cubic meter are calculated from (23). The calculated temperature distribution is showed in Fig. 10 from which it can be seen that the highest temperature is 354 K or 81°C, which is almost equal to the designed temperature of 80°C. During simulation an ambient temperature of 20°C was assumed.

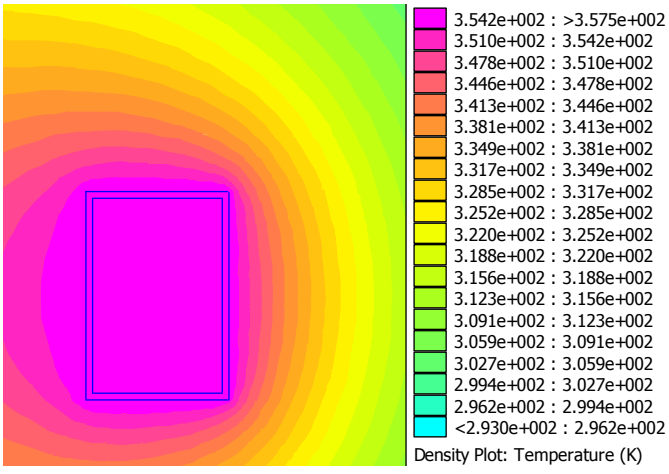


Figure 10. Temperature distribution

It may seem that the model shown in Fig. 10 is oversimplified, but with current level of design, without experimental verification, it would be unreliable to assume special insulation types for conductors and special wooden or metal platforms for the metal egg to spin on them which may influence the basic calculation.

V. CONCLUSIONS

In this paper a full design procedure for calculating the parameters of Tesla's two-phase inductor is explained. Starting from basic concepts of electrical machine design, a new method dedicated to special machines with large air gaps and toroidal inductors is developed. The main problem of the proposed design procedure is found to be the selection of magnetic flux density and current density values. While magnetic flux density, required to lift a small metal egg, is determined experimentally, the value for current density is selected as a result of detailed thermal analysis. Even such analysis implies complicated set of equations and is parameter sensitive, it is shown that for reasonable values of operating and ambient temperature, the current density is found to be in a relatively small range of values. It is shown that the analytical model gives very similar results as the FEM computer simulation. It should also be noted that the total copper losses of such a device are around 60 W which is equivalent to a light bulb. Even it seems that the larger values of current densities can be allowed without overheating the winding, the FEM simulation showed that with only 60 W of losses the temperature rise is close to given 80°C limit. The next step would be to build the prototype based on the proposed calculation. Only by experimental verification, the the design process will be completed.

REFERENCES

[1] Tesla universe, "Tesla's Egg of Columbus" [Online]. Available: http://www.teslauniverse.com/nikola-tesla-article-teslas-egg-of-columbus

[2] N. Tesla, "System of Electrical Distribution," U.S. Patent 0 381 970 May 1, 1888.

[3] RT17, "Tesla's Egg of Columbus–Production of working replicas" [Online]. Available: http://www.rt17.hr/teslas-egg-of-columbus/

[4] P. Matić, M. Milanković, G. Kondić, M. Radivojević, "An laboratory model for induction machine practice (in Serbian)," in Proc. Symposium Industrial Electronics INDEL 2004, Banja Luka, Republic of Srpska – Bosnia and Herzegovina, Sep. 11-12, 2004, Vol. V, pp. 65-68

[5] D. Jagodić, A. Ilišković, „Forming of rotary magnetic field (in Serbian)," in Proc. Symposium Industrial Electronics INDEL 1997, Banja Luka, Republic of Srpska – Bosnia and Herzegovina, Sep. 24-26, 1997, Vol. 1, pp. 154-158

[6] V. Petrović, *Uput u proračun asinhronog motora*, vol. I. Beograd Naučna knjiga, 1974, pp. 81-88

[7] B. Mitraković, *Asinhrone mašine,* vol. IV. Beograd Naučna knjiga, 1989, pp. 105

[8] V. Morgan, *Thermal behaviour of electrical conductors*, vol. I. Taunton Research Studies Press, 1991

Particle Swarm Optimization Based Energy Efficiency Method for High Speed IPMSM Drives

Marko A. Gecić, Mirna N. Kapetina, Vladimir M. Popović, Darko P. Marčetić
Faculty of Technical Science
Novi Sad, Serbia

Abstract— **This paper describe a novel method for increasing energy efficiency of interior permanent magnet synchronous motor (IPMSM) drives. In order to minimize the controllable electrical losses of IMPSM the *dq*-axes armature current is calculated based on particle swarm optimization (PSO). The method are tested for the wide speed range and different load condition. Simulation results of high speed IPMSM drives are presented and discussed.**

Keywords- IPMSM, high speed, PSO, energy efficiency

NOMENCLATURE

v_d ,v_q	– stator *d*- and *q*- axis input voltages
v_{od} ,v_{oq}	– stator *d*- and *q*- axis airgap voltages
i_d, i_q	– stator *d*- and *q*- axis input currents
i_{do}, i_{qo}	– stator *d*- and *q*- axis airgap currents
i_{dc}, i_{qc}	– stator *d*- and *q*- axis equivalent iron loss currents
R_s	– stator phase resistance
R_c	– iron losses resistance
Ψ_m	– permanent magnet flux
L_d ,L_q	– stator *d*- and *q*- axis self inductances
ω	– actual rotor angular speed
m_{el}	– electromagnetic torque
m_m	– load torque
J	– motor inertia
p	– number of pole pairs
ρ	– saliency ratio (L_q/L_d)
T	– torque
P_{Cu} ,P_{Fe}	– copper and iron losses
P_L	– total electric power losses

I. INTRODUCTION

The 60% of electric energy generated in industrialized countries is expended on electromechanical conversion. Losses occur during any such conversion, and optimization is required for their minimization [1]. Energy conservation and profitability are the basis of a faster development of digitally regulated electrical drives in a wide range of speeds, which are seeing more and more use in industrial processes.

Last decade the permanent magnet synchronous motor (PMSM) are widely used. PMSM combine a high power density, good heat transfer, and, lastly, a greater efficiency compared to other types of motors used for energy conversion. At the same time, synchronous motors boast favorable control characteristics, making them suitable for various applications, such as hybrid vehicles, servo-drives, household appliances etc [2]. The past few decades experienced rapid development of microcontrollers of vast capabilities, enabling full digital control of electromechanical conversions. Great effort has been expended in resolving problems of digital electrical drives. Synchronous motor energy efficiency can be further increased with improvements in digital control algorithms, while additional expenses can be reduced by increasing the rotor revelation speed [3], [4].

Drive losses are consisted of converter losses and motor losses. Motor losses are consisted of losses in stator windings, mechanical losses, and iron losses. The past several years saw the development of a number of loss optimization methods for regulated permanent magnet synchronous motor drives. These methods can be divided into two basic groups: methods based on search algorithms [1] - [3], and model-based methods [4] - [14]. The first group is independent on the motor model used, and includes inverter losses, but may, in some cases, cause ripples to appear in steady-state torque. The second group requires the knowledge of motor (as well as converter) parameters during drive operation.

In search algorithms, the input power is measured, and then minimized through alteration of system variables. A required property for optimization is constant output power. Search algorithms are most often used in steady-state operation, but, optionally, can be combined with model-based methods during transient states. Authors in [1] have, based on measured currents and DC circuit voltage, estimated input power, and use algorithms to determine the optimal *d*-axis current vector component for steady-state operation. An adaptive algorithm for on-line IPMSM loss optimization is presented in [2]. The algorithm functions in steady-state only. The authors of [3] present an algorithm suitable for scalar PMSM control in battery powered electric vehicle drives. Input power calculations utilize DC current and voltage. The minimal power is achieved through regulating output voltage.

Model-based algorithms require the modeling of motor and converter losses and utilization of those models during optimization. Parameters must be known, and in most cases necessitate the consideration of magnetic core saturation [4]. In [8], the authors utilize the stator flux vector as an independent variable both the torque equation and the voltage equations, and propose loss reduction through voltage angle correction. The proposed solution takes into consideration both voltage and current limitations, also expressed through the stator flux vector. In addition to loss optimization, flux estimation is necessary to improve the control algorithm dynamic performance [9]. The choice of optimal currents is enabled by look-up tables, generated off-line using various program suites.

This paper will consider the synchronous motor constructed in such a way that the induced electromotive force is sinusoidal, while the permanent magnets ate imprinted into the

This research was funded by the Ministry of Education, Science and Technological Develop. of Republic of Serbia under contract No. III 042004.

iron core. High magnetic anisotropy is a staple of such motor, that is to say, such construction greatly reduces the amount of iron in the d-axis, making the L_d inductance significantly lesser than L_q inductance (L_d<L_q). The optimization was performed by using particle swarm optimization (PSO) algorithm. PSO algorithm is a swarm intelligence optimization technique that has found its basis in natural, special in interactions of flocks of birds and swarms of insects. It was first introduced by Kennedy and Eberhart (1995) [15]. The proposed PSO algorithm attempts to reduce copper and iron losses both in constant field and field weakening areas of operation. Optimal currents are recorded in look-up tables. For a given speed and torque, optimal currents are read and used as input for current regulators. The proposed algorithm is compared a standard i_d= 0 control.

II. EQUIVALENT CIRCUIT AND BASIC EQUATION

Fig. 1 shows the *d*- and *q*-axis equivalent circuits of IPMSM.

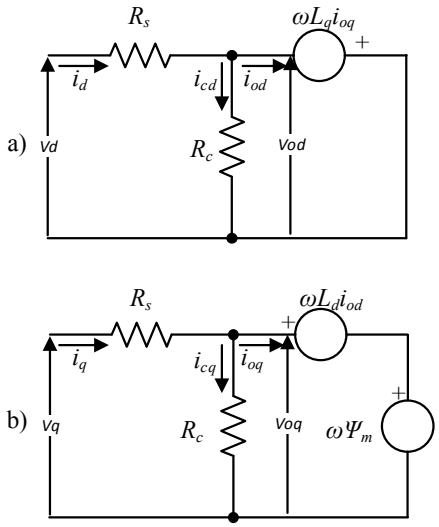


Figure 1. d- and q-axis equivalent circuits of IPMSM a) d-axis equivalent circuit, b) q-axis equivalent circuit

Based on Fig. 1 the mathematical equations of the equivalent *dq* axis model of IPMSM in the rotor reference frame are expressed as [7]:

$$\begin{bmatrix} v_d \\ v_q \end{bmatrix} = R_s \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \left(1 + \frac{R_s}{R_c}\right) \begin{bmatrix} v_{od} \\ v_{oq} \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} v_{od} \\ v_{oq} \end{bmatrix} = \begin{bmatrix} 0 & -\omega\rho L_d \\ \omega L_d & 0 \end{bmatrix} \begin{bmatrix} i_{od} \\ i_{oq} \end{bmatrix} + \begin{bmatrix} 0 \\ \omega\psi_m \end{bmatrix} \tag{2}$$

where,

$$i_{od} = i_d - i_{cd}, i_{oq} = i_q - i_{cd} \tag{3}$$

$$i_{od} = -\frac{\omega\rho L_d i_{oq}}{R_c}, i_{oq} = \frac{\omega(\Psi_m + L_d i_{od})}{R_c} \tag{4}$$

A. Torque production

The electromagnetic torque of the IPMSM has two components: fundamental magnetic torque (which is proportional to the product of the magnet flux and *q*-axis stator current), and the reluctance torque (which is dependent on the saliency ratio and to the product of *dq*-axis stator current components). Based on Fig.1 torque can be expressed as:

$$T = \frac{3}{2} p\left(\Psi_m i_{oq} + (1-\rho) L_d i_{od} i_{oq}\right) \tag{5}$$

B. Controllable losses

The copper losses are proportional to square of current and can be estimated using circuit in Fig.1:

$$P_{Cu} = \frac{3}{2} R_s \left(i_d^2 + i_q^2\right) \tag{6}$$

The iron losses in the machine consist of two components: hysteresis and eddy current losses. The entire no-load losses are assumed to be dominantly due to the iron losses and are modeled by a parallel resistance called R_c (which is function of speed) [4]:

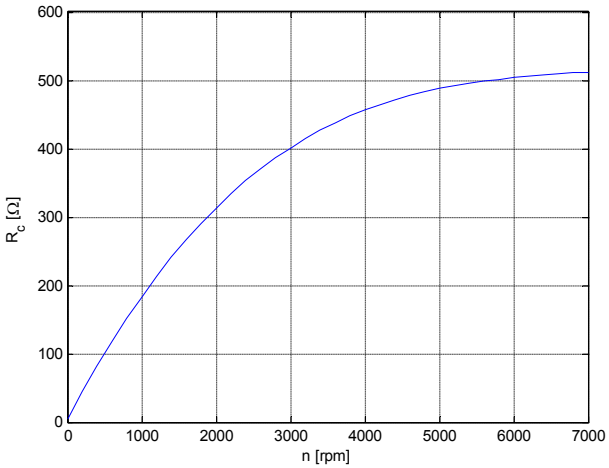


Figure 2. R_c estimation

The iron losses can be estimated using circuit in Fig.1:

$$P_{Fe} = \frac{3}{2} R_c \left(i_{cd}^2 + i_{cq}^2\right) \tag{7}$$

The mechanical losses are not controllable. The electrical losses are controllable by means of current vector control. The electrical losses are consisted of copper and iron losses:

$$P_L = P_{Cu} + P_{Fe} \tag{8}$$

The electrical losses can be expressed as function of i_{od}, T and ω. Minimal electrical losses can be derived by differentiating function of electrical losses with respect to i_{od} and equating the derivates to zero. For IPMSM there is no analytical solution [4].

$$P_L = f\left(i_{od}, T, \omega\right) \tag{9}$$

III. PARTICLE SWARM OPTIMIZATION

The PSO main goal is to explore the search space of interest using groups made of particles. A group of particles makes swarm, which is identified with a population in evolutionary terms. Each particle is characterized with its position which representing the potential solution of the optimization problem and velocity. Velocity is the difference between the current and previous positions. Particle remembers its best personal position in the history of the search, while swarm remembers best global position. The basic idea of the PSO algorithm is that the particles move guided by the personal and global best position through search space, while calculating a new value of velocity in each iteration. A new position of the particle is described by the following expressions:

$$v[k+1] = w \cdot v[k] + cp \cdot rp[k] \cdot (p[k] - x[k]) + cg \cdot rg[k] \cdot (g[k] - x[k]) \quad (10)$$

$$x[k+1] = x[k] + v[k+1] \quad (11)$$

The parameters w, cp and cg represent inertial, cognitive and social component. Their value is changed in order to improve performance which led to different modifications of the PSO algorithm [16] - [18]. In this analysis the Generalized PSO (GPSO) was used [18]. GPSO is inspired by linear control theory. The authors have identified particles swarm with dynamical system of second order with two inputs and one output and then analyzed its stability. The input represented by personal and global position of the particle, and the output of system is the current position of the particle.

IV. SIMULATION RESULTS

As described in Section II the electrical losses P_L are controllable by current vector control. If the electrical losses can be driven to a minimum value by the optimal current vector control, the efficiency becomes maximum. So, optimal problem is to find value of current components that minimalizes the loss.

Fig. 3 shows a block diagram of optimal model-based method for wide range speed control of IPMSM. Rotor speed and position are estimated or measured using a rotary encoder. Speed regulator output is the torque required for the requested speed. Based on the current speed and required torque, the model-based loss optimization block (PSO) generates the optimal reference currents that are further routed to an current regulated voltage source inverter (CRVSI).

Fig. 4 and 5 show the sum of controllable losses dependent on i_{od} current for a given speed and multiple various load torques. The diagram displays the existence of such i_{od} currents that will produce optimal controllable losses. As there is no analytical solution of equating the derivates of (9) to zero, a PSO is used to generate currents that will result in minimal losses.

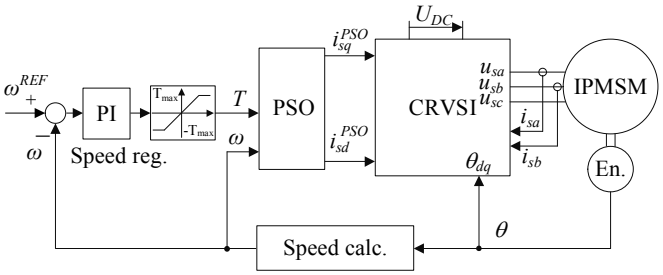


Figure 3. PSO based energy efficient IPMSM drive

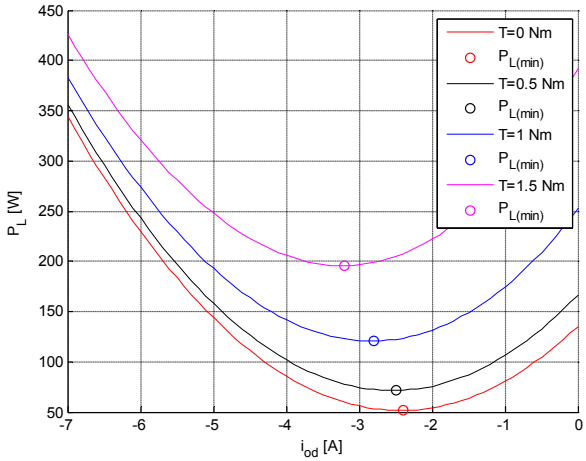


Figure 4. Electrical losses as function of i_{od} and T at 7000rpm

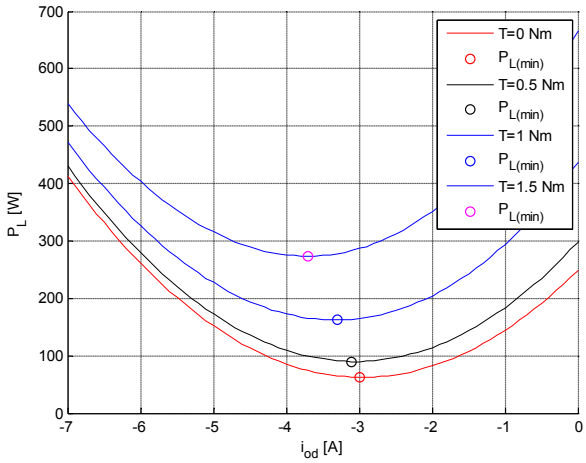


Figure 5. Electrical losses as function of i_{od} and T at 14000rpm

Optimal values of current for wide speed and torque range (speed: 0 - 14000rpm, with step 100rpm; torque: 0 - 1.5Nm, with step 0.1Nm) are found by using GPSO modification [18] of the typical PSO algorithm. One algorithm running find optimal current value for fixed speed and torque, and (9) is used as fitness (criteria) function of optimization algorithm. Number of iteration used in optimization process was 40 and number of the particles in populations was 25. Initial populations are random numbers on the interval (-10, 0). Algorithm is evaluated 2256 times, and the number of calculated optimal current values are the same. The flowchart of algorithm is shown in Fig 6.

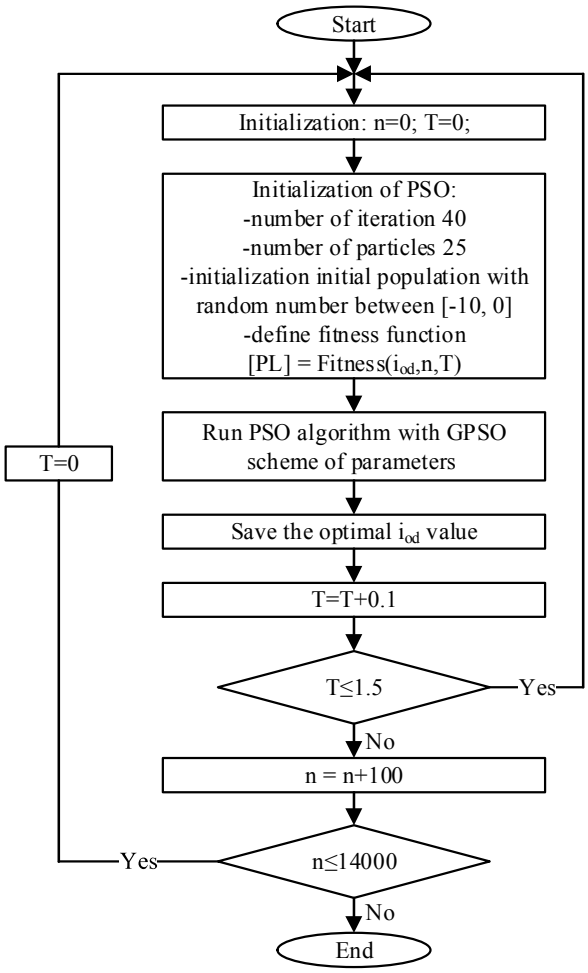


Figure 6. Optimal i_{od} for wide speed and load range

Calculated results presented optimal *d*-axis current component value for wide speed and torque and there are shown in Fig. 7.

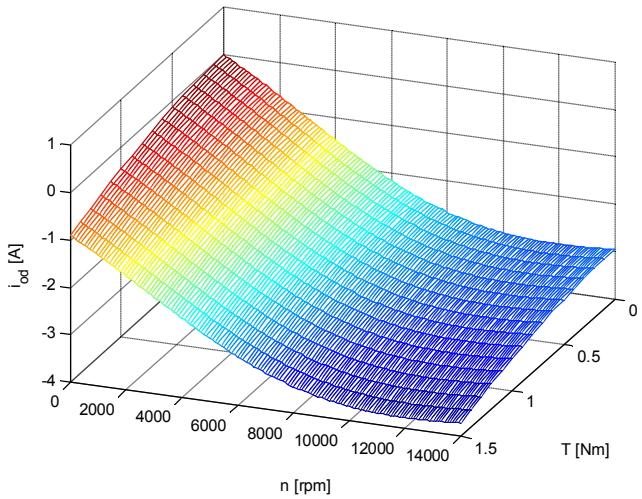


Figure 7. Optimal i_{od} for wide speed and load range

Fig. 8 shows electrical losses as a function of load at given speed for the two control types: the standard, i_{od} = 0 (blue, dashed), and the PSO algorithm (red, solid). The decreased controllable losses of the new algorithm can be noted. The existence of the negative d axis current causes a reduction in motor flux, and, thus, the iron losses. As such losses are exacerbated by high speeds, the PSO algorithm yields greater benefits in high speed operation.

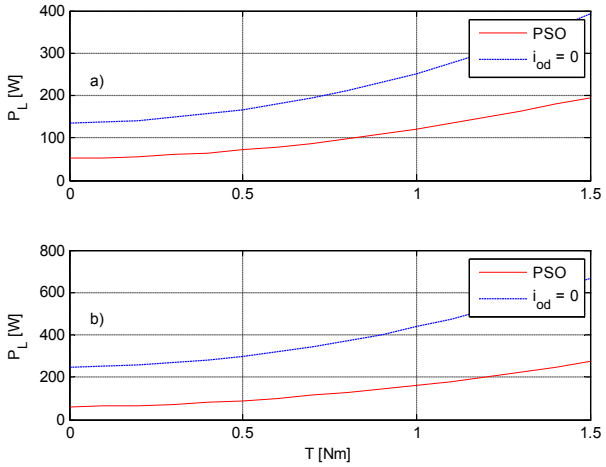


Figure 8. Electrical losses for wide load range at a) 7000 rpm, b) 14000 rpm

TABLE I. MOTOR PARAMETERS

Parametrs	IPMSM	
	Value	***Unit***
P	1	kW
Ψ_m	0.080074	Wb
L_d	20.33	mH
L_q	30.54	mH
R_s	3.575	Ω
p	4	-

Fig. 9 shows copper and iron losses ratio as a function of load for two different speed, 7000rpm (red, solid) and 14000rpm (blue, dashed). The PSO algorithm (Fig.9b) providing higher copper losses and lower iron losses, compared with standard i_{od} algorithm (Fig. 9a), but total losses are lower at the end.

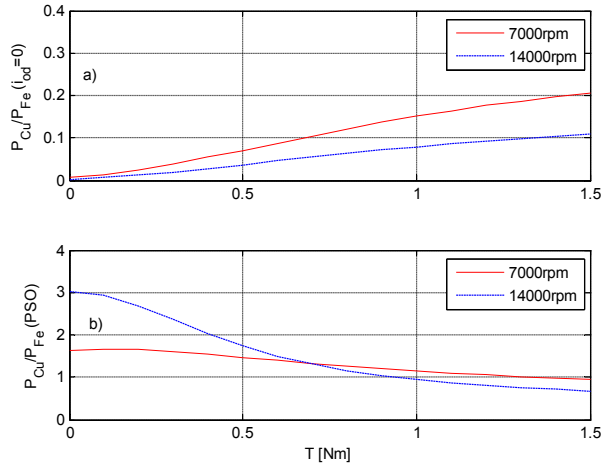


Figure 9. P_{Cu}/P_{Fe} ratio for wide load range at 7000 rpm and 14000 rpm: a) standard i_{od} = 0 algorithm, b) PSO algorithm

Fig. 10 shows electrical losses as a function load and speed for the i_{od} = 0 algorithm. Fig. 11 shows electrical losses as a function load and speed for the PSO algorithm. With increasing the speed at given load losses are increased as expected, but unlike the i_{od} = 0 control these electrical losses are optimal.

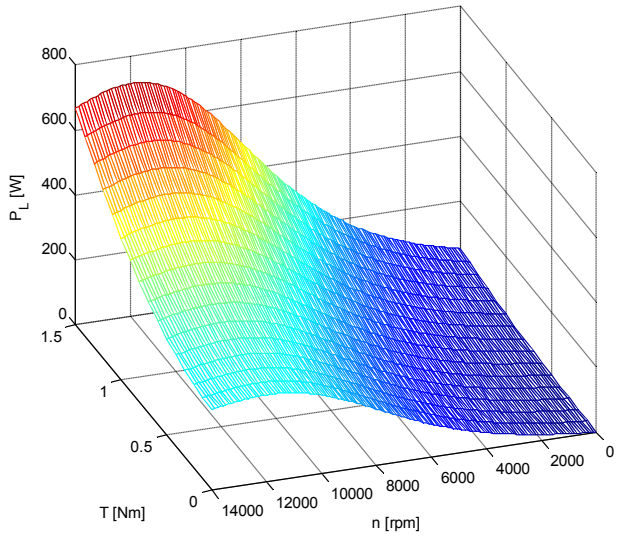


Figure 10. Electrical losses for wide speed and load range with i_{od}=0

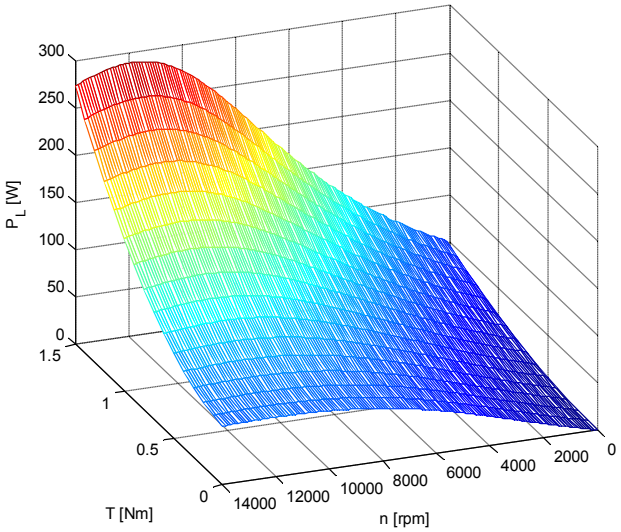


Figure 11. Optimal electrical losses for wide speed and load range

V. CONCLUSION

The paper proves that controllable electrical losses of IPMSM can be minimized by the optimal control of current vector angle. In particular, the proposed particle swarm optimization drastically minimizes controllable electrical losses. The lookup table is generated offline and can be easily implemented in existing IPMSM control algorithm. This method can be further improved by including *dq*-axis inductance dependences on current magnitudes during the lookup table creation. Generated lookup table can be implemented in the DSP-based digital control system with the enough memory space.

REFERENCES

[1] Z. Zhu and L. Gong, "Improved sensorless operation of permanent magnet brushless AC motors based on online optimal efficiency control," IEEE Int. Elec. Mach. Drives Conf. (IEMDC), vol., no., pp.1591–1596, May 15-18, 2011.

[2] S. Vaez, V. John and M. Rahman, "An on-line loss minimization controller for interior permanent magnet motor drives," IEEE Trans. Energy Conv., vol. 14, no. 4, pp. 1435–1440, December 1999.

[3] C. Chan and K. Chau, "An advanced permanent magnet motor drive system for battery-powered electric vehicles," IEEE Trans. Veh. Tech., vol. 45, no. 1, pp. 180–188, February 1996.

[4] M. Gecić, P. Matić, V. Katić, I. Krčmar, D. Marčetić, S. Cvetićanin, "Evaluation of Energy Efficiency of HighSpeed PMSM Drives," 17th International Symposium on Power Electronics – Ee 2013, Novi Sad, Serbia, October 30th- November 1st, 2013

[5] R. Bojoi, M. Pastorelli, J. Bottomley, P. Giangrande and C. Gerada, "Sensorless control of PM motor drives — A technology status review," IEEE Work. Elec. Mach. Design Contr. Diag. (WEMDCD), vol. 48, no. 4, pp. 168–182, March 11-12, 2013.

[6] K. Jang-Mok and S. Seung-Ki, "Speed control of interior permanent magnet synchronous motor drive for the flux weakening operation," IEEE Trans. Ind. Appl., vol. 33, no. 1, pp. 43–48, January/February 1997.

[7] S. Morimoto, Y. Tong, Y. Takeda, and T. Hirasa, "Loss minimization control of permanent magnet synchronous motor drives," IEEE Trans. Ind. Electron., vol. 41, no. 5, pp. 511–517, October 1994.

[8] J. Sung-Yoon, H. Jinseok and N. Kwanghee, "Copper loss minimizing torque control of IPMSM based on flux variables," IEEE Int. Elec. Mach. Drives Conf. (IEMDC), vol., no., pp. 1174–1179, May 12-15, 2013.

[9] B. Patel and M. Uddin, "Development of a nonlinear loss minimization control of an IPMSM drive with flux estimation," IEEE Int. Elec. Mach Drives Conf. (IEMDC), vol., no., pp. 1196–1203, May 12-15, 2013.

[10] H. Aorith, J. Wang and P. Lazari, "A new Loss Minimization Algorithm for Interior Permanent Magnet Synchronous Machine drives," IEEE Int. Elec. Mach Drives Conf. (IEMDC), vol., no., pp. 526–533, May 12-15, 2013.

[11] C. Botan, M. Ratoi, F. Ostafi and V. Horga, "Minimum energy control of servo drive systems with PMSM," Int. Symp. Pow. Electron., Elec.l Drives Aut. and Motion (SPEEDAM), vol., no., pp. 19–23, June 14-16, 2010.

[12] A. Rabiei, T. Thiringer and J. Lindberg, "Maximizing the energy efficiency of a PMSM for vehicular applications using an iron loss accounting optimization based on nonlinear programming," Int. Conf. Elec. Mach. (ICEM), vol., no., pp.1001–1007, September 2-5, 2012.

[13] D. Pohlenz and J. Bocker, "Efficiency improvement of an IPMSM using Maximum Efficiency operating strategy," 14th Int. Pow. Electr. Motion Con. Conf. (EPE/PEMC), vol., no., pp. T5-15–T5-19, September 6-8, 2010.

[14] L. Junggi, N. Kwanghee, C. Seoho, K. Soonwoo, "A Lookup Table Based Loss Minimizing Control for FCEV Permanent Magnet Synchronous Motors," IEEE Veh. Pow. Prop. Conf., (VPPC), vol., no., pp. 175–179, September 9-12, 2007.

[15] J. Eberhart and R. Kennedy: Particle Swarm Optimization, Proc. of IEEE Int. Conf. Neural Networks, pp. 1942–1948, 27 Nov.-1 Dec, Perth, Australia, 1995

[16] A. Ratnaweera, K.H. Saman and H.C. Watson, "Self-Organizing Hierarchical Particle Swarm Optimizer With Time-Varying Acceleration Coefficients," IEEE Transactions on Evolutionary Computation, pp. 240-2558 , 2004

[17] M. Rapaić and Ž. Kanović: "Time-varying PSO - convergence analysis,convergence related parameterization and new parameter adjustment schemes," Information Processing Letters, vol. 109, pp.548-552, 2009

[18] Z. Kanovic, M. Rapaic and Z. Jelicic, "Generalized particle swarm optimization algorithm - Theoretical and empirical analysis with application in fault detection, " Applied Mathematics and Computation, vol. 217, 2011

Evaluation of Luenberger Observer Based Sensorless Method for IM

Vladimir M. Popovic, Marko A. Gecic, Veran V. Vasic, Djura V. Oros, Darko P. Marcetic
Faculty of Technical Science
Novi Sad, Serbia

Abstract—**System approach for analysis of Luenberger observer for vector-controlled induction motor drives without shaft sensor is presented in this paper. Important aspects of control algorithm for induction motor system were described in detail. Realization of control algorithm based on Luenberger observer calculation mehod for estimation of unknown induction motor states is performed within appropriate digital signal controller (DSC). Verification is given through simulation and experimental results of vector–controlled induction motor sensorless drive.**

Keywords-vector control; Luenberger observer, induction motor

NOMENCLATURE

$\boldsymbol{u_s}$	– stator voltage vector
$\boldsymbol{i_s}$	– stator current vector
$\boldsymbol{\psi_r}$	– rotor flux vector
R_s	– stator phase resistance
R_r	– rotor phase resistance
L_m	– mutual inductance
L_r	– rotor self inductance
T_r	– rotor circuit time constant
L_σ	– equivalent stator circuit leakage inductance
ω_r [rpm]	– actual rotor angular speed
T_e	– electromagnetic torque
T_m	– load torque
J_m	– motor inertia coefficient
P	– number of motor pole pairs

I. INTRODUCTION

Various methods based on sensorless vector control (VU) of induction motor (IM) drives have been studied and proposed over past few decades, [1−6]. One of the very versatile techniques based on observer characteristic, [3−6], are in advantage to open-loop techniques, [1] and [2], because of feedback control law inherited within observation phenomena.

Digital signal controller (DSC) based IM vector–controlled drives without shaft sensor in very first period of their exploitation were exclusively used for low performance drive applications. Specified algorithms for estimation of unknown machine states are demanded in the case of VU *sensorless* drives and such algorithms could only be implemented in high-speed DSC platforms. Concept of VU for performance improvements of these drives also requires the use of power electronic devices with high frequency switching capabilities. Stability and robustness of vector−controlled *sensorless* drives are crucial attributes for quality drive operation. For the fulfillment of these conditions the use of quality and often expensive micro– and power–electronic devices is imposing as inevitable fact.

Today, with the grown–up trend of development and price reduction in area of electronics goes the increase in the use of *sensorless* VU drives. More complex and sophisticated control algorithms can be transferred to DSC and high performance characteristics of IM drives can be achieved. This enables the utilization of these drives in high performance drive applications. Lack of shaft sensor leads to decrease in cost and the increase in reliability and security of considered drives. Complicated installation of optical incremental encoder or resolver cables and connectors magnifies the risks of various faults. Besides, mounting of a shaft sensor is not applicable in many areas of practical usage, [9].

Global needs for power efficiency often impose the utilization of specified algorithms within control unit for better evaluation of IM states. One of the most common used algorithms for estimation and improvement in control performance of IM drive is based on mathematical model of IM called Luenberger observer (LO). In this paper, evaluation of LO based *sensorless* method for IM, theoretical introduction in system observability as well as the construction method for appropriate LO and DSC realization are considered. Validation is given through simulation and experimental results at various operational conditions.

II. INTRODUCTION TO OBSERVERS

The availability of entire state vector is essential for performance improvement of general system *S* governed by

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t). \tag{1}$$

Optimal control law implementation in form of

$$\boldsymbol{u}(t) = \varphi(\boldsymbol{x}(t), t), \tag{2}$$

can be constructed for relevant system *S* if its entire state vector is known e.g. through measurement. In practice this differs for the most of complex systems where some items within state vector remain unknown because of measurement disability nature of corresponded items. In that case, one of the most common approach for construction of control law deduced in form of (2) includes the estimation of unknown part of state vector in the manner of system observability phenomena.

System observability characteristic involves the fact of linear tracking the state vector of supervised system S_1 via state vector of second system S_2 which is, in the most cases, driven by available outputs of system S_1. Such driven system is called the observer to S_1 system. As a result, unknown part of the S_1 state vector can be derived for control law implementation or other *sensorless* control methods within general system.

This research was funded by the Ministry of Education, Science and Technological Develop. of Republic of Serbia under contract No. III 042004.

Here, free observation system theorem will be introduced for better understanding of a system observability characteristic. Further theoretical analysis will be restricted to linear time–invariant systems without of loss in generality.

Theorem 1: *Let S_1 be a free system, $\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t)$, which drives S_2 $\dot{\boldsymbol{z}}(t) = \boldsymbol{F}\boldsymbol{z}(t) + \boldsymbol{H}\boldsymbol{x}(t)$. Suppose there is a transformation $\boldsymbol{T}$ which satisfy $\boldsymbol{T}\boldsymbol{A} - \boldsymbol{F}\boldsymbol{T} = \boldsymbol{H}$. If $\boldsymbol{z}(0) = \boldsymbol{T}\boldsymbol{x}(0)$, then $\boldsymbol{z}(t) = \boldsymbol{T}\boldsymbol{x}(t)$ for all $t \geq 0$. Or more generally,*

$$\dot{\boldsymbol{z}}(t) = \boldsymbol{T}\boldsymbol{x}(t) + e^{\boldsymbol{F}t}[\boldsymbol{z}(0) - \boldsymbol{T}\boldsymbol{x}(0)]. \quad (3)$$

Proof: We may write immediately

$$\dot{\boldsymbol{z}}(t) - \boldsymbol{T}\dot{\boldsymbol{x}}(t) = \boldsymbol{F}\boldsymbol{z}(t) + \boldsymbol{H}\boldsymbol{x}(t) - \boldsymbol{T}\boldsymbol{A}\boldsymbol{x}(t). \quad (4)$$

Substituting $\boldsymbol{T}\boldsymbol{A} - \boldsymbol{F}\boldsymbol{T} = \boldsymbol{H}$ this becomes

$$\dot{\boldsymbol{z}}(t) - \boldsymbol{T}\dot{\boldsymbol{x}}(t) = \boldsymbol{F}[\boldsymbol{z}(t) - \boldsymbol{T}\boldsymbol{x}(t)] \quad (5)$$

which has (3) as a solution, [7].

One of the most convenient observer is one where transformation matrix $\boldsymbol{T}$ is the identity matrix $\boldsymbol{I}$ ($\boldsymbol{T} = \boldsymbol{I}$). This observer is called the identity observer because state vector of free system $\boldsymbol{x}(t)$ is identically related to the observer state vector $\boldsymbol{z}(t)$ as shown in the form

$$\boldsymbol{z}(t) = \boldsymbol{x}(t). \quad (6)$$

As a result, those systems have same dynamic order. Thus, specification of the observer rest on the specification of matrix $\boldsymbol{H}$. Again, if we consider observer system S_2 in altered form deduced as

$$\dot{\boldsymbol{z}}(t) = \boldsymbol{F}\boldsymbol{z}(t) + \boldsymbol{G}\boldsymbol{y}(t), \quad (7)$$

where $\boldsymbol{y}(t)$ is available (often through measurement) part of system state vector $\boldsymbol{x}(t)$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t), \quad (8)$$

with constant matrix $\boldsymbol{C}$, under the assumption of identity transformation employment within observation system, the same observation system can be written in form of

$$\dot{\boldsymbol{z}}(t) = (\boldsymbol{A} - \boldsymbol{G}\boldsymbol{C})\boldsymbol{z}(t) + \boldsymbol{G}\boldsymbol{y}(t). \quad (9)$$

Any matrix $\boldsymbol{G}$ leads to the observer system but the estimation dynamic of observer depends on the selection of matrix $\boldsymbol{G}$ coefficients. For stable observation of unknown system state, eigen–values of observer system matrix (poles of observer) are often allocated to be proportionally negative semi–definite according to roots of free system. Convergence dynamic within observation of a free system state vector is upgraded and as a result, state vector of system is available for the purposes of e.g. control improvements in a form of optimal control law (2), [7].

III. Mathematical Models of Induction Motor and Luneberger Observer

LO belongs into category of deterministic observers because it is based on mathematical model of considered system. Base goal in this section is to describe the construction of LO for unknown states and rotor speed observation of IM in a purpose of performance improvements of IM *sensorless* drive.

A. Mathematical model of induction motor

Mathematical model of IM in stationary reference $\alpha\beta$ frame with stator current and rotor flux vectors as IM state vector is suitable for VU concept development. It is presented in extended matrix form as

$$\begin{bmatrix} \dot{\boldsymbol{i}}_s \\ \dot{\boldsymbol{\psi}}_r \end{bmatrix} = \begin{bmatrix} -(R_s + \frac{R_r L_m^2}{L_r^2})/L_\sigma & -\left(\frac{L_m}{L_\sigma L_r}\right)(-1/T_r + j\omega_r) \\ L_m/T_r & (-1/T_r + j\omega_r) \end{bmatrix} \begin{bmatrix} \boldsymbol{i}_s \\ \boldsymbol{\psi}_r \end{bmatrix} + \frac{1}{L_\sigma}\begin{bmatrix} \boldsymbol{u}_s \\ \boldsymbol{0} \end{bmatrix}. \quad (10)$$

which has a simplified form in state space as presented in (1), [8].

Stator current vector is also a measurable vector and as a part of IM state vector it can be written in shortened matrix form as

$$\boldsymbol{i}_s(t) = \boldsymbol{C}\boldsymbol{x}(t), \quad (11)$$

where $\boldsymbol{C} = [\boldsymbol{I}\ \boldsymbol{0}]$ is $2x4$ dimension constant matrix.

Electromagnetic torque of IM in a form of vector equation is considered by

$$T_e = \frac{3}{2}P\frac{L_m}{L_r}(\boldsymbol{\psi}_r \times \boldsymbol{i}_s). \quad (12)$$

Complete mathematical model of IM is enclosed by adding the mechanical system differential equation which describes mechanical phenomena in IM. This is governed by

$$T_e - T_m = J_m \frac{d}{dt}\omega_m, \text{ [8].} \quad (13)$$

B. Mathematical model of Luenberger observer

According to the IM model attributes, complete model of LO for IM state observation in shortened matrix form is given as

$$\dot{\hat{\boldsymbol{x}}}(t) = \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{G}(\hat{\boldsymbol{i}}_s - \boldsymbol{i}_s), \quad (14)$$

where variables with ^ represents observed variables from LO and matrices $\hat{\boldsymbol{A}}$ and $\boldsymbol{B}$ represents LO system and LO control matrices with $4x4$ and $4x2$ dimension rate, respectively. It should be advised that, in general, those matrices do not have to fit in original to corresponding IM system matrices.

It is shown in section II that the construction of LO observer when the system model is known rest on definition of matrix $\boldsymbol{G}$, so called feedback action matrix. This matrix is multiplied with the stator current error vector and in summary, this has affect in the decrease in estimation error due to the mismatch in rotor speed or parameters values within IM and LO models. It is given in a form of

$$\boldsymbol{G} = \begin{bmatrix} g_1\boldsymbol{I} + g_2\boldsymbol{J} \\ g_3\boldsymbol{I} + g_4\boldsymbol{J} \end{bmatrix}, \boldsymbol{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{J} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (15)$$

where g_1, g_2, g_3 and g_4 are coefficients of matrix $\boldsymbol{G}$.

In practice, the most common approach for designing the observer is to place the eigen–values of LO system matrix $\boldsymbol{\lambda_0}$ in proportional negative semi–definite correlation due to the roots of IM system matrix $\boldsymbol{\lambda}$ where IM is a stable system for itself. In [5] it is proven that gain selection of matrix $\boldsymbol{G}$ which constitutes the observer system matrix has an influence on convergence dynamic of IM states. For stable and more robust estimation, matrix gains are chosen in a way of

$$\boldsymbol{\lambda_0} = k\boldsymbol{\lambda}, \tag{16}$$

where k is coefficient of proportionality of IM and LO system poles.

Due to the previous, matrix $\boldsymbol{G}$ gains are derived as

$$g_1 = -(k-1)(\alpha + 1/T_r) \tag{17}$$

$$g_2 = (k-1)\hat{\omega}_r \tag{18}$$

$$g_3 = -(k^2-1)(\alpha/\beta + \gamma) + 1/\beta(k-1)(\alpha + 1/T_r) \tag{19}$$

$$g_4 = -1/\beta \cdot (k-1)\hat{\omega}_r \tag{20}$$

where $\alpha = (R_s + L_m^2/(L_r T_r))/L_\sigma$, $\beta = L_m/(L_r L_\sigma)$ and $\gamma = L_m/T_r$, [5].

Classical version of LO for IM drive assumes that the rotor speed is a variable parameter. In the essence of *sensorless* IM drives lies the Lyapunov function candidate governed by

$$V(\boldsymbol{e}) = \boldsymbol{e}^T\boldsymbol{e} + \frac{(\hat{\omega}_r - \omega_r)^2}{\lambda}, \tag{21}$$

where $\boldsymbol{e}$ represent error estimation vector and λ is a positive constant, [5].

If LO system matrix eigen–values are negative semi–definite the proposed estimation algorithm will be asymptotically stable. Based on the equalization of first derivative of this criterion function with zero

$$\frac{dV(\boldsymbol{e})}{dt} = \boldsymbol{e}^T[(\boldsymbol{A}+\boldsymbol{GC})^T + (\boldsymbol{A}+\boldsymbol{GC})]\boldsymbol{e} - \hat{\boldsymbol{x}}^T\Delta\boldsymbol{A}^T\boldsymbol{e} - \boldsymbol{e}^T\Delta\boldsymbol{A}\hat{\boldsymbol{x}} + \frac{2\Delta\omega}{\lambda}\frac{d\hat{\omega}_r}{dt} = 0, \tag{22}$$

proposed speed observation is maintained when the sum of second, third and fourth term of the right side of (22) equals to zero.

Solution of previous determines the adaptive mechanism for rotor speed estimation governed by

$$\hat{\omega}_r = \int \frac{\lambda}{c}(\hat{\psi}_{\beta r}e_{i\alpha s} - \hat{\psi}_{\alpha r}e_{i\beta s})dt \tag{23}$$

where $c = L_\sigma L_r/L_m$, [5].

Thus, IM system is linearized and for better dynamic response of observed rotor speed, proportional term is added so block diagram of adaptive mechanism for evaluation of rotor speed has ultimate form presented in Fig. 1.

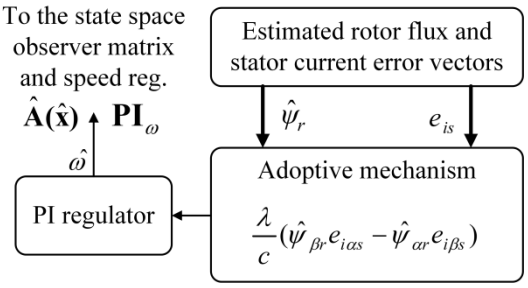


Figure 1. Adaptive mechanism for rotor speed estimation of IM

Finally, block diagram of complete LO for observation of unknown states of IM and estimation of rotor speed is represented in Fig. 2.

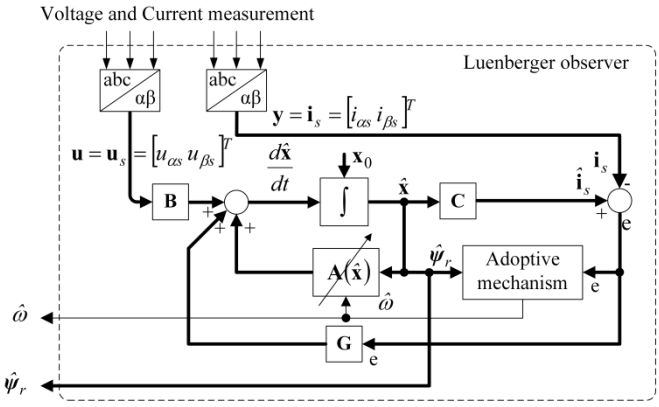


Figure 2. Block diagram of proposed LO for observation of unknown IM states and rotor speed

IV. DSP IMPLEMENTATION OF LUENBERGER OBSERVER FOR INDUCTION MOTOR

VU concept of IM control and various complicated algorithms for estimation and observation within IM drives requires exclusively DSC–based systems for quality control implementation. Those systems implement the algorithm law in discrete–time sequences. Due to that, it is of crucial importance to discretize the continuous model (14) in the meaning of magnitude and time discretization methodology. Also, *real time* applications requires execution of control law in limited–time sequences so model normalization is required for the purposes of utilization of *fixed–point* calculation method that saves, in the most occasions critical, CPU time.

First step for DSC realization of LO algorithm involves the method for normalization of state and control variables of IM. In case of a DSC implemented control of IM it is the most common practice to choose base state values to have maximum values which can appear in operation mode of IM drive. Motive for this comes from the fact that the problem of DSC registers saturation is then overcame.

Next step involves time discretization where the most common approach is based on approximation method which implies the substitution of complex s variable with complex z variable of relevant system represented in complex plane. For simplicity, it is often used *Forward Euler* approximation of complex s variable with z variable in the form of

$$s \to (z-1)/T \tag{24}$$

where T represents sample period which matches the PWM rate period. In the case of simulation and experiments of considered drive described in the next section, value of sample period correspond to the PWM frequency of 8 kHz. Therefore

$$T = T_{pwm} = 125\ \mu s. \tag{25}$$

Third and final step of adapting the model (14) for DSC realization is based on magnitude discretization problem. Analog–to–Digital DSC unit automatically converts analog signal to digital.

Applying these three steps into LO continuous model (14) in complex plane (time derivation of state vector is substituted with complex variable s), discrete equations are obtained like:

$$\boldsymbol{i_s^*}(kT) = A_{r11}\boldsymbol{i_s^*}(kT-T) + (A_{r12} - jA_{i12}\omega_r^*)\boldsymbol{\psi_r^*}(kT-T) + B_r\boldsymbol{u_s^*}(kT-T) + T(g_1^* + \boldsymbol{j}g_2^*)(\widehat{\boldsymbol{i_s^*}}(kT-T) - \boldsymbol{i_s^*}(kT-T)) \quad (26)$$

$$\boldsymbol{\psi_r^*}(kT) = [A_{r22} + jA_{i22}\omega_r^*]\boldsymbol{\psi_r^*}(kT-T) + A_{r21}\boldsymbol{i_s^*}(kT-T) + T(g_3^* + \boldsymbol{j}g_4^*)(\widehat{\boldsymbol{i_s^*}}(kT-T) - \boldsymbol{i_s^*}(kT-T)) \quad (27)$$

where $A_{r11} = 1 - \omega_b T\left(\frac{r_s}{l_\sigma} + \frac{l_m^2}{l_r^2}\frac{r_r}{l_\sigma}\right)$, $A_{r12} = \omega_b T\frac{l_m}{l_r}\frac{r_r}{l_r}\frac{1}{l_\sigma}$, $A_{i12} = \omega_{b1}T\frac{l_m}{l_r}\frac{1}{l_\sigma}$, $A_{r21} = \omega_b T r_r\frac{l_m}{l_r}$, and $A_{r22} = 1 - \omega_b T\frac{r_r}{l_r}$.

For calculation of (26) and (27) it is necessary to determine the values of motor parameters. One of the basic ways for extracting the parameter values from IM is through IM experiment at no–load and short–circuit state, [2].

In Table I parameter values of used IM are shown.

TABLE I. IM PARAMETER VALUES

R_S [Ω]	R_r [Ω]	L_s [H]	L_r [H]	L_m [H]
3.26	1.05	0.078	0.078	0.074

V. SYSTEM CONFIGURATION

Whirlpool three–phase IM (Maytag Whirlpool Factory Washer Motor W10171902 J58GTC–1132) is controlled via DSC MC56F8245 of FreeScale manufacturer which generates pulses with variable filling factor (PWM–pulse width modulated) to "GATE DRIVER" block that drives inverter. Vector control algorithm is implemented in DSC module. Measured variables are stator current from shunt in inverter DC circuitry and rotor shaft speed from tacho generator. In case of LO, reconstructed stator currents are used for feedback action and utilized in adaptive mechanism for estimation of rotor speed. In that case, measured rotor speed serves only for system monitoring purposes. Interface between the user and drive system is, in the manner of programmability and control management sense, implemented through serial link. Fig. 3 represents simplified block diagram of digital VU IM sensorless drive.

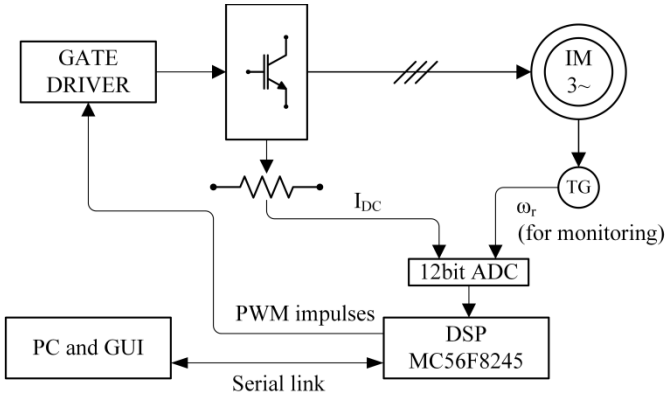


Figure 3. Simplified block diagram of IM drive

VI. SIMULATION AND EXPERIMENT

A. Simulation analysis of induction motor drive

Simulation analysis of LO algorithm performance within considered drive from Fig. 3. were implemented in *Simulink tool* of MATLAB software. Estimation results within basic observer modes such as open–loop and closed–loop modes were presented as well as the influence in discrete calculation methodology. Simulation analysis were performed at no–load condition of considered IM.

Observation responses of IM and continuous LO states in open–loop mode (pure rotor flux observer where rotor speed is measured and used from sensor) are presented in Figs. 4 and 5 at rotor speed references of 500 rpm and 2500 rpm.

From Figs. 4 and 5 it is noticeable the match of corresponding $\alpha\beta$ components of IM and LO states. From subplots in Figs. 4 and 5, blue– (correspond to $i_{s\alpha}^*$ from upper and $\psi_{s\alpha}^*$ from lower subplot) and green– (correspond to $i_{s\beta}^*$ from upper and $\psi_{s\beta}^*$ from lower subplot) colored graph components coincide the corresponding red– (correspond to $i_{s\alpha}$ from upper and $\psi_{s\alpha}$ from lower subplot) and yellow– (correspond to $i_{s\beta}$ from upper and $\psi_{s\beta}$ from lower subplot) dotted graph components. States with * represents observed states from LO.

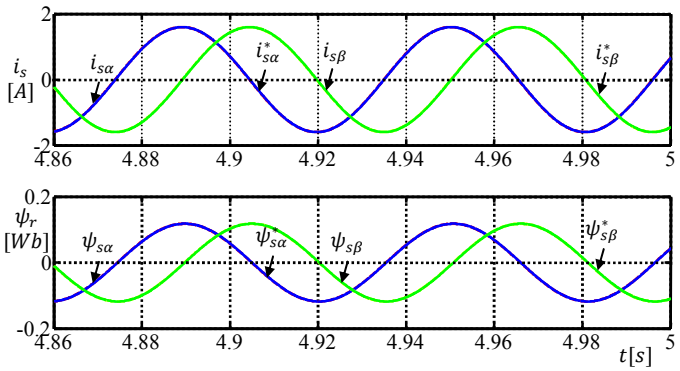


Figure 4. Current (upper subplot) and flux (lower subplot) responses of IM and continuous open–loop LO in $\alpha\beta$ frame, speed reference 500 rpm

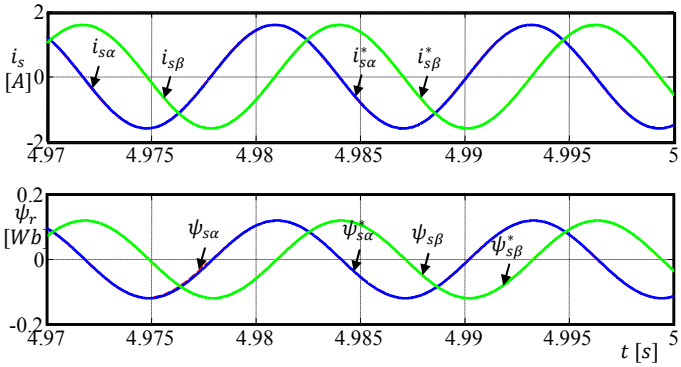


Figure 5. Current (upper subplot) and flux (lower subplot) responses of IM and continuous open–loop LO in $\alpha\beta$ frame, speed reference 2500 rpm

Observation responses of IM and discrete LO states in open–loop mode (rotor flux and speed observer where rotor speed is evaluated from adaptive mechanism, Fig. 1) are shown

in Figs. 6 and 7 at rotor speed references of 500 rpm and 2500 rpm, respectively.

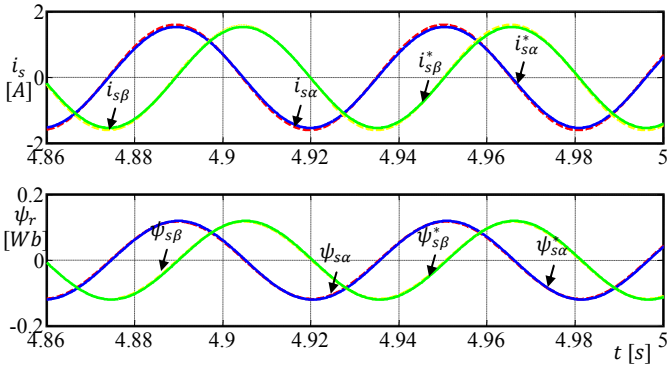


Figure 6. Current (upper subplot) and flux (lower subplot) responses of IM and discrete open–loop LO in $\alpha\beta$ frame, speed reference 500 rpm

Comparison of responses from Figs. 6 and 7 to the responses from Figs. 4 and 5 at identical simulation conditions shows bad influence of discrete calculation approximation method based on (24) to the observation of IM states. Proper magnitude as well as the phase loss of LO states particularly noticeable at higher speed reference values are repercussions from discrete equation utilization (26)–(27) within LO algorithm.

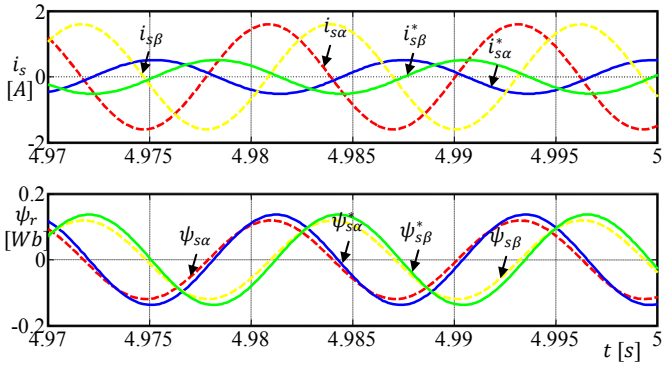


Figure 7. Current (upper subplot) and flux (lower subplot) responses of IM and discrete open–loop LO in $\alpha\beta$ frame, speed reference 2500 rpm

Responses of discrete LO states at closed–loop mode are similar to those at open–loop mode.

One of the solution for fixing the problem of magnitude and phase loss of LO state includes the use of feedback action within LO via stator current vector error.

Feedback action within LO inquires the use of non–zero gains of feedback action matrix G. If coefficient k from (16) is set on the value which differs from one than feedback is included in LO algorithm. It is often convenient to select this value to be greater than one mainly because the observation response converges faster to the stationary state of observation system.

Figs. 8 and 9 represents the observation responses of IM and discrete LO states in open– and closed–loop modes of LO algorithm with implemented feedback action at rotor speed reference of 2500 rpm, respectively.

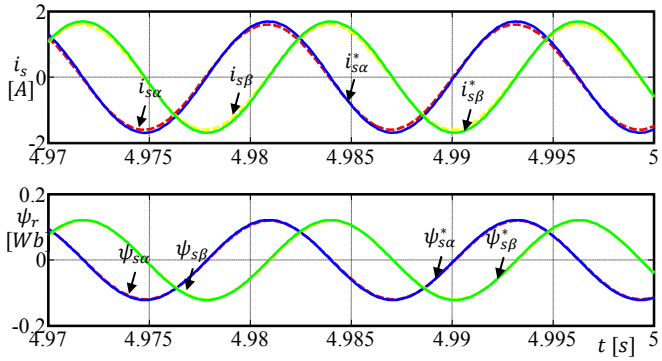


Figure 8. Current (upper subplot) and flux (lower subplot) responses of IM and discrete open–loop LO in $\alpha\beta$ frame with implemented feedback action $k = 3$, speed reference 2500 rpm

It is obvious from previous Fig. that the implementation of feedback within LO algorithm greatly reduces the magnitude and phase error between IM and observed LO states. In closed–loop mode coefficient k from (16) is set at value smaller than one, $k = 0.5$, because of stability issue disruption problem.

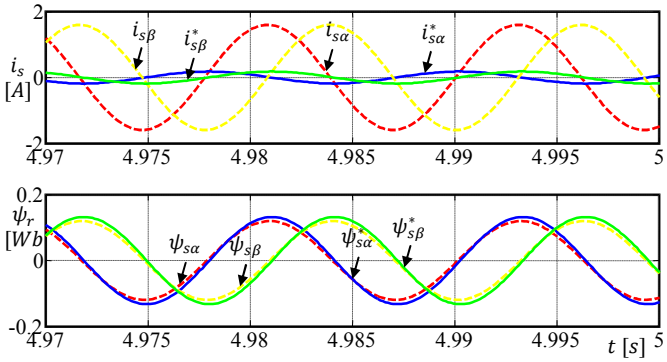


Figure 9. Current (upper subplot) and flux (lower subplot) responses of IM and discrete closed–loop LO in $\alpha\beta$ frame with implemented feedback action $k = 0.5$, speed reference 2500 rpm

Compare to the Fig. 7, it is clear from Fig. 9 that feedback action decrease the estimation error of observable states due to IM states but, compare to the Fig. 8 this correction has much smaller effect. Main reason lies in fact that feedback action matrix gains (17)−(20) are rotor speed–dependent. Rotor speed signal is estimated from adaptive mechanism (23). Discrete LO calculation method based on LE approximation has an inherent problem due to the appearance of observation error. Such erroneous states are utilized in adaptive mechanism and the aftermath is the inaccuracy of estimated speed signal (see Fig. 10.) and error accumulation problem in observation algorithm. Drawback of gain selection method based at (17)−(20) in closed–loop mode of LO is that the observer feedback action cannot drive the observation error to zero which affects the dynamic performance response and also stability in *sensorless* drives especially at high speed region.

Fig. 10 represents the rotor speed response from IM and adaptive mechanism implemented within LO (upper subplot) and also zoomed to the steady state condition (lower subplot), rotor speed reference value 2500 rpm. PI regulator gains

within adaptive mechanism were set at values $k_p = 30$ and $k_i = 60$.

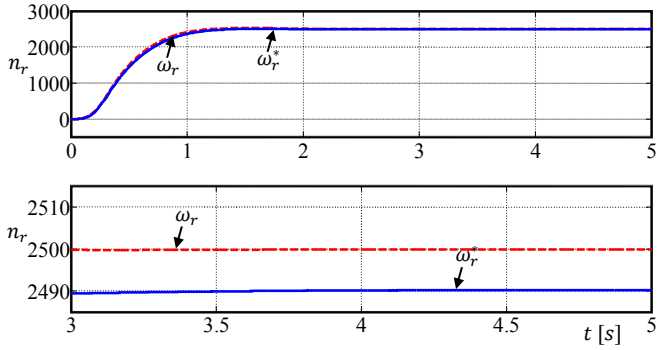


Figure 10. Speed responses from IM and adaptive mechanism within LO (upper subplot) and zoomed to the steady state conditions (lower subplot), speed reference 2500 rpm

B. *Experimental analysis of induction motor drive*

Validation of examined simulation results is given through appropriate experimental tests on subscribed IM drive, section V. Analysis are carried at identical no–load conditions of IM drive like in case of simulation tests.

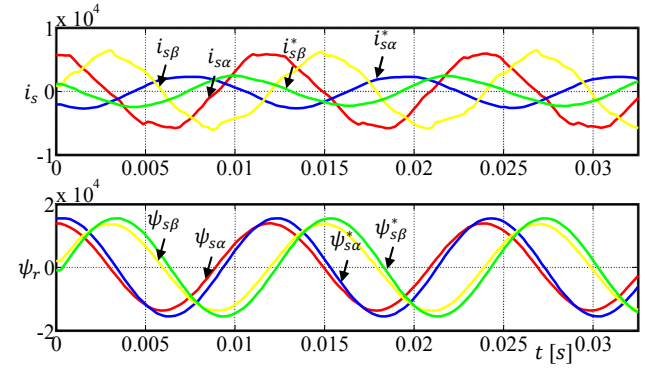


Figure 11. Current (upper subplot) and flux (lower subplot) responses of IM and discrete closed–loop LO in $\alpha\beta$ frame, speed reference 2500 rpm

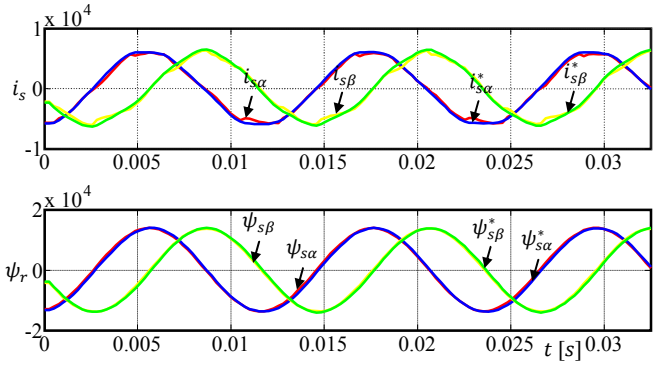


Figure 12. Current (upper subplot) and flux (lower subplot) responses of IM and discrete open–loop LO in $\alpha\beta$ frame, speed reference 2500 rpm

Fig. 11 shows the experimental responses of stator current and rotor flux of IM and discrete closed–loop LO at speed reference 2500 rpm. It should be advised that IM flux is derived from MRAS estimator which is also performed in specified software for control management of considered drive. States with * represents observed states from LO. All variables from Fig. 11 are in relative domain and represented as fixed–point variables like in DSP registers.

The similar behavior at identical operating conditions is obvious from Figs. 7 and 11. The same statement valid at open loop mode of LO with feedback action $k = 3$. Fig. 12 confirms that conclusion.

Feedback action in closed–loop mode was not implemented because of critical stability issue of considered system.

VII. CONCLUSION

Performance evaluation of LO based *sensorless* method for IM is introduced in this paper. It is shown that discrete calculation method based on (24) and performed within LO algorithm has dominant effect on the accuracy of IM state and rotor speed observation. Drawback of the LO model approximate implementation algorithm (26)–(27) is the occurrence of degradation in both magnitude and phase of LO states to the IM states with greater effect at high speed region. Proposed method which includes the use of feedback law within LO is proven to be suitable in open–loop mode Therefore, LO can be used as robust flux observer for IM drive performance improvements because it cancels bad effect of loss in magnitude and phase of observed LO states due to the IM states. In closed–loop mode where rotor speed is estimated in addition to the observed states, it is proven that feedback action did not cancel this effect. Actually small improvements in observation are conducted in that case and that is the main reason why this construction of LO is not suitable for appliance in high performance IM drives at wide speed range. Finally, the significant similarity is shown between simulation and experiment tests of considered IM drive.

REFERENCES

[1] J. Holtz, "Sensorless control of induction machines-with or without signal injection," IEEE Trans. Ind. Electron., Vol. 53, No. 1, pp. 7–30, February 2006.

[2] C. Schauder. "Adaptive speed identification for vector control of induction motors without rotational transducers," IEEE Trans. Ind. Appl., Vol. 28, No. 5, pp. 1054–1061, September/October 1992.

[3] J. Maes and J. Melkebeek, "Speed sensorless direct torque control of induction motors using an adaptive flux observer," IEEE Trans. Ind. Appl., Vol. 36, No. 3, pp. 778–785, January 2000.

[4] H. Rehman, A. Derdiyok, M. K. Güven and L. Xu, "A new current model flux observer for wide speed range sensorless control of an induction machine," IEEE Trans. Power Electron., Vol. 17, No. 6, pp. 1041–1048, June 2002.

[5] H. Kubota, K. Matsuse and T. Nakano, "DSP-based speed adaptive flux observer for induction motor," IEEE Trans. Ind. Appl., Vol. 29, No. 2, pp. 344–348, March/April 1993.

[6] Z. Yongchang and Z. Zhengming, "Speed Sensorless Control for Three-Level Inverter-Fed Induction Motors Using an Extended Luenberger Observer," IEEE Veh. Pow. Prop. Conf. (VPPC), September 3-5, 2008, Harbin, China

[7] D. G. Luenberger, "An introduction to observers," IEEE Trans. Aut. Cont., Vol. AC–16, No. 6, pp. 596–602, December 1971.

[8] S. N. Vukosavic, "Elektricne masine," Elektrotehnicki fakultet, Beograd, 2010.

[9] D. P. Marcetic, "Mikroprocesorsko upravljanje energetskim pretvaracima", FTN Novi Sad, 2012.

IPMSM Inductances Calculation Using FEA

Dejan G. Jerkan, Marko A. Gecić and Darko P. Marčetić
Department for Power, Electronic and Telecommunication Engineering
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
dejan.jerkan@uns.ac.rs

***Abstract* – Accurate determination of interior permanent magnet synchronous machine's (IPMSM) inductances is very important issue, especially in areas of high-performance drives and systems. This paper presents the method for calculation of the direct and quadrature inductances of permanent magnet synchronous machine using finite element analysis (FEA), where the calculation of these parameters is based on the determination of flux linkages. Two types of IPMSMs are investigated, with tangentially and radially magnetized permanent magnets. The results of the calculated inductances are presented by diagrams and they are discussed and compared with those obtained by measurements.**

Keywords – Finite element analysis, direct and quadrature axis inductances, magnetic flux density, interior permanent magnet synchronous machine

NOMENCLATURE

v_d ,v_q	– stator *d*- and *q*- axis voltages
i_d, i_q	– stator *d*- and *q*- axis currents
$i_{a,b,c}$	– stator phase currents
R_s	– stator phase resistance
Ψ_m	– permanent magnet flux
L_d ,L_q	– stator *d*- and *q*- axis self inductances
Ψ_{ds}	– stator *d*-axis flux
Ψ_{qs}	– stator *q*-axis flux
ω	– actual rotor angular speed
m_{el}	– electromagnetic torque
m_m	– load torque
J	– motor inertia
p	– number of pole pairs
ρ	– saliency ratio (L_q/L_d)
J_s	– current density vector
A_z	– *z* component of the magnetic vector potential
μ	– permeability of material

I. INTRODUCTION

The interior permanent magnet synchronous motors (IPMSM) have many advantages, such as high power density and possibility for speed regulation in wide range of speeds [1], [7], [9]. IPMSMs are widely used in high-performance drives such as industrial robots and high-performance machine tools because of their advantages on high-torque with additional reluctant component. In recent years, the magnetic and thermal capabilities of the PM have been enhanced by employing permanent magnets with high coercitivity [1]. IPMSMs are used in more and more applications because of their small volume, very good efficiency, lower moment of inertia, rotor without heat problem, etc. [2]. Because of demands of high-performance drives it is very important to calculate as accurate as possible the values of the parameters of the IPMSM. Of the most important significance are the direct- and the quadrature- axis inductances, as they are determining corresponding synchronous reactances [3]. Also, they are the most important parameters when steady state and dynamic models of IPMSM are developed [4]. Unlike surface PM motors, which have the same value of inductance in direct and quadrature axes and where all the torque is produced by the magnet flux, interior permanent magnet motors have different direct and quadrature inductances which results in an additional torque component called reluctance torque [5]. The conventional methods of testing for determination of synchronous machine parameters are often inappropriate in the case of permanent magnet machines, because magnetic field produced by the permanent magnets cannot be canceled during measurements, and its field affects the total level of saturation of iron during experiments. On the other hand, finite element method provides great opportunities for accurate numerical analysis of IPMSM, because using FEA algorithms allows calculations of fundamental field quantities (such as flux linkages and stored magnetic energy), and also the fields produced by PMs can be canceled easily [6]. Standard experiments are reconstructing parameters of the machine based on quantities which can be measured through electrical connection of the machine, without knowledge of field distribution inside of the machine. During the last two decades the finite element method proved to be the most appropriate numerical method in terms of modeling, flexibility and accuracy to solve the nonlinear Poisson's equation governing the magnetic field who's concerned a principal element in calculation of machine parameters [6].

II. EQUIVALENT CIRCUIT AND BASIC EQUATIONS

Fig. 1 shows the *d*- and *q*-axis equivalent circuits of IPMSM in which magnetic losses due to variable magnetic field in stator core are neglected. Based on Fig. 1 the mathematical equations of the equivalent *dq* axis steady state model of IPMSM in the rotor reference frame are given with:

$$\begin{bmatrix} v_d \\ v_q \end{bmatrix} = R_s \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} 0 & -\omega L_q \\ \omega L_d & 0 \end{bmatrix} \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} 0 \\ \omega \psi_m \end{bmatrix} \tag{1}$$

The electromagnetic torque of the IPMSM has two components: fundamental magnetic torque (which is proportional to the product of the magnet flux and *q*-axis stator current), and the reluctance torque (which is dependent on the saliency ratio and to the product of *dq*-axis stator current components). It is essential to determine IPMSM inductances to predict reluctance torque as an additional torque component. Based on Fig.1 torque can be expressed as:

$$T = \frac{3}{2} p\left(\Psi_m i_q + (1-\rho) L_d i_d i_q\right) \tag{2}$$

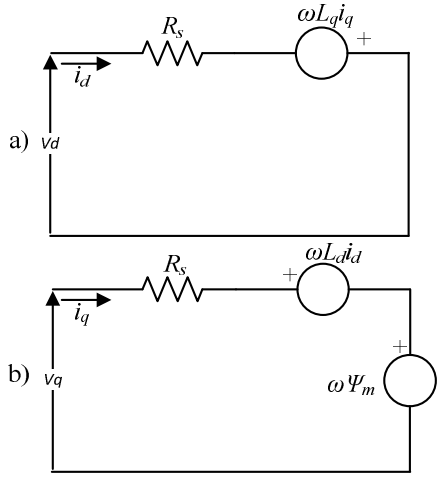


Figure 1. d- and q-axis equivalent circuits of IPMSM a) d-axis equivalent circuit, b) q-axis equivalent circuit

III. FINITE ELEMENT ANALYSIS OF IPMSM

The finite element analysis is used in many areas of technical sciences, such as magnetics, electrostatic problems, heat transfer, fluid dynamics etc. All FEA algorithms are based on solution of field equations over domain of interest using division of that domain with small segments of simple geometric shapes, called finite elements (in 2-D FEA the most common are triangular shapes), in order to reconstruct the field of the entire domain. Values of field variables inside any of the elements are represented using 2-D interpolation functions [6]. These functions are defined on each element using the values of the calculated variable in each node. Knowing the value of variable of interest in every node of the region, combined with the usage of interpolation functions allows complete definition of the behavior of the variable field on each element. The precision of the method depends not only on the dimensions of elements and their number but also on the type of the interpolation function. As for the numerical method, the FEA algorithm converges to the exact solution provided to increase the number of subdivisions of the solution domain and to ensure continuity of the interpolation function of its first derivatives along the borders of adjacent elements [6]. FEA algorithms used in the problems of electrical machines analysis are based on the evaluation of the magnetic vector potential A. As already mentioned, 2-D FEA algorithm will be used in this research, and the starting equation for FEA is given by:

$$-\frac{\delta}{\delta x}\left(\frac{1}{\mu}\left(\frac{\delta A_z}{\delta x}\right)\right) - \frac{\delta}{\delta y}\left(\frac{1}{\mu}\left(\frac{\delta A_z}{\delta y}\right)\right) = J_S \tag{3}$$

Equation (3) is used for magneto-static solutions, so there is no time varying of currents or motion of rotor.

Adequate usage of FEA algorithms demands definition of proper boundary conditions, which determine how lines of magnetic potential vector pass through areas that separate regions with different magnetic properties. Analyzed machines have small area of 2-D cross-section (please see Table 1 for machine's dimensions), so only one boundary condition needs to be defined for this type of FEA, so-called *Dirichlet* boundary condition, and the most common use of it is to define A_z=0 along outer stator surface. This means that with this boundary condition we force the magnetic field to stay inside boundary defined by stator outer diameter. For larger geometries it is useful to define additional boundary conditions which allow usage of only one slice of machine's geometry, and the width of that slice is defined by pole width, which means that for large machines with large number of pole pairs these boundary conditions can speed-up the simulations considerably [6].

FEA algorithm solves equation (3) for every node of the mesh created by division of region of interest with finite elements using some of well-known numerical algorithms, such as Newton-Raphson's, for instance. After finding values of A_z in all nodes of meshed region, calculating flux linkages of specific phase windings can be done easily. The flux linkage Ψ_{ij} of the j-th winding when the i-th winding is supplied with current can be expressed by [7]:

$$\Psi_{ij} = p\frac{1}{S_j}\left(\int_{\Omega_j^+} A_z d\Omega - \int_{\Omega_j^-} A_z dS\right) N_j l_j \tag{4}$$

where N_j is the number of turns of j-th winding, l_j is the length of the stator core and S_j is the cross-section of the coil

region. Ω_j^+ and Ω_j^- represent the positive and negative trace of the winding in (*x*,*y*) plane. The flux vector in d-q domain can be formed as follows:

$$\Psi = \Psi_{ds} + j\Psi_{qs} = \frac{2}{3}(\Psi_a + a\Psi_b + a^2\Psi_c) \quad (5)$$

$$\Psi_{ds} = \frac{2}{3}(\Psi_a - \frac{1}{2}\Psi_b - \frac{1}{2}\Psi_c) \quad (6)$$

$$\Psi_{qs} = \frac{1}{\sqrt{3}}(\Psi_b - \Psi_c) \quad (7)$$

IV. INDUCTANCE CALCULATION

Accurate inductance calculation of PMSM is a relevant topic, since the inductances determine large part of the electrical machine behavior [1]. The inductances estimation in the d-q axis is crucial not only for determination of the torque and flux weakening capability but also for designing control systems in order to optimize the efficiency, power factor, etc. [3]. In this paper two types of IPMSM geometries will be analyzed. Both types have identical stator (three-phase four-pole concentrated winding), and the rotors are chosen in such a way to represent two most significant types of PM orientation used in IPMSMs. First type, here called IPMSM-T (Figure 2a) is with tangentially magnetized PMs, and the second one, here called IPMSM-R (Figure 2b) is with radially magnetized PMs. Figure 3 and 4 respectively show PMs air-gap flux distribution for the two mentioned types of motors. It is well known that in the type IPMSM-R quadrature inductance Lq is greater than direct inductance Ld, because there is more iron along q axis (Figure 2b). However, such conclusion is not so straightforward for the type IPMSM-T, because PMs flux path is between two neighboring magnets, which means through iron (which increases the inductance), but the flux path is much longer (which decreases the inductances).

The first FEA simulation will be used to calculate Ld. The magnet flux is turned off by setting the magnet remanence Br to zero. The current vector must be aligned with the d axis. Variation of inductance as a function of current amplitude is of interest, so series of magneto-static simulations with different levels of current excitation will be conducted in order to obtain these relationships. For example, if the magnitude of stator current is chosen to be 1A, then to align the current axis with phase a axis the instantaneous phase currents have to be defined as:

$$i_a = 1A,\ i_b = i_c = -\frac{i_a}{2} = 0.5A \quad (8)$$

The d and q components of the current vector are then defined as:

$$i_d = 1A,\ i_q = 0A \quad (9)$$

Figures 5 and 7 respectively show the field solution for this case, for both types of IPMSM. The flux linkages of phases a, b and c are then calculated using equation (4). The flux vector is constructed and it's direct and quadrature components are calculated using expressions (6) and (7). It is important to state that the 2-D FEA simulation neglects the leakage flux in the end region, because it cannot be included without usage of 3-D FEA algorithm [8], [9]. Analyzed machines have concentrated winding, which means that they have very short end connections, especially in comparison with distributed windings. Because of that, one may expect very small end turn leakage inductances. However, their influence can be included by adding some analytical expressions [7] on calculated inductances, or by comparing the 2-D FEA results with measurements, because measured inductances include those effects. In this paper end connections leakage inductances have been neglected.

The inductances Lq are calculated in a similar manner, only this time the current vector needs to be aligned with the q axis. The phase currents are then given as:

$$i_a = 0A,\ i_b = -i_c = \frac{\sqrt{3}}{2}A \quad (10)$$

The d and q components of the current vector are then defined as:

$$i_d = 0A,\ i_q = 1A \quad (11)$$

Figures 6 and 8 show the field solution for this case, for both types of IPMSM. After calculating the flux components, the inductances for one turn per coil are given by:

$$L_d = \frac{\Psi_{ds}}{i_d},\ L_q = \frac{\Psi_{qs}}{i_q} \quad (12)$$

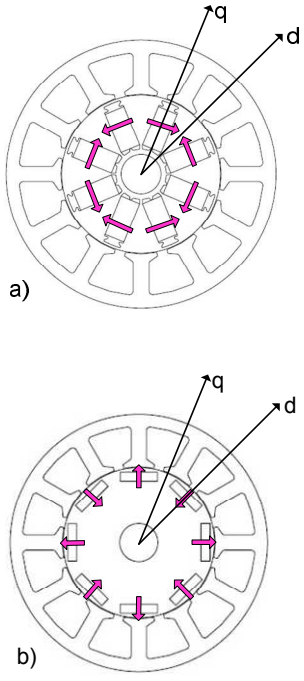


Figure 2. Two types of IPMSM: a) IPMSM with tangentially magnetized PMs (IPMSM-T), b) IPMSM with radially magnetized PMs (IPMSM-R)

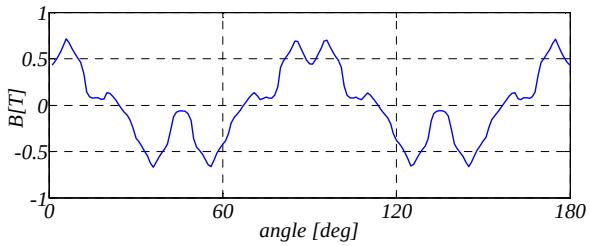


Figure 3. Flux distribution of permanent magnets in air gap of IPMSM-T

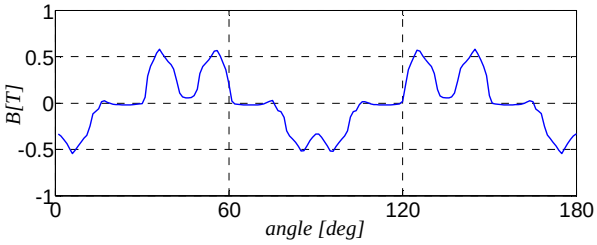


Figure 4. Flux distribution of permanent magnets in air gap of IPMSM-R

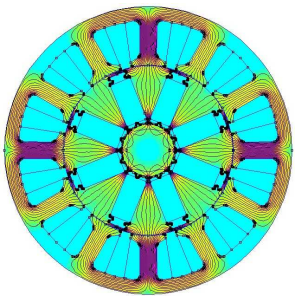


Figure 5. Distribution of flux lines in IPMSM-T when there is only d axis current in stator windings

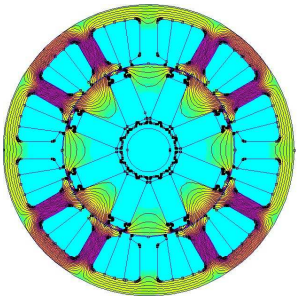


Figure 6. Distribution of flux lines in IPMSM-T when there is only q axis current in stator windings

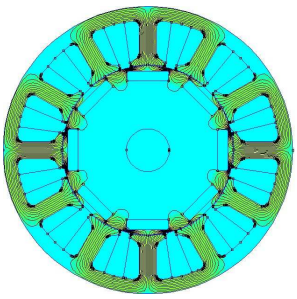


Figure 7. Distribution of flux lines in IPMSM-R when there is only d axis current in stator windings

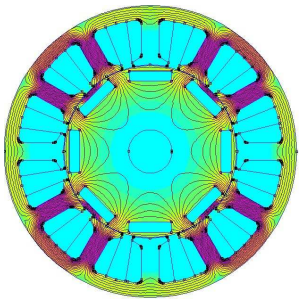


Figure 8. Distribution of flux lines in IPMSM-R when there is only q axis current in stator windings

V. RESULTS OF INDUCTANCE CALCULATIONS

Rated parameters for IPMSM-T machine are given in Table 1, and IPMSM-R machine was generated for the sake of comparison with IPMSM-T configuration. As mentioned above, both machines have the same stator. Results of calculations described in previous chapter are shown in Figures 9 (for inductance Ld) and 10 (for inductance Lq). It is interesting to notice that for both types of IPMSMs quadrature inductance Lq is approximately 50% greater than direct inductance Ld, and that inductance Ld is more influenced by saturation. Results are showing good match with experimental results (existent only for type IPMSM-T), but there are some difference between simulations and experiments (both Ld and Lq calculated with FEA are approx. 10% larger from those obtained with experiments). This can be explained by the fact that PMs cannot be excluded from experiments without destroying the rotor, and their field is added on time-varying field produced by excitation during experiments, which has influence on saturation level during measurement. Also, for the purpose of FEA simulations authors did not have precise data for steel used for stator stack (relative permeability as a function of field strength).

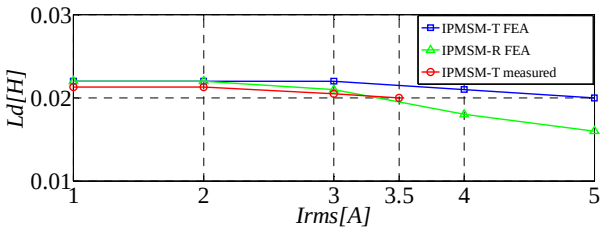


Figure 9. Results of FEA simulations and measurements for direct axis inductance Ld

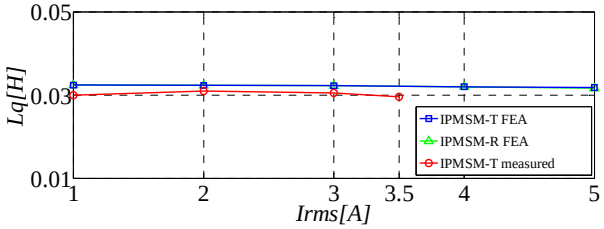


Figure 10. Results of FEA simulations and measurments for quadrature axis inductance Lq

TABLE I. IPMSM-T NAMEPLATE DATA

IPMSM-T rated data	Values
Number of stator slots	12
Number of poles	8
Stator core outside dimension [mm]	102
Stator core stack thikness [mm]	42
Stator core inside diameter [mm]	60.6
Air gap width [mm]	0.3
Winding specification	φ0.65×84T, Y ,Al
Resistance (2 phase) 20°C	7.5Ω
Direct and quadrature inductances [mH]	Ld=20.8 Lq=30.1
Rated speed [rpm]	500
Rated torque [Nm]	1.00

VI. CONCLUSION

In this paper, the method for calculation of inductances for two types of IPMSMs was presented. The method is based on the FEA calculation of flux linkages. It was shown that the quadrature inductance Lq is greater than direct inductance Ld for both types of motors. Further research using FEA software combined with measurements may be useful for determination of end connection leakage inductance influence, and also to investigate saturation of q axis flux path because of the presence of permanent magnets. Results for type IPMSM-T are showing good match with the results from the manufacturer of the machine. Presented method can be useful for detail analysis of different constructions of IPMSM with possibility to separate influence of excitation from influence of permanent magnets field, which cannot be done easily during measurements.

ACKNOWLEDGMENT

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under project III42004.

VII. REFERENCES

F. Shiferl, T.A. Lipo, “Power Capability of Salient Pole Permanent Magnet Synchronous Motors in Variable Speed Drive Applications”, IEEE Trans. Ind. Appl., vol. 26, pp. 115–123, Jan./Feb. 1990.

Y. Li, X.P. Yan, “The perspective and status of PMSM electrical servo system,” *Micromotors Servo Technique*, vol. 4, pp. 30-33, 2001.

M. Pastorelli, J. Bottomley, P. Giangrande and C. Gerada, “Sensorless control of PM motor drives — A technology status review,” IEEE Work. Elec. Mach. Design Contr. Diag. (WEMDCD), vol. 48, no. 4, pp. 168–182, March 11-12, 2013.

K. Jang-Mok and S. Seung-Ki, “Speed control of interior permanent magnet synchronous motor drive for the flux weakening operation,” IEEE Trans. Ind. Appl., vol. 33, no. 1, pp. 43–48, January/February 1997.

D. Zarko, “A Systematic Approach to Optimized Design of Permanent Magnet Motors With Reduced Torque Pulsations”, Pd.D. thesis, University of Winsconsin-Madison, USA, 2004.

P. Silvester, “Finite Elements for Electrical Engineers”, Cambridge University Press, 1990.

[7] J. Cros and P. Viarouge, “Synthesis of High Performance PM Motors With Concentrated Windings”, IEEE Trans. Energy. Conv., vol. 17, no. 2, pp. 248-253, Jun 2002.

[8] T. Cox, F. Eastham, and J. Proverbs, “End Turn Leakage Reactance of Concentrated Modular Winding Stators,” IEEE Transactions on Magnetics, vol. 44, no. 11, pp. 4057-4061, November 2008.

[9] A. M. EL-Refaie, “Fractional - Slot Concentrated-Windings Synchronous Permanent Magnet Machines: Opportunities and Challenges”, IEEE Trans.Ind. El., vol. 57, no. 1, January 2010.

A suplementary tool to the standard teaching metod of induction motor drive control

Vlado Porobić
Department of Power, Electronics, Communications
Faculty of Technical Sciences
Novi Sad, Serbia
poroba@uns.ac.rs

Evgenije Adžić, Zoran Ivanović, Marko Vekić, Stevan Grabić
Department of Power, Electronics, Communications
Faculty of Technical Sciences
Novi Sad, Serbia

***Abstract*—This paper demonstrates an modern approach to teaching motor drive control, which can be used as additional tool to the standard education methods or even completely self contained. Emulation-based virtual laboratory is proposed for control engineering education purpose. By using the example of induction machine digital control, there is shown that emulation experiments can give students an industrially relevant educational experience. Both digital controller design steps and suitable emulated power stages, together with experiment results are presented in detail.**

Keywords- Controller, Motor drive, Hardware-in-the-Loop, Education

I. INTRODUCTION

Education in applied fields, such as control engineering, can take many forms ranging from highly theoretical to deeply vocational. Typically, theoretical books cover the "what" rather than the "how." Thus, lecturers are left with considerable flexibility regarding practical, i.e. demonstration and experiment process. On the other hand, graduates put different emphases in their education. For example, those who go on to complete a higher degree will typically need a strong foundation in basic scientific principles, whereas those who choose industrial career typically need a greater emphasis on practical relevance.

A well-rounded course in control engineering needs to combine the appropriate theoretical knowledge with a strong emphasis on relevance. Thus, basic theory is best presented in a directed fashion with the going through suitable examples. It means, that lecture courses need to be supported by practical laboratory sessions. However, this leads to a dilemma for educators, namely how to achieve an appropriate tradeoff between cost, flexibility and safety of laboratory. In that way, there has been substantial ongoing interest by engineering educators and education researchers throughout the world. A comprehensive review of the literature relating to hands-on, simulated and remote laboratories is given in [1]. Paper [2] refers constant pressure on universities and the emergence of new technologies in the creation of novel engineering systems in education, especially simulation and remote-access laboratories. Proponents of simulations assert that physical laboratories needlessly consume university space and funds, while advocates of hands-on laboratories argue that students should be exposed to real environments. Remote laboratories have appeared as a third option. Authors in [3] discuss about factors of a separation, both physical and psychological, between the students and the laboratory hardware and a technology-mediated interface that is used to close this distance.

Practical side of education in control engineering field, especially electrical drive control is one of the most challenging tasks. Modern electrical drives require fast, accurate and robust control system. Powerful digital signal processors (DSP) offer a variety of different functions providing implementation of sophisticated algorithms. These devices have to provide high drive performance regarding many aspects such as accuracy, low energy consumption, safety, motor parameter mismatch and different motor size handling. Given the cost of purchasing, operating and maintaining motor drives, some educators question whether it is actually necessary to have experimental work at all. On the other hand, students strongly perceive the need to relate the theoretical content of their courses to the real world. Therefore, emergence of a highly flexible and safe platform for education is the key driver in courses comprising digitally controlled drives. In that sense, the real-time emulators could be very useful, since they offer motor drives control studying in an effortless manner. They provide variety of different power converter topologies, motor types, power levels, torque load profiles, etc. Considering mentioned, the ultra low latency Hardware-In-the-Loop (HIL) platform has recognized as a high end prototyping tool [4]-[5], which can be successfully utilized for motor drive control development process also [6]-[7].

In this paper, main practical aspects of motor control are shown on the typical and comprehensive example with indirect field oriented control (IFOC) of induction machine drive. Instead of using expensive, inflexible and potentially unsafe real hardware, DSP interacts directly with Hardware-In-the-Loop system. In that manner, all DSP control signals are fed to real-time emulating system comprising power converter and induction machine, while all feedback signals are returned to the controller. In the first part of the paper, sinusoidal and space vector modulation strategy for voltage generation is tested on a emulated three phase resistive-inductive load. Simplicity of the proposed system is then illustrated in the

paper second part in the example of IFOC induction machine drive. For the current and speed control loops, system step responses are shown for a standard parameters selection procedure. Verification of reference frame angle calculation in IFOC case is presented also.

II. EDUCATIONAL SETUP DESCRIPTION

Complete control algorithm is implemented on digital signal processor as in a real environment case, while total power stage is realized on a emulator. Proposed system for the education is shown in Fig. 1.

Figure 1. Educational setup with DSP controller and HIL

Digital signal controller is placed in appropriate docking station which adjust signal levels between it and HIL device. For the motor control and power electronics applications a devoted DSP, having peripherals such as PWM, ADC, quadrature encoder and capture unit is chosen (TMS320F2808, Texas Instruments). DSP software development environment, Code Composer Studio is used for DSP control education purpose. Lego like software blocks (such as Park&IPark transformation, space vector PWM modulation, PI regulator, speed calculation, etc.), written in widely spread "C" programming language are used.

Power stage of the considered scheme, comprising grid voltage, rectifier, inverter and motor is completely emulated in real time on the HIL system. HIL used in this paper is based on scalable, custom, ultra low latency processor design implemented on FPGA chip. This approach results in I/O latency of the order of 1μs for the typical two and three-level topology of PE converters. Fidelity confirmation as well as details about power electronics converters modeling used in this HIL platform is based on the work in [8]-[11]. The used HIL platform is flexible enough to cover not only a motor drive, but also a range of power electronics (PE) applications and in this way facilitates rapid system-prototyping. Furthermore, modeling environment, using a graphical user interface is intuitive and easy to use. Software tool-chain consists of the three basic elements, Fig. 2. It includes a schematic editor, where scheme is easily drawn from rich library of predefined switching models of PE elements and linear models of electrical machines, electrical sources, and passive elements. The second part is circuit compiler, which translates schematic in the HIL acceptable format. The third part is emulator control panel which controls emulation during runtime, e.g. its start/stop, change source values/frequencies, open/close contactors, etc. These tools provide a flexible environment for model editing, compiling and running the emulation.

The aim of the described experimental setup is to guide students and researchers in obtaining hands on experience to build, develop and verify control algorithm in a highly flexible and safe manner.

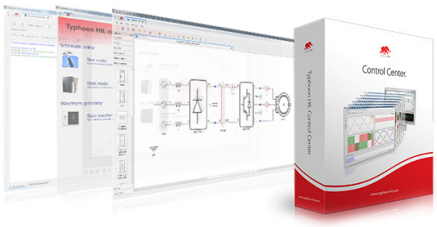


Figure 2. HIL control center toolchain

III. EXPERIMENTAL WORK IN A VIRTUAL LABORATORY

In order to demonstrate effectiveness of utilized system, standard control for the IFOC induction machine drive, shown in Fig. 3, is considered here.

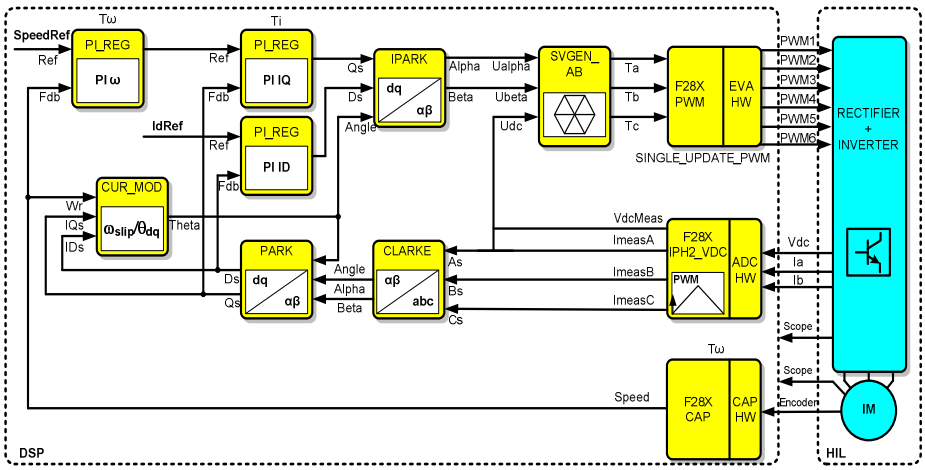


Figure 3. Example: block diagram of indirect field oriented control of induction motor, DSP and HIL side

IFOC is widely used control, based on transformation from stationary to synchronously rotating coordinates which leads to a simple control structure similar to that of a separately excited DC machine. Position of the reference frame tied to machine rotor flux is calculated as integral term of sum of electrical rotor speed and estimated slip frequency.

There are proposed several steps that thoroughly introduce students to the world of digital motor control. In addition to being used for the education, they can be encountered in the industrial environment also. They are summarized in Table I. In the following text, the most significant details of them will be presented.

TABLE I. TEST PHASES OF THE IFOC MOTOR DRIVE CONTROL

Software Module	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
RAMPCTRL	√	√	√	√	√
RAMPGEN	√	√	√	√	√
IPARK	√√	√	√	√	√
SVGEN	√√	√	√	√	√
PWM	√√	√	√	√	√
CLARKE		√√	√	√	√
PARK		√√	√	√	√
CAPTURE			√√	√	√
SPEED_PR			√√	√	√
PI (I_d, I_q reg)			√√	√	√
CURMOD (θ_{dq})				√√	√
PI (spd reg)					√√

Note: the symbol √ means this module is using and the √√ means this module is testing in this phase

The first, usual step in a depicted motor drive control is verification of PWM modulation strategy, no matter which type is used. Although space vector modulation is commonly used comparing to sinusoidal, just for educational reason both will be presented here. For that purpose, block diagram of control code (i.e. DSP side) shown in Fig. 4. is used. Output voltage amplitude is set by constant $V_{dTesting}$ and $V_{qTesting}$ variables, while output frequency is defined by constant *SpeedRef* value, which is smoothed and transferred to ramp signal by software blocks RAMPCTRL and RAMPGEN, respecively.

Appropriate power stage is emulated on HIL side (Fig. 5.): three-phase inverter, DC constant voltage source in converter DC-link circuit, RL load at converter output and some measurement elements for feedback values. In general, additional measurement equipment such as for line voltage measurement and filtering here are not necessary since this signals are automatically present in HIL safe and user friendly environment.

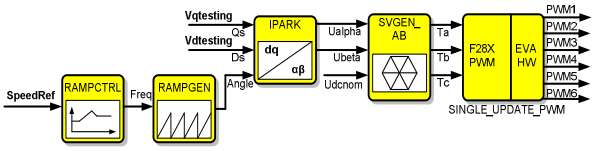


Figure 4. Block diagram of control code for verification of PWM modulation strategy (DSP)

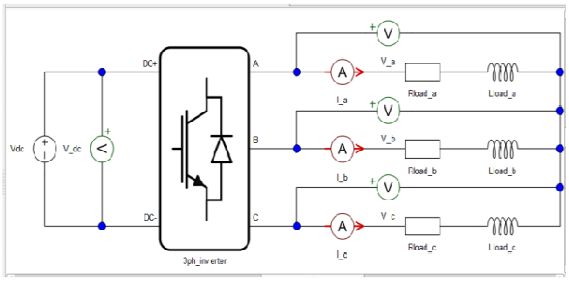


Figure 5. Emulated three phase inverter with RL load (HIL)

Expected load phase current value and load impedance can be calculated as in (1).

$$i_{out}^{amp} = \frac{v_{out}^{amp}}{z_{load}} = \frac{v_{out}^{amp}}{\sqrt{R_{load}^2 + x_{load}^2}} = \frac{v_{out}^{amp}}{\sqrt{R_{load}^2 + (2\pi f L_{load})^2}} \quad (1)$$

In a case of sine PWM modulation strategy, phase output voltage is calculated as in (2).

$$v_{out}^{amp} = \mathrm{mod}_{ind} \frac{v_{dc}}{2} \quad (2)$$

For this case, the calculated duty-cycle values for each phase legs and corresponding voltage reference in controller program are shown on the upper part of Fig. 6. On the lower part of this figure, motor phase voltage and currents at HIL output are given (modulation index 0.5, f = 50Hz).

In a case of space vector PWM modulation strategy, we have better utilization of available dc link voltage comparing to sinusoidal strategy. The phase output voltage is calculated as in (3).

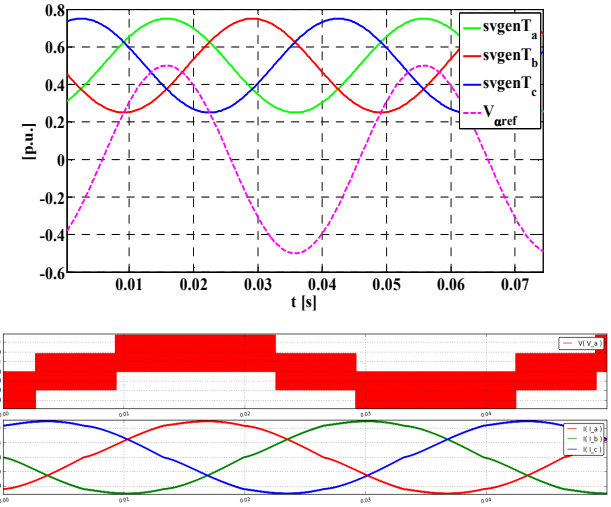


Figure 6. Sine PWM modulator: recorded controller variables (duty-cycle values for each phase legs, voltage reference) and captured HIL model data:

three-phase currents (Ia, Ib, Ic) and phase voltage (Va) for reference frequency 50 Hz and modulation index 0.5

$$v_{out}^{amp} = \mathrm{mod}_{ind} \frac{v_{dc}}{\sqrt{3}} \tag{3}$$

For this strategy, the calculated leg's duty-cycle values and corresponding alpha voltage reference in controller are shown on the upper part of Fig. 7. On the lower part of this figure, motor phase voltage and currents at HIL output are given (modulation index 0.5, f = 50Hz). One can note that DSP calculated duty-cycle values have typical waveforms which deviates from sinusoidal, with two humps in the area of its maximum and minimum values.

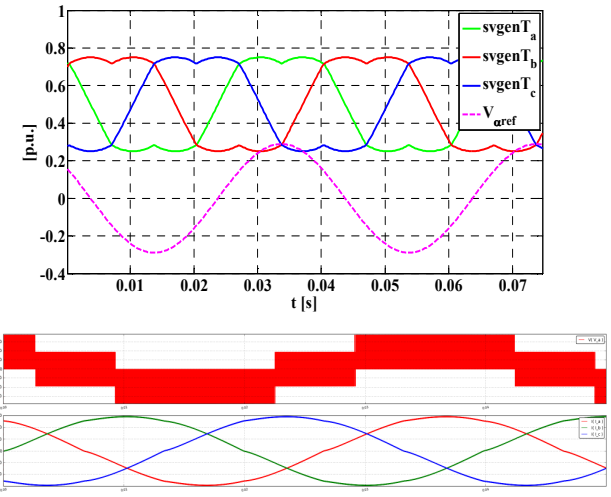


Figure 7. Space vector PWM modulator: recorded controller variables (duty-cycle values for each phase legs, alpha voltage reference) and captured HIL model data: three-phase currents (Ia, Ib, Ic) and phase voltage (Va) for reference frequency 50 Hz and modulation index 0.5

Both given examples are given for schematic load resistance 1Ω, inductance 1mH and dc-link voltage 563V (230·√3). Therefore, there can be easily verified that inverter current amplitude matches theoretical results, Fig. 6. and Fig. 7.

After successful PWM modulation test, an open loop control of induction machine (Fig. 8) can be easily employed. For the given voltage frequency and amplitude, machine rotor without torque load will rotates at synchronous speed divided by pole pairs number.

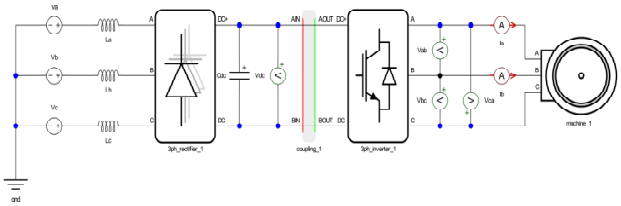


Figure 8. Emulated power stage of motor drive (HIL)

The next important step in motor drive control design is to setup current control loop. In that sense, limits and gains of current loop PI controllers parameter arrangement is the most sensitive task. In order to eliminate side effects, constant DC bus (therefore DC-link voltage ripple is zero), inverter zero dead time and locked rotor (high potential torque load and/or high inertia coefficient) is easily selected in the HIL environment.

According to induction motor voltage and flux equations, expressions for electrical part of the machine in a case of vector control can be obtained (4), (5).

$$U_{sd} = R_s i_{sd} + L_\sigma \frac{di_{sd}}{dt} - \omega_s L_\sigma i_{sq} \tag{4}$$

$$U_{sq} = R_s i_{sq} + L_\sigma \frac{di_{sq}}{dt} + \omega_s L_\sigma i_{sd} + \omega_s \frac{L_m}{L_r} \psi_{rd} \tag{5}$$

After applying decoupling strategy and neglecting the last, slow changing term in (5), current regulation scheme is obtained, Fig. 9. In a case of permanent magnet synchronous machine, parameter L_σ has to be replaced with L_s or L_{sd}/L_{sq} depending on whether the magnets are surface or interior mounted. Similar, in a case of DC machine parameter L_σ should be replaced with armature inductance parameter L_a. Variables denoted with ^ sign are controller side motor parameters, i.e. expected ones. In the case of this HIL utilization, they are set to be equal with real ones, but in general they can be changed easily for test purposes.

Motor current loop response according a number of different procedures for PI regulator parameters can be easily observed. Here, results of one of the most commonly used, Dahlin algorithm [12] is given. For the first order system transfer function (6), where T_i is object time constant (here T_σ) and τ is inverter transport delay (here T_{PWM}), parameters are given in (7) and (8).

$$G_{ob}(S) = \frac{1}{1+sT_i} e^{-\varpi} \tag{6}$$

$$K_p = \frac{1 - e^{-\lambda T}}{K \cdot (e^{\frac{T}{T_1}} - 1) \cdot \left[1 + N \cdot \left(1 - e^{-\lambda T}\right)\right]} \tag{7}$$

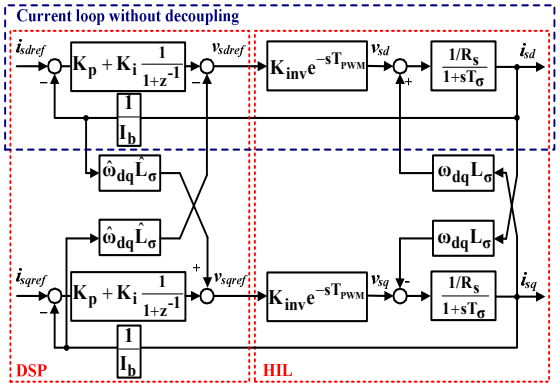


Figure 9. Current regulation scheme

$$K_i^z = K_p \frac{T}{T_i} = K_p (e^{\frac{T}{T_1}} - 1) \qquad (8)$$

Total object gain is $K = K_{inv} \cdot (1/R_s) \cdot (1/I_b)$, where is $K_{inv} = V_{dcbus}/\sqrt{3}$ for windings star connection; number of delay periods is $N=1$. Parameter λ defines system response having dominant time constant $T_d = 1/\lambda$ and here is chosen to be $5/T_\omega$ which means that current will achieve reference in one speed period loop (T_ω). For the calculated parameters, f_{PWM} = 20kHz, current loop period 1/4kHz and speed period T_ω = 5ms, direct axis current step response of the stalled motor is given in Fig. 10, upper part. One can notice that referenced current is achieved without overshoot (aperiodic) and in a expected time of $5 \cdot T_d = T_\omega$ what is in accordance with Dahlin's algorithm. On the lower part of the Fig., zoomed signals of phase current (ripple) and input voltages captured on HIL side are presented.

In a case of induction motor indirect field oriented control with *dq* coordinate system tied to rotor flux, the next important step is to test the reference frame angle calculation. For calculation of this angle, the fundamental thing is electrical slip frequency estimation, given in (9). According this equation, slip frequency can be easily numerically evaluated for several values of quadrature current i_{sq} under the same condition (locked rotor) for the machine with nominal flux (referenced nominal direct current i_{sdn}). In Fig. 11, motor phase currents and electrical torque for fluxed machine with stepped quadrature current i_{sq} is shown. In this experiment, the same value for both *d* and *q* stator current is chosen, resulting that slip frequency is equal to reciprocal of rotor time constant.. Frequency of currents in locked rotor case is the same as calculated DSP slip frequency given in (9), since rotor speed is zero.

$$\omega_s = \frac{L_m}{T_r} \frac{i_{sq}}{\psi_{rd}} = \frac{1}{T_r} \frac{i_{sq}}{i_{sd}} \qquad (9)$$

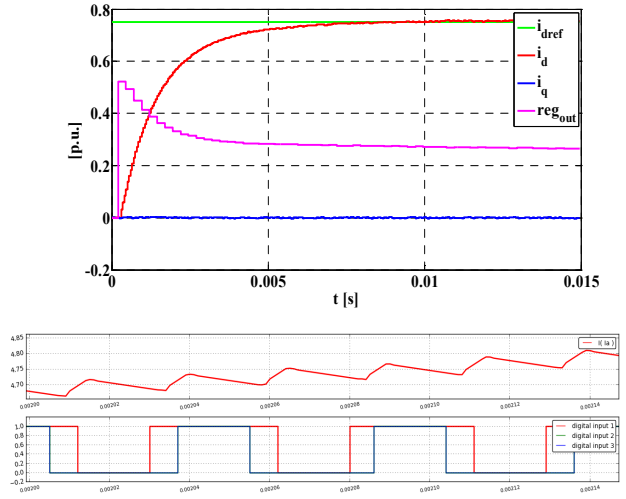


Figure 10. *Current step response (0.75 p.u.): controller variables (reference/actual currents and regulator output) and captured HIL model data: phase current and inverter leg voltages (zoomed).*

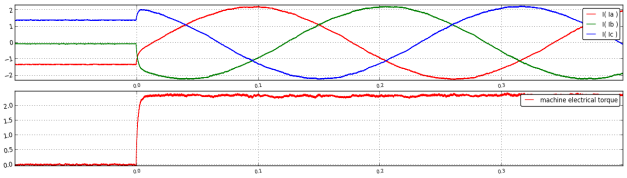


Figure 11. *IFOC, dq reference frame angle calculation test (i_q current step 0.2 p.u., constant i_d=0.2 p.u.); captured HIL model data: phase currents and electrical torque.*

The last step in drive control design is test of speed control loop. According to block diagram in Fig. 12, parameters of speed PI regulator are calculated [13]. In this case, speed is reached without overshoot and for minimal possible time. Implemented PI regulator has proportional gain in local path in order to avoid abrupt speed variation. Since speed information is averaged per a loop period, measurement transfer function is given as in (10), while gain K_m in vector drive as in (11).

$$W_{meas}(z) = \frac{1 + z^{-1}}{2} \qquad (10)$$

$$T_{el} = \frac{3}{2} P \frac{L_m}{L_r} \psi_{rd} i_{sq}^{[A]} = \frac{3}{2} P \frac{L_m^2}{L_r} i_{sd}^{[A]} I_b i_{sq}^{[rel]} == K_m i_{sq}^{[rel]} \qquad (11)$$

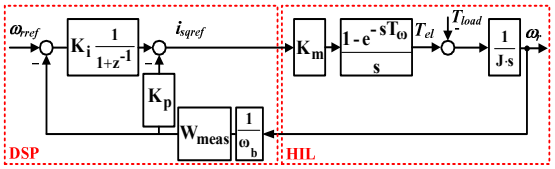


Figure 12. *Speed regulation scheme*

Speed step response is shown in Fig. 13. This results are completely in accordance with theoretical expectations also.

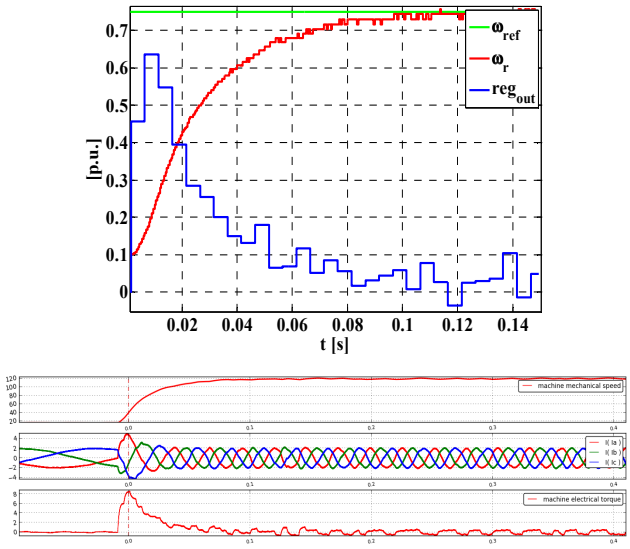


Figure 13. *Speed step response (0.75 p.u.): controller variables (reference/actual speed and regulator output) and captured HIL model data: speed, phase currents and electrical torque, respectively*

IV. CONCLUSION

This paper discussed how to overcome gap between basic theory on the one hand and engineering design reality on the other. The paper identified one efficient solution for bringing students closer to reality by the use of emulation based virtual laboratory. This type of laboratory gives students exposure to the reality of industrial control system design in a structured and user-friendly environment. Using Hardware-in-the-Loop system, students are capable to learn, test and verify every aspect of their control procedure. Beside great flexibility, this proposed system presents highly safe environment. In that manner, it allows user to completely focus on software-based functionalities without any hardware concerns.

ACKNOWLEDGMENT

This research was partially co-funded by the Ministry of Education, Science and Technological Development of Republic of Serbia under contract No. III 042004 and by the Provincial Secretariat for Science and Technological Development of AP Vojvodina under contract No. 114-451-3508/2013-04.

REFERENCES

[1] J. Ma and J. V. Nickerson, “Hands-on, simulated, and remote laboratories: A comparative literature review,”Comput. Surv., vol. 38, no. 3, p. 7, 2006.

[2] J. V. Nickerson, J. E. Corter, S. K. Esche, and C. Chassapis, “A model for evaluating the effectiveness of remote engineering laboratories and simulations in education,”Comput. Educ., vol. 49, no. 3, pp. 708–725, 2007.

[3] E. Lindsay, S. Naidu, and M. Good, “A different kind of difference: Theoretical implications of using technology to overcome separation in remote laboratories,”Int. J. Eng. Educ., vol. 23, no. 4, pp. 772–779, 2007.

[4] S. Lentijo, S. D'Arco, A. Monti, “Comparing the dynamic performances of power hardware in the loop interfaces”, *IEEE Trans. Ind. Electron.*, Apr. 2010, vol. 57, no. 4, pp. 1195–1208.

[5] W. Dirk, and M. Kratz, "A real time development platform for next generation of power systems control functions," *IEEE Trans. Ind. Electron.*, Apr. 2010, vol. 57, no. 4, pp. 1159-1167.

[6] S. Abourida and J. Belanger, "Real-time platform for the control prototyping and simulation of power electronics and motor drives", *Proc. 3rd Int. Conf. on Modeling, Simulation, and Applied Optimization*, Sharajah, U.A.E., Jan. 2009.

[7] E. Adžić, M. Adžić, V. Katić, D. Marčetić, N. Čelanović, "*Development of High-Reliability EV and HEV IM Propulsion Drive With Ultra-Low Latency HIL Environment*", IEEE Transaction of Industrial Informatics, Vol. 9, No. 2., pp. 630-639, May 2013.

[8] D. Majstorović, I. Čelanović, N. Teslić, N. Čelanović, V.A. Katić, "Ultra-Low Latency Hardware-in-the-Loop Platform for Rapid Validation of Power Electronics Designs”, *IEEE Trans. Ind. Electron.*, Oct. 2011, vol. 58, no. 10, pp. 4708-4716.

[9] M. Vekić, S. Grabić, D. Majstorović, I. Čelanović, N. Čelanović, V. Katić "*Ultralow Latency HIL Platform for Rapid Development of Complex Power Electronics Systems*", IEEE Transaction of Power Electronics, Vol. 27, No. 11., pp. 4436-4444, Novembar 2012.

[10] J. Poon, M. Kinsy, N. Pallo, S. Devadas, I. Čelanović, *Hardware-in-the-Loop Testing for Electric Vehicle Drive Applications*, Proc. of 27 Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 2576–2582, Feb. 2012.

[11] J. Poon, E. Chai, I. Čelanović, A. Genić, E. Adzić, *High-Fidelity Real-Time Hardware-in-the-Loop Emulation of PMSM Inverter Drives*, IEEE Energy Conversion Congress, Denver, 2013.

[12] E. B. Dahlin, “Designing and Tuning Digital Controllers,” Instruments and Control systems, vol. 41, No. 6, 1968.

[13] S. Vukosavic, "Digital Control of Electrical Drives", Springer 2007, ISBN-13: 978-1441938541.

Session TO-5
MEASUREMENT METHODS AND SYSTEMS

Design of Monitoring and Data Acquisition System for Environmental Sensors

Mitar Simić
NORTH Point Ltd, Member of the NORTH Group
Trg Cara Jovana Nenada 15/8, 24000 Subotica
Republic of Serbia
mitarsimic@yahoo.com

Abstract—In this paper description of monitoring and data acquisition system for environmental sensors is described. Device can be used as standalone device but it also can be connected to the PC through serial communication. Color display is used for local graphical presentation of sensors readings while for online PC analysis specialized application was developed. Application provides alarm occurrence log with e-mail notifications. System is scalable and easily can be expanded with new sensors. System verification was done with SHT11 measurement device and system worked stable and reliable.

Keywords-data acquisition, environmental sensors, monitoring systems, relative humidity and temperature measurement.

I. INTRODUCTION

Real time monitoring and data acquisition have important role in industry but also in everyday life. For example, in literature can be found description of systems used in pollutant detection [1], storing and transportation of medicinal products [2], clothes and books storing [3], food transportation [4] and greenhouses air quality monitoring [5].

In last few decades attention is especially focused on environment (air, soil and water) pollutant monitoring. These systems require sensing elements with high accuracy, selectivity and sensitivity. Data acquisition in these systems is also very important and systems have to be able to store data in long time period (in some applications even more than decade) for analysis and predictions what can be expected in future. Collecting of measurement results must be done with appropriate sampling time and reliable alarm system.

Sensors developed within SENSEIVER project [6] are characterized and mostly tested in laboratory [7]-[12]. Sensors showed excellent characteristics so they can be implemented in the field in real time pollutant monitoring systems. That means that specific equipment needs to be developed for in the field data acquisition [13]-[20] and material characterization [21]-[23].

Therefore in this paper is presented study on such data acquisition system with real time monitoring and online analysis capabilities. System presents improved design of systems reported earlier [18]-[24].

In the first version of the system [18], only local monitoring and offline analysis on PC were possible but that system made core for further upgrades. System was powered by rechargeable batteries with it's own monitoring system and charger unit. Device had two modes for data collecting: continuous and manual. In continuous mode, the device collected data from sensor automatically with defined sampling time and stored it on micro SD card. In manual mode, the user had an option to manually choose the moment of recording the results. Created report of measured values with date and time stamp was stored in format which is compatible with MS Excel. Device had keypad with two navigation and two confirmation keys so full hand-held configuration and operation capabilities were possible. Connection between PC and device was available through serial communication interface: RS-232 or USB cable. Device also had alphanumeric LCD for displaying results and configuration in the field. System verification was done with SHT11 measurement device.

All good features of system presented in [18] are kept and implemented with optimization regarding dimensions and power consumption in new redesigns of the system. In every new design additional option was implemented. For example, in [19] real time analysis was added with PC application developed in LabVIEW. System was later improved also with embedding a PC based web server into the microcontroller [20].

Main contribution of work presented in this paper is adding local real time monitoring with color display and new PC application developed in Microsoft Visual Studio. Microsoft Visual Studio was chosen because it is now free for usage even in systems with commercial purposes [24].

II. DEVICE STRUCTURE

Proposed structure of redesigned system for monitoring and data acquisition of environmental sensors is presented in Fig. 1. Main hardware improvement is replacement of LCD display used before [18]-[20] with TFT color display [25] which allows realization of local system for real-time monitoring and data acquisition with graphical analysis of obtained results. New system also can be USB powered and has smaller dimensions.

This work was financially supported by European Commission in the framework of the FP7 project SENSEIVER, grant number 289481.

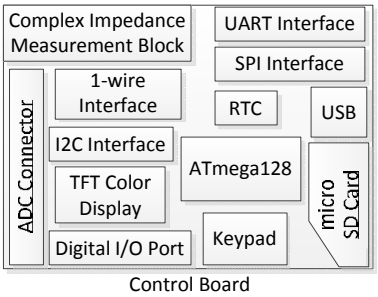


Figure 1. Block-scheme of proposed device

The main part of the Control board is microcontroller AVR ATmega128 [26] that interfaces with user via display and keypad. Other peripherals on the board are micro SD card, real time clock DS1307 [27] and TFT color display.

Measured values of relative humidity and temperature are stored on 2 GB micro SD card. FAT16 file system is used to allow manipulation of file with results on PC and off-line processing. The main advantage for using a micro SD card as data storage system is the ease of transferring data directly to other electronic devices which support the FAT format as a file system.

Real time clock DS1307 is used to provide time measuring. Time and date values are used for presentation on display and for additional information for every result in report stored on micro SD card.

A thin-film-transistor liquid-crystal display (TFT LCD) is a variant of a liquid-crystal display (LCD) that uses thin-film transistor (TFT) technology to improve image qualities such as addressability and contrast. A TFT LCD is an active-matrix LCD, in contrast to passive-matrix LCDs or simple, direct-driven LCDs with a few segments. TFT LCDs are used in appliances including television sets, computer monitors, mobile phones, handheld video game systems, personal digital assistants, navigation systems and projectors [28].

Keypad with two navigation and two confirmation keys is connected to the Control board to ensure easy manipulation through menu system for configuration and measurements. Menu system is based on finite state machine and it can be accessed by external keyboard with two navigation keys (Up and Down) and two confirmation keys (OK and Esc).

The Control board also has connectors for:

- analog to digital conversion (8 inputs),
- 8 digital general input/output pins,
- one-wire devices interface,
- complex impedance measurement,
- UART interface,
- I2C interface and
- SPI interface.

Connector with 8 digital input/outputs ensures that sensor device SHT11 [29] can be attached directly to the microcontroller. The sensors integrate sensor elements plus signal processing and provides a digital output so no additional calibration is needed. SHT11 is a Sensirion's family of surface mountable temperature and relative humidity sensors. Relative humidity is measured by capacitive sensor element while band-gap sensor is used for temperature measurement. The SHT11 has operating range of temperature between -40 and +123.8 °C and range for relative humidity between 0 and 100 %RH.

A hardware outcome of Control board with attached SHT11 sensor is shown in Fig. 2.

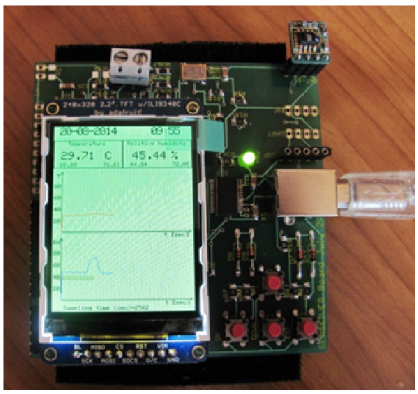


Figure 2. A prototype hardware outcome of Control board with connected SHT11 sensor

Main purpose of developed system presented in the Fig. 2 is to be used as a standalone device for real time monitoring and graphical presentation of obtained results in the field. But USB power supply option can also be used to provide real time connection of the board with PC and more detailed analysis of obtained results. For example, as was reported earlier in [25], LabVIEW can be used for PC data acquisition. In this work PC application was developed in Microsoft Visual Studio.

In Fig. 3 main screen of PC application for temperature and relative humidity monitoring is presented.

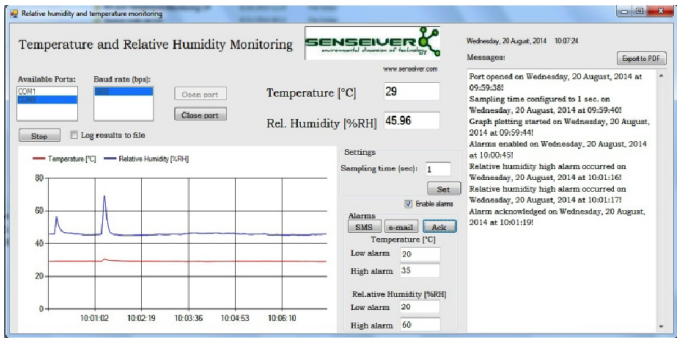


Figure 3. Main screen of the PC application

As can be seen from Fig. 3, PC application consists of few main blocks:

- Communication
- Sampling time configuration
- Alarms
- Graphs
- Logging to file

- Messages log

Communication. User can choose communication PC port which is connected to Control board. Baud rate is fixed and can not be changed because it is set to 9600 bps in the microcontroller program.

Sampling time configuration. After the Control board is connected with PC, user has to define sampling time between measurements. This value can be changed during operation.

Alarms. User can set alarm values for temperature and relative humidity. These values can be easily changed during operation and when any alarm occurs in message log that information is added. User has to acknowledge alarm with "ACK" button. If alarm is acknowledged but actual value is still above defined limit, appropriate filed will remain to be red but no new messages will be added in message log. There are also two buttons for alarm notifications: SMS and e-mail. SMS notification is not yet implemented because SMS gateways were not available in Serbia but with additional hardware (GSM modem) this can be done easily. E-mail notification is implemented and user can choose if e-mail should be sent by clicking to button "E-mail". Example of created e-mail message is presented in Fig. 4.

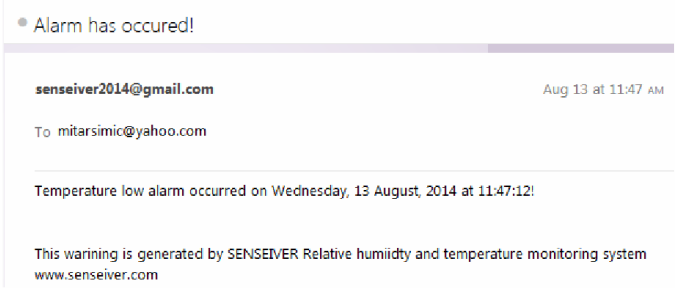


Figure 4. Example of generated e-mail alarm notification

Graphs. User has option to start and stop real time plotting of obtained results in any moment. Time stamp is added on x-axis.

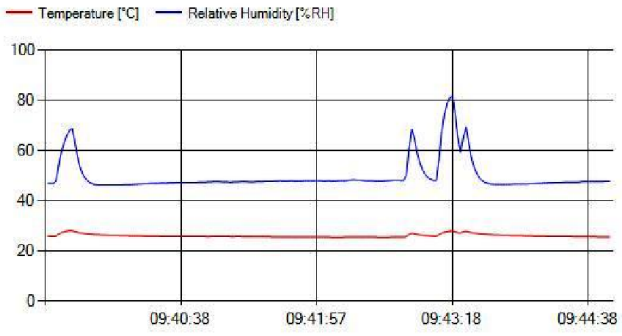


Figure 5. Examples of created graphs

Logging to file. The Control board performs storing of measurement results on the micro SD card, as noted before, but sometimes option for logging on PC hard drive can be useful and because of that it is implemented. In every moment user can disable logging to file and later enable it again. Example of created report is presented in the Fig. 6.

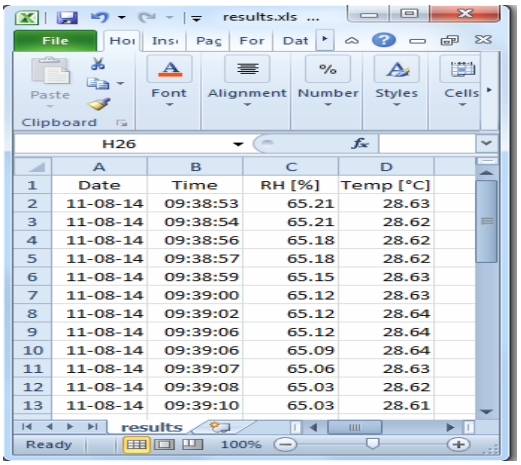


Figure 6. Example of created report

Messages log. All operations such as connecting/disconnecting from device, enabling/disabling of alarms or graph plotting, changing of sampling time value are logged and presented in right corner of PC application. There is button "Export to PDF" which generates report of logged activities in PDF file format. Example of created report is presented in Fig. 7.

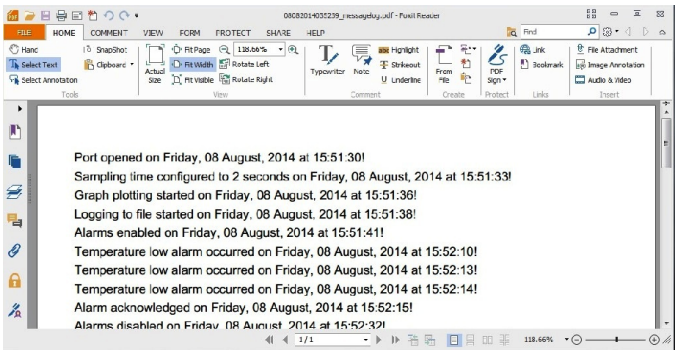


Figure 7. Example of message log

A hardware outcome of complete system is shown in Fig. 8.

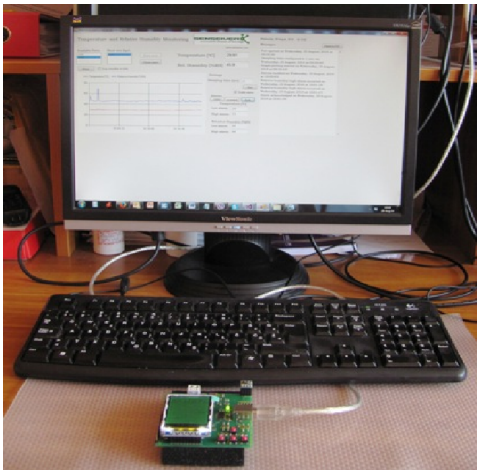


Figure 8. Hardware outcome of device for temperature and relative humidity monitoring

III. CONCLUSION

The main task in this study was to provide direct access to the sensor readings in real time. Device can be used as standalone or with PC connection. Local presentation of results is performed on color display with graphs and information about minimum and maximum measured values. If device is connected to the PC, more detailed analysis is possible with PC application.

The designed system was tested with SHT11 as measurement device under different conditions and in all tests the system worked stable and accurately. Main application of developed system is to be used for monitoring and data acquisition with environmental sensors developed within SENSEIVER project.

REFERENCES

[1] S. S. Chandrasekaran, S. Muthukumar and S. Rajendran, "Automated Control System for Air Pollution Detection in Vehicles", presented at the 4th International Conference on Intelligent Systems, Modelling and Simulation, 2013.

[2] T. H. Khan, K. A Wahid, "An advanced physiological data logger for medical imaging applications", URASIP Journal on Embedded Systems, pp. 1-14, 2012.

[3] N. Li, Y. Tang, "Library Temperature and Humidity Remote Control System Based on Micro Controller Unit", In proceedings of the 2nd International Conference on Computer Science and Electronics Engineering ICCSEE, 2013, pp. 3096-3098.

[4] K. H. Eom, C. W. Lee, N. T. Van, K. K. Jung, J. W. Kim and W. S. Choi, "Food Poisoning Prevention Monitoring System based on the Smart RFID Tag System", International Journal of Multimedia and Ubiquitous Engineering, Vol. 8, No.5, pp. 213-222, 2013.

[5] M. Omid, A. Shafaei, "Temperature and relative humidity changes inside greenhouse", Int. Agrophysics, Vol. 19, pp. 153-158, 2005.

[6] http://www.senseiver.com

[7] S. Toskov, A. Maric, N. Blaz, G. Miskovic, G. Radosavljevic, "Properties of LTCC Dielectric Tape in High Temperature and Water Environment", International Journal of Materials, Mechanics and Manufacturing, Vol. 1, No. 4, 2013.

[8] L. Manjakkal, K. Cvejin, J. Kulawik, K. Zaraska, D. Szwagierczak, "A Low-Cost pH Sensor Based on RuO2 Resistor Material", Nano Hybrids, Vol. 5, pp. 1-15, 2013.

[9] K. Cvejin, L. Manjakkal, J. Kulawik, K. Zaraska and D. Szwagierczak, "Synthesis of perovskite Sr doped lanthanide cobaltites and ferrites and application for oxygen sensors: a comparative study", Key Engineering Materials, Vol. 605, pp. 483-486, 2014.

[10] S. Toskov, G. Radosavljevic, "Water Temperature Sensor Built up in LTCC Technology", in Conference Presented at the 29th International Technical Conference on Circuits/Systems, Computers and Communications, Phuket, Thailand, 2014.

[11] G. Mišković, S. Toškov, A. Maric, N. Blaz, G. Radosavljevic, "Characterization of LTCC Tapes in Water Presence", Presented at the ICAMR, Dubai, UAE, 2013.

[12] M. Zawadzka, J. Kulawik, D. Szwagierczak, K. Zaraska, "Free-grown polypyrrole sensors", Presented at the 37th International Conference of IMAPS-CPMT Poland, 2013.

[13] A. Trandabat, M. Pislaru, C. Lorenz, M. Sireteanu, "Internet Based Virtual Laboratory in Bioengineering Field", Presented at the 2nd International Conference on Nanotechnologies and Biomedical Engineering ICNBME, Chisinau, Republic of Moldova, April 18-20, 2013.

[14] S. Ajkalo, G. Stojanovic, "Laboratory Prototype of Wireless Sensor System for Air Quality Parameters Monitoring", In Proceedings of the INFOTEH-JAHORINA Vol. 12,, 2013, pp. 1113-1117.

[15] A. Iavorschi, N. Zoric, S. Ajkalo, "Wireless Communication System for Temperature Monitoring", Presented at the 2nd Int. Conference on Nanotechnologies and Biomedical Engineering, Chisinau, Republic of Moldova, 2013.

[16] A. Iavorschi, V. Sontea, "Microcontroller Based Data Acquisition System for Environmental Monitoring", Presented at the 7th International conference and exposition on electrical and power engineering EPE, Iasi, Romania, October 25-26, 2012.

[17] M. Sireteanu, A. Iavorschi, L. Manjakkal, J. F. B. Villalba, "Design of data acquisition system for environmental sensors manufactured in LTCC Technology", Presented at the International conference and exposition on Electrical and Power Engineering EPE, October 16-18, Iasi, Romania, 2014. (accepted).

[18] M. Simić, "Microcontroller Based System for Measuring and Data Acquisition of Air Relative Humidity and Temperature", Presented at the 37th International Conference of IMAPS-CPMT Poland, 2013.

[19] M. Simić, M. Sireteanu, "Real Time Temperature and Relative Humidity Monitoring System using LabVIEW", In Proceedings of the International scientific conference of Metrology and Quality in Production Engineering and Environmental Protection – ETIKUM, 2014, pp. 67-70.

[20] M. Simić, "Design and Development of Air Temperature and Relative Humidity Monitoring System With AVR Processor Based Web Server", Presented at the 8th International conference and exposition on electrical and power engineering EPE, Iasi, Romania, October 16-18, 2014. (accepted).

[21] M. Simić, "Complex Impedance Measurement System for the Frequency Range from 5 kHz to 100 kHz", Presented at the 4th International Conference on Materials and Applications for Sensors and Transducers IC-MAST, Bilbao, Spain, June 2014.

[22] M. Simić, "Realization of Complex Impedance Measurement System Based on the Integrated Circuit AD5933", In Proceedings of the 21st Telecommunications forum TELFOR, 2013, pp. 573-576.

[23] M. Simić, "Realization of Digital LCR Meter", Presented at the 8th International conference and exposition on electrical and power engineering EPE, Iasi, Romania, October 16-18, 2014. (accepted).

[24] http://www.microsoft.com/express/support/faq/default.aspx

[25] http://www.adafruit.com/products/797

[26] ATmega128, datasheet

[27] DS1307, datasheet

[28] http://www.tftcentral.co.uk/

[29] SHT11, datasheet

Software Supported Procedure applied to Testing of Instruments for High-order Harmonics Measurement

Milan Simić, Dragan Živanović, Dragan Denić and Goran Miljković
Department of Measurement
University of Niš, Faculty of Electronic Engineering
Niš, Serbia
milan.simic@elfak.ni.ac.rs

***Abstract*—Software supported procedure, applied for generation of the test voltage waveforms with certain level of the standard harmonic disturbances, is presented in this paper. Procedure is functionally based on the virtual instrumentation concept, which includes control application in LabVIEW software environment and data acquisition board NI PCIe 6343. Variation of the basic parameters for definition, presentation and signal generation is provided by various control functions and switches, implemented on front panel of the developed virtual instrument. For specific harmonic disturbances is possible to define percentage amounts of the harmonic amplitude levels, nominal frequency variations, amplitude fluctuations, start and stop times, rising and falling times of the disturbances. Described acquisition system is verified by testing of the three-phase power quality analyzer Fluke 435 Series II. By this generation system is possible to provide various voltage test waveforms with typical harmonic disturbances. In this specific case, for testing purpose are used some characteristic test waveforms with harmonic disturbances. Basic measurement results and some recorded voltage signals, obtained from testing procedure, are presented and analyzed in this paper.**

Keywords-Software supported test procedure; signal harmonic disturbances; virtual instrumentation software; voltage waveforms.

I. INTRODUCTION

Increased using of the power electronic components and powerful switching devices directly causes degradation of electrical power quality (PQ), which affects on the production process costs and reduces reliability of the customer electrical devices and equipment. In order to avoid these problems and to increase total energy efficiency level, electricity suppliers must provide appropriate quality of the power distribution networks. For purpose of the customer protection, satisfactory PQ level is prescribed by the relevant international quality standards and documents [1,2]. Optimal PQ level is defined with acceptable interval values of the standard quality parameters and some typical network disturbances. Relevant information, necessary for assessment of the power distribution network quality level, can be provided by measurement and detailed analysis of the basic quality parameters at some specific locations in the power distribution network. Different types of devices and equipment for measurement and software based processing of the standard PQ parameters are available at the market. These instruments are developed to perform continuous monitoring of the power supply quality level at specific locations in power distribution networks. By measurement of the relevant quality parameters, which includes performing of software supported statistical and diagnostic activities in single or three-phase power distribution networks, these instruments are capable to verify compliance of the measured parameters with relevant PQ standards [3,4].

In order to satisfy specified parameters and measurement accuracy level, these measurement devices must be followed by an appropriate metrological traceability chain. Metrological verification and testing of such instruments must be performed inside appropriate metrological laboratories. Reference devices, such as voltage and current calibrators, are available in various functional and constructive solutions. Such voltage and current calibrators are sources of the reference test signals with high accuracy levels, which correspond to the secondary standards, laboratory and industrial standards in metrological traceability assurance chain. Also, there are special calibration instruments for some specific types of the PQ meters, such as solutions of the multifunctional calibrators Fluke 5520A and Fluke 6100B, supported by some special PQ calibration functions [5].

Acquisition system described in this paper is developed in LabVIEW [6] software package for generation of the reference voltage test signals, including special functions for simulation of the high-order harmonic disturbances, typical for real power distribution networks. Signal generation process is functionally based on the virtual instrumentation concept, which includes control software application and acquisition board PCIe 6343. Software support provides definition and presentation of the reference voltage waveforms. Number of the control functions, implemented on control panel, provides adjustment of the basic parameters for definition, presentation and signal generation. Test waveforms, defined directly from control block diagram according to requirements of the European PQ standard EN 50160, can be used for generation of the various test sequences. For this specific purpose are generated voltage test waveforms with some typical combinations of signal high-order harmonic disturbances. Generated voltage test waveforms are applied for testing of the three-phase PQ analyzer Fluke 435 Series II [7].

II. LABVIEW SOFTWARE SUPPORT OF THE PROCEDURE

Data acquisition system for generation of the standard PQ parameters and high-order harmonic disturbances is developed. It can generate test waveform sequences including all types of the PQ disturbances defined by European standard EN 50160.

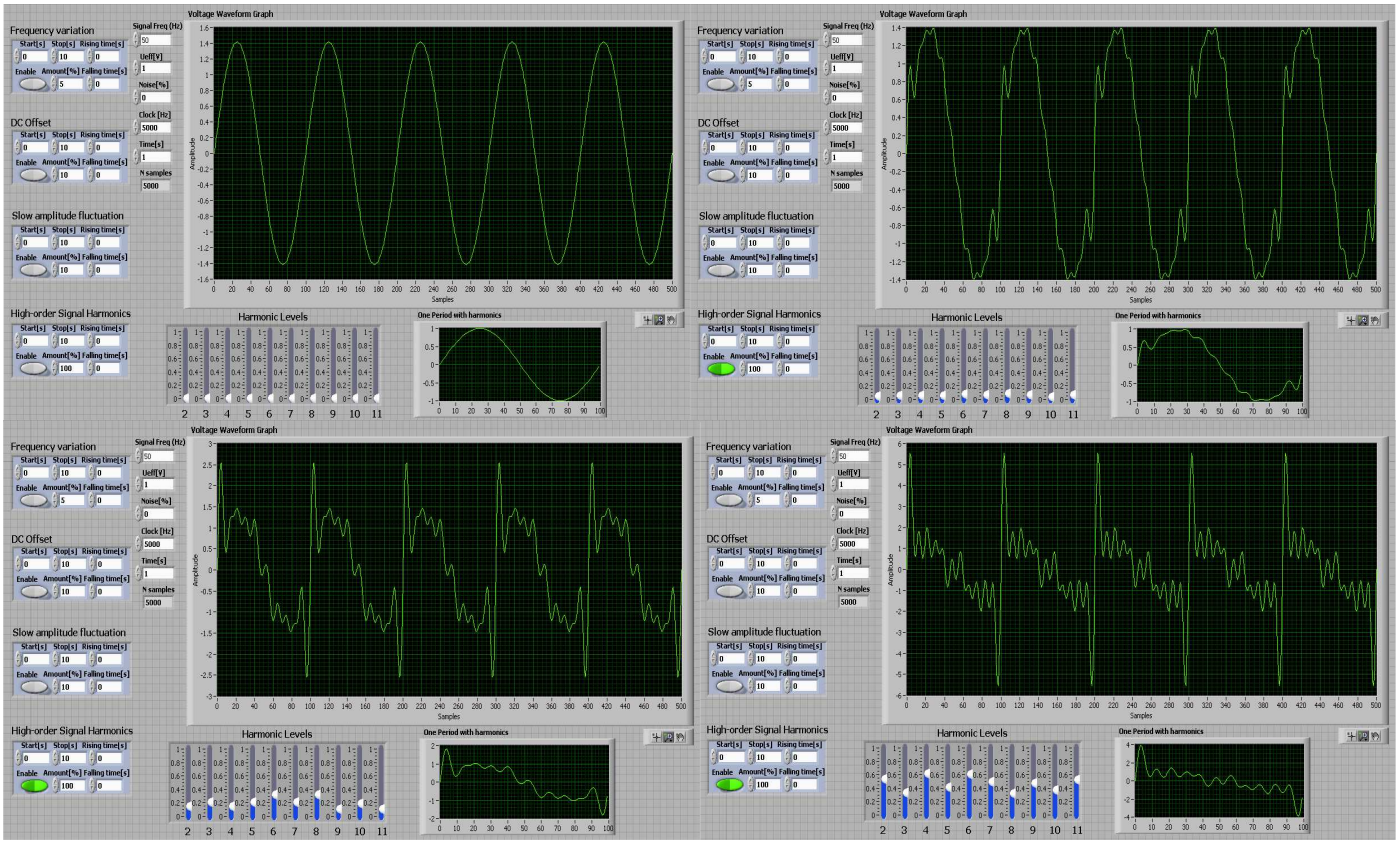


Figure 1. Control front panel in LabVIEW environment for generation of the voltage test waveforms (various levels of the high-order signal harmonics)

Main requirements were to enable combining a number of the various PQ disturbances in one complex signal test sequence, but also to include certain level of the noise, small voltage and frequency variations. Developed generation procedure includes standard computer with control programming application in LabVIEW software environment and data acquisition board NI 6343, equipped with standard block for connection SCB-68A. Entire procedure includes two connected functional part. First part of this procedure provides definition and simulation of the standard signal waveforms, with selected maximum levels of the specific high-order harmonic disturbances. Definition of the basic signal parameters for different types of the disturbances can be performed directly inside control front panel and block diagram of the LabVIEW virtual instrument [8]. Control front panel enables fast and simple correction of the basic waveform parameters, according to some specific user requirements and demands. Second functional part of this generation process is focused on real-time generation of the previously defined test voltage waveforms with standard harmonic disturbances, using analog outputs of the data acquisition board NI 6343. This is 32-channel PCIe acquisition card, with digital to analog signal conversion, output signal range of ±10V and 16-bit resolution.

Control front panel of the virtual instrument in LabVIEW software environment, developed for generation and graphical presentation of the voltage test waveforms, is shown in Fig. 1. This control front panel provides selection and variation of the basic signal parameters, by number of the control functions and knobs, implemented in block diagram of the virtual instrument. Some important functions are: definition of the nominal signal amplitude and frequency, definition of the signal sample rate and duration of the final test sequence, adding of the Gaussian noise in generated test signal, slow variation of the nominal signal frequency value, slow variation of the signal amplitude value with defined variation frequency and changing signal DC offset, voltage swell and voltage sag. Separated segment of the control functions on this front panel is used for selection and variation of the specific amplitude levels related to individual high-order harmonic components. Content of the specific high-order harmonics can be precisely determined by number of the control knobs for regulation of the harmonic amplitude levels. Shown test waveforms are generated with nominal frequency value of 50Hz and normalized RMS voltage value of 1V. Here are presented four different cases of the voltage test waveforms with various levels of the high-order signal harmonics. In order to be more realistic in generation of the signal waveforms, for particular harmonic disturbances is enabled separate definition of the nominal frequency variation, signal DC offset, amplitude fluctuations, start and stop times of the disturbances, rising and falling times and percentage amount of the harmonic amplitude levels from 2^{nd} to 11^{th} high-order signal harmonic components.

III. Presentation of Experimental Test Results

Previously described data acquisition system can be used as software based generator of the reference voltage waveforms, applicable in testing of various instruments for measurement and processing of the standard PQ parameters and high-order harmonic disturbances. As practical example, in this paper is described experimental procedure for testing of the three-phase

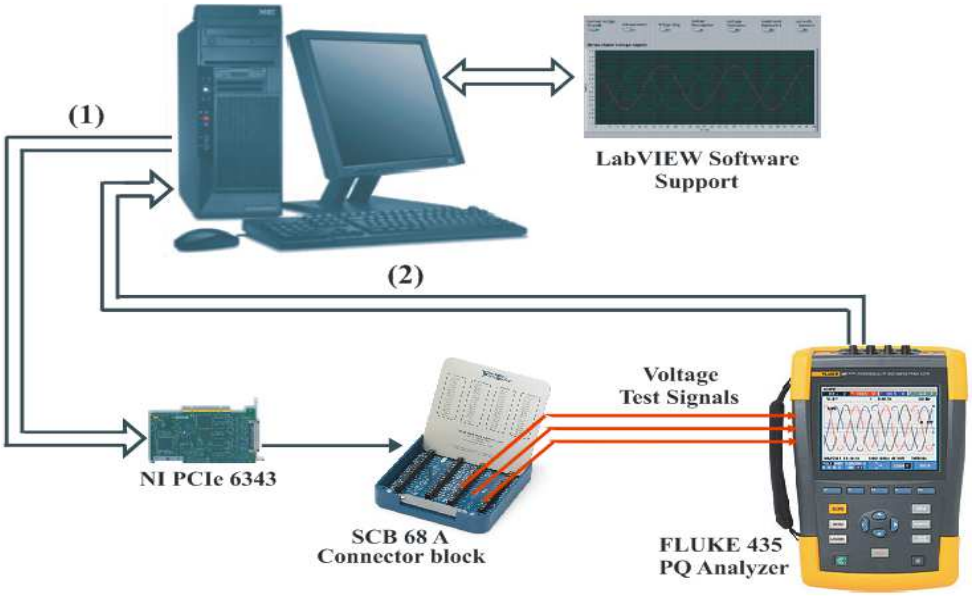


Figure 2. Basic hardware configuration of the software supported procedure applied to testing of three-phase PQ analyzer Fluke 435

PQ analyzer Fluke 435 Series II, in operational mode for high-order harmonics measurement. Basic hardware configuration of this test procedure is presented in Fig. 2. Reference voltage test signals, previously defined and generated with certain level of the harmonic disturbances, can be sent directly to the voltage inputs of the tested measurement instrument Fluke 435, using analog output channels of the data acquisition board NI PCIe 6343 and standard connector block SCB 68A (communication line 1). Instrument need to be set to work in operational mode for measurement of the high-order harmonics. Direct two-way communication between instrument outputs and computer is provided using standard USB interface (communication line 2). Obtained measurement results from instrument outputs can be simply transferred directly to computer, recorded into database and processed for graphical presentation. Some of experimental measurement data are shown in following section of this paper.

Software controlled test procedure includes three functional segments. First functional segment is recording of the voltage test waveforms on a graphical display of the tested instrument Fluke 435. Second part of this process includes measurement of the percentage values for individual high-order harmonic components in test waveforms. Third segment of the procedure is recording of the graphs for measured individual high-order signal harmonics, related to test signals. Four examples of the voltage test signals, recorded on graphical display of the tested analyzer Fluke 435, are shown in Fig. 3. These specific signals correspond to the test waveforms generated with various levels of the high-order disturbances, previously illustrated in Fig. 1.

Measurement results, related to percentage RMS values of the high-order harmonic components for test signals, obtained by measurement instrument Fluke 435, are presented in Fig. 4.

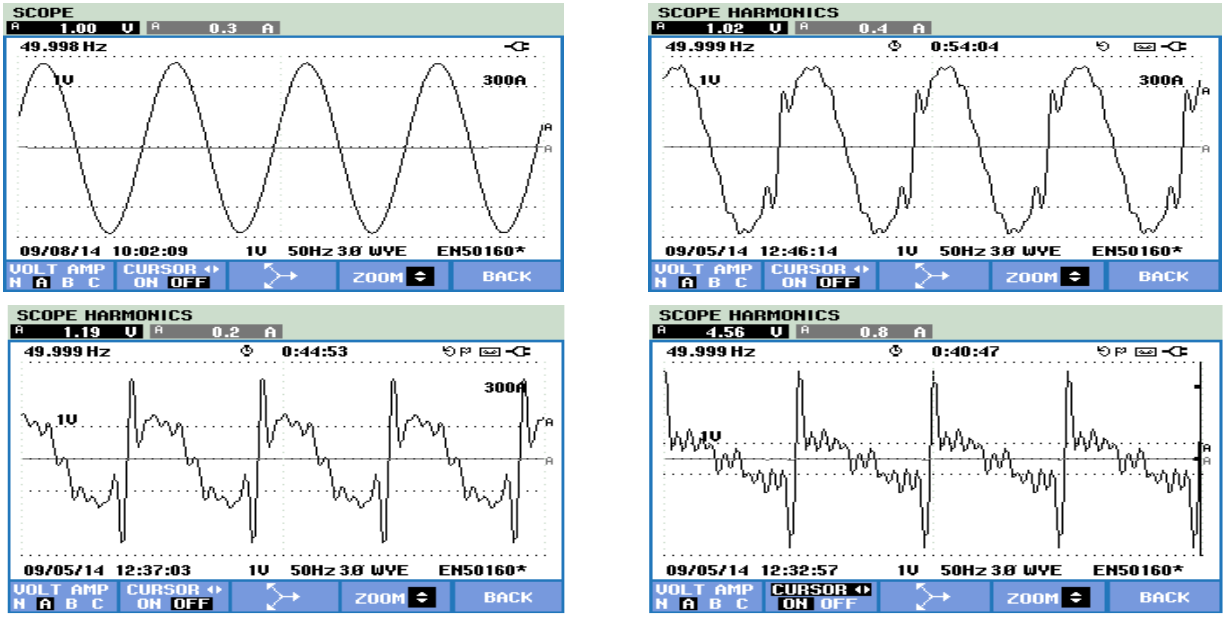


Figure 3. Test voltage waveforms recorded on graphical display of the measurement instrument - PQ analyzer Fluke 435

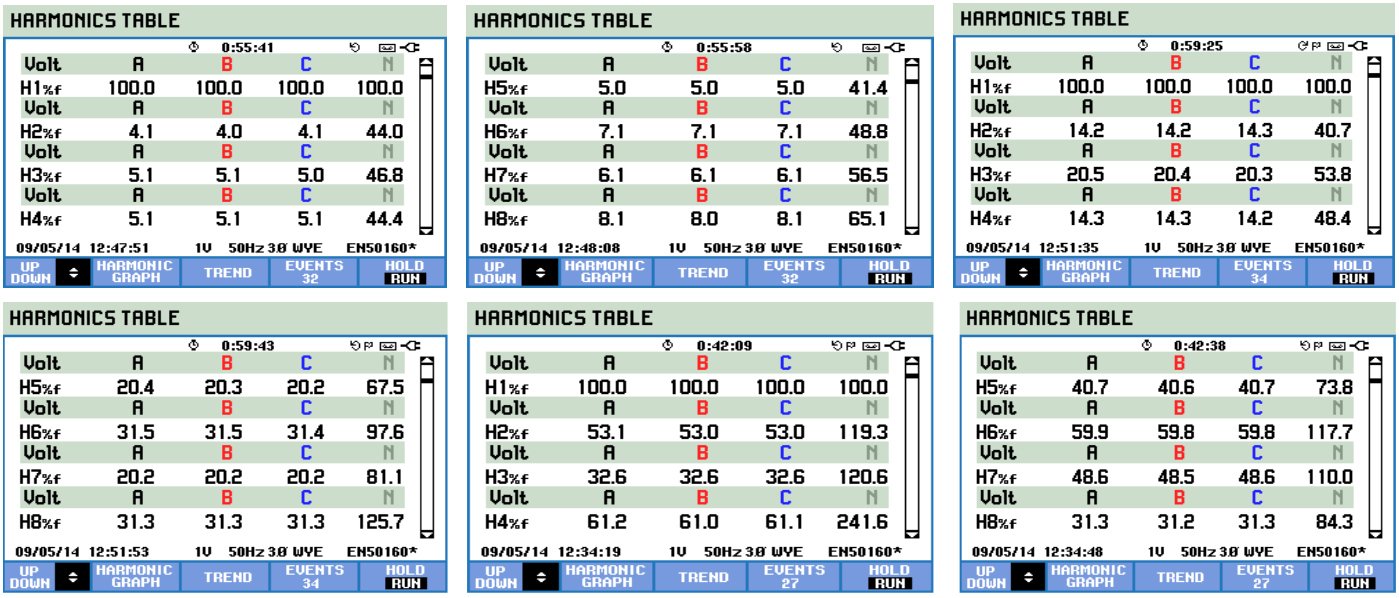


Figure 4. Percentage values of the high-order harmonics in test signals measured by instrument - PQ analyzer Fluke 435

Here are only shown measured percentage values of individual voltage harmonics from 2nd to 8th high-order harmonics, which includes three test voltage signals generated with various levels of the harmonic disturbances, previously presented in Fig. 3.

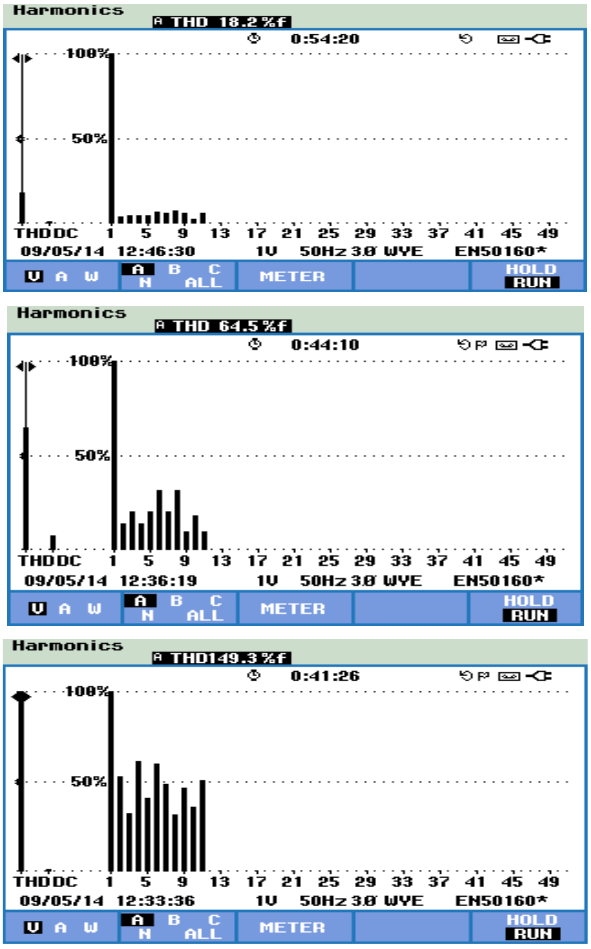


Figure 5. Recorded graphs of the measured high-order signal harmonics

Graphs of the measured percentage RMS values for 2nd to 11th high-order harmonics, recorded on a graphical display of the tested instrument for presented three test waveforms with certain level of the harmonic disturbances, are shown in Fig. 5. These graphs indicate maximum measured percentage values of individual high-order voltage harmonics in test waveforms. Percentage harmonic levels on instrument display, obtained by measurement procedure, fully correspond to defined harmonic values in test signal waveforms previously presented in Fig. 1.

Generally, developed LabVIEW based acquisition system for generation of the reference voltage signals is applicable in testing of the various instruments and equipment designed for measurement of the PQ parameters and network disturbances, defined according to relevant international quality standards.

IV. Conclusion

Acquisition system for generation of the reference voltage test waveforms with standard high-order harmonic disturbances is described in this paper. This software supported solution for signal generation, functionally based on LabVIEW application software, includes standard computer and acquisition board NI PCIe 6343. Basic purpose of this system is to provide standard signals applicable in testing of instruments for measurement of basic electrical power quality parameters and high-order signal harmonics, according to European quality standard EN 50160. Specific amplitude levels of the high-order harmonics in test voltage waveforms can be individually defined and generated. Continuous variation of the percentage amplitude level in input control section can be performed separately for each individual harmonic component or at the same time for all selected high-order harmonic components. Developed solution is practically verified by software controlled procedure applied to testing of instrument for measurement and processing of the high-order harmonic components – three-phase PQ analyzer Fluke 435 Series II. Some test waveforms and experimental measurement data, obtained by testing procedure, are presented in the paper.

ACKNOWLEDGMENT

This work was supported by national scientific project, with reference number TR 32019, sponsored by Serbian Ministry of Education, Science and Technological Development.

REFERENCES

[1] E. F. Fuchs, and M.A.S. Masoum, Power Quality in Power Systems and Electrical Machines, Academic Press, USA, February 2008.

[2] Power Quality Application Guide, Voltage Disturbances, Standard EN 50160, Copper Development Association, 2004.

[3] L. F. Auler, and R. D'Amore, "Power quality monitoring controlled through low-cost modules," IEEE Transactions on Instrumentation and Measurement, vol. 58, no. 3, pp. 557–562, 2009.

[4] B. D'Apice, C. Landi, A. Pelvio, and N. Rignano, "A multi-DSP based instrument for real-time energy and PQ measurements," Metrology and Measurement Systems, vol. 14, no. 4, pp. 495–506, 2007.

[5] 5520A-PQ Power Quality Option for the 5520A Calibrator – Technical Data, Fluke Corporation, Test Equipment Depot, Melrose, USA, 2007.

[6] LabVIEW User Manual, National Instruments Corporation, 2012.

[7] Fluke 435 Series II - Three-phase Energy and Power Quality Analyzer Users Manual, Fluke Corporation, January 2012.

[8] M. Simić, D. Denić, D. Živanović, D. Taskovski, and V. Dimcev, "Development of a data acquisition system for the testing and verification of electrical power quality meters," JPE - Journal of Power Electronics, vol. 12, no. 5, pp. 813-820, September 2012.

Hardware realisation of data logger system for inland excess water

Đorđe Obradović
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
djordjeo@uns.ac.rs

Miodrag Brkić
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
brkic.miodrag@uns.ac.rs

Viktor Dogan
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
dr.wigan@gmail.com

Miloš Živanov
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
zivnov@gmail.com

Barta Karoly
Természeti Földrajzi és Geoinformatikai Tanszék
SZTE, University of Szeged
Szeged, Hungary
barta@geo.u-szeged.hu

***Abstract*—Inland excess water cause considerable economic, social and environmental problems. In northern parts of Vojovodina, where inland excess waters occurs regularly, continuous measurement of groundwater level on several measurement points over long periods of time is needed for realization of hydrodynamic models. With data collecting system management innovative geographic information methods and observation techniques these models can be developed to predict inland excess water. In this paper hardware realization of data logger is described in details. Data logger system was developed and used on locations of interest in northern Vojvodina.**

Keywords- logger; measurement; inland excess water;

I. INTRODUCTION

Inland excess waters cause numerous problems in Vojvodina. As a predominantly agricultural region, the problems are related to obstruction of agricultural activities on the fields as well as causing damages to the crops (rotting, occurrence of fungal infections due to increased moisture and other). Beside these problems inland excess waters cause damage to houses, buildings, flooding from sewer pits and canals, soil contamination, impeding of local traffic and transport [1].

The occurrence of inland excess water in Vojvodina is the consequence of numerous natural and anthropogenic factors. The following are the most important: lower position of the endangered surfaces relative to the main river of the catchment; morphology and terrain inclination (predominantly flat area), the vicinity of rivers and canals; climatic conditions (successive rainy years and uneven yearly distribution of the precipitation); shallow first aquifer and impervious layer; reduced flow on watercourses.

This system is developed for continuous measurement of groundwater level on several measurement points over long periods of time (several years), which is needed as input data for hydrodynamic models. With this model some excess water types could be forecasted, and actions could be taken [2].

In this paper acoustic measurement is used for determination of water level in wells. Acoustic measurement was chosen over industry solutions because of its low cost and small dimensions of the sensors. This type of measurement is very similar to ultrasonic, but lower frequencies of measurement signal were chosen because lowering the frequency decreases reflection from the walls of the well .

II. DATA LOGGER SYSTEM

Data logger system is shown in **Fig**. 1. To measure level of inland water, a series of wells are drilled on locations where excess water can appear. Measurement system is set at the top of the well, and by measuring time measure time of travel of sound to and from level of water inside wells information of level of groundwater is obtained [3].

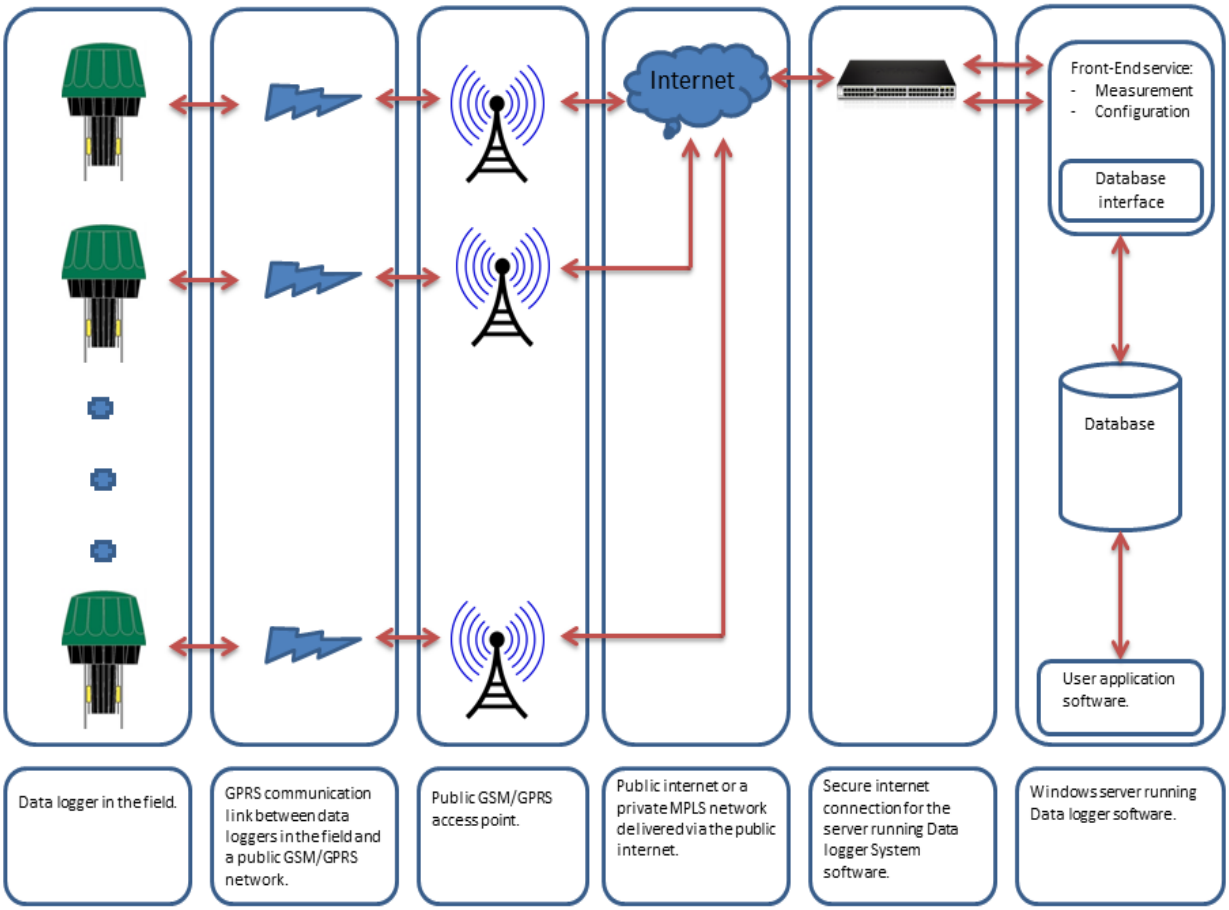


Figure 1. Data logger system

III. HARDWARE REALIZATION

Data logger consists of several electronic subsystems :

- Transmitter and receiver electronic for acoustic measurement
- MCU electronic for control of data logger
- Power electronic for stable power supply

A. *Transmitter electronics*

Schematic of receiver electronics is shown in **Fig**. 2. KPEG-272 buzzer is used as transmitter. Since this buzzes is working with voltages up to 18V, and battery ih this system is around 3.6V, step up "boost" converter is used to supply buzzer with higher voltage of 18V, achieving higher power of audio signal. Step up convertor is realized with U1 (TPS61040) and auxiliary passive components. U7 (UCC27524) is mosfet driver, used in this circuit as switch to control power to buzzer.

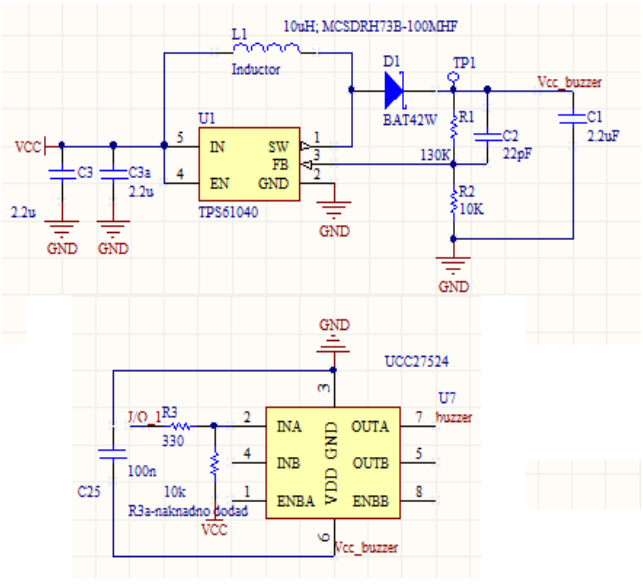


Figure 2. Schematic of transmitter electronics

B. *Receiver electronics*

Schematic of receiver electronics is displayed in **Fig**.3. Microphone ABM-713 is used as receiver. Since electret microphone is used, DC biasing is needed, realized with resistors R5, R6. operational amplifier (OP) MCP60022 is used as buffer and preamplifier (U2). Passive band pass filter consisted of R13, C19, C12, C1, R15 is used for first level of filtering, at frequencies of 4KHz. Additional 6-pole band pass filters is used for filtering, this level of filtering is needed since receiver microphone is wide range with bandwidth of 100Hz-20KHz. For software control of the amplification U2B and U5 is used, U2B is used as inverting amplifier whereas U5 (AD5234) is digital potentiometer connected in feedback loop of U2B. Resistance of U5 can be changed in range from 0 to 100Kohm which changes amplification of this amplifier, SPI pins of U5 connected to main microcontroller which controls its value. Voltage reference 3312 is used to generate voltage of 1,25V which is used as virtual ground in all amplifiers in this circuits, so there is no need for negative voltage supply.

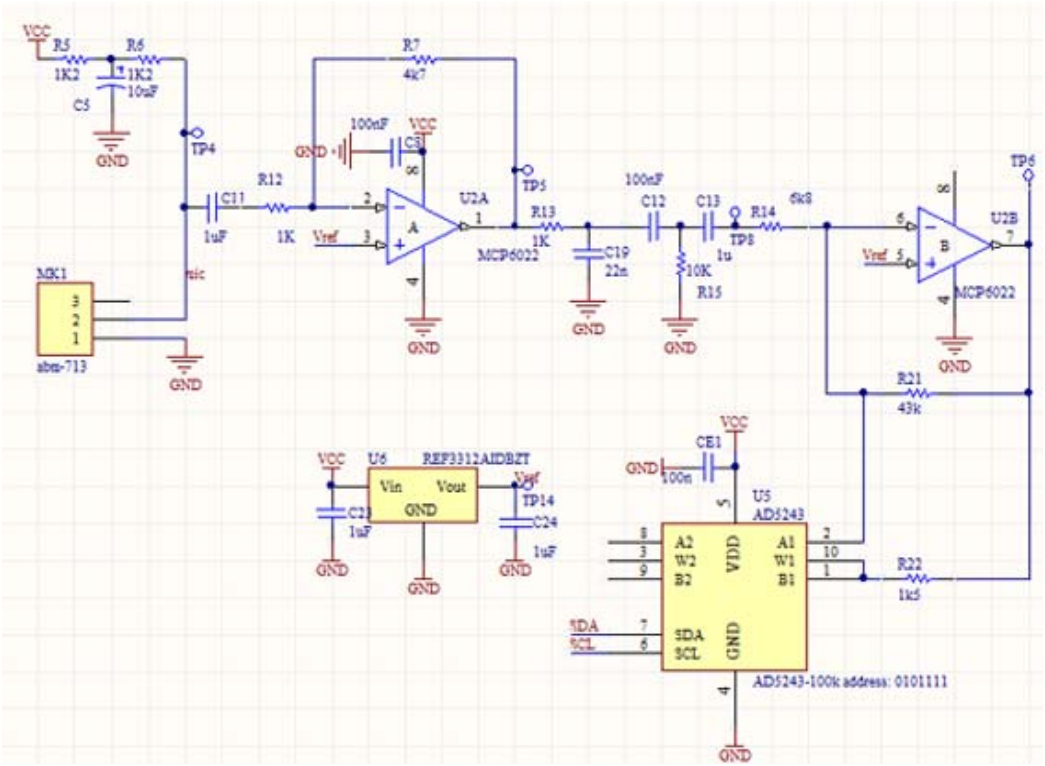


Figure 3. Schematic of receiver electronics

A measuring frequency of 4KHz is selected for measurement signal, this frequency is selected as optimum for this system after testing of well pipes with wide range of measurements signals. When signal with higher frequency was used, significant reflections from joints of pipes appeared making it impossible to distinguish them from useful signals. For lower frequencies received signals had tendency of uniform amplitude, making it hard to precisely measure the travelling time of measurement signal, which significantly decreased measurement resolution of the system.

IV. MCU BLOCK

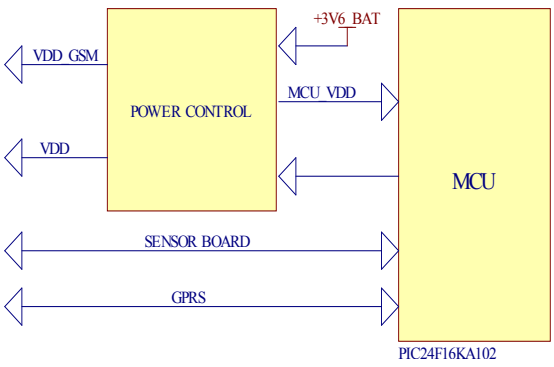


Figure 4. MCU block

MCU block is shown in **Fig**.4. It has several functions:

- To communicate with GPRS module by sending and receiving data over serial port.
- To control the transmitter on sensor board, and to perform A/D conversion of receiver signal
- To control power for the rest of the circuitry

MCU used in this system is PIC24F16KA102, extreme low power consumption microcontroller. It has the ability to go into "deep sleep" mode, where current consumption is under 1uA. This MCU has internal 10-bit ADC converter, more than adequate for this system.

In order to minimize consumption, and so prolong battery life, between measurements the power for all the rest of the electronics, including the modem and the sensor block is completely shut down. By using the power control circuit shown in **Fig**. 5, power supply to other blocks is allowed only when necessary.

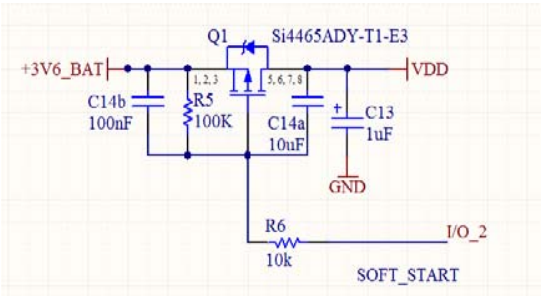


Figure 5. Schematic of power control circuit

Si4465 mosfet is used as a power switch because of its low on resistance and low leakage current [4]. C14a and R6 provide a slower turn on to limit the initial inrush current

Battery used in the system is 3,6V, so it can directly supply the MCU. Also the MCU can be supplied from a usb/serial converter which is occasionally used to read data from the controller in the debugging process. Having this external supply attached, it is possible to take the battery out and change it without losing real-time clock data. Also, during the modem operation, the boosted modem power supply VDD_GSM is brought to the MCU to additionally backup the battery and prevent controller resets due to voltage drops caused by current surges of the GPRS module. The multiple power sources are automatically managed by a sort of a wired OR diode circuit which is shown in **Fig**. 6.

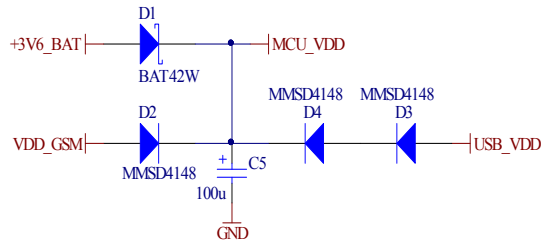


Figure 6. Schematic of wired OR power suppy managing circuit

V. POWER SUPPLY

Schematic of power supply is shown in **Fig**. 7. Power supply used in this circuit is step-up convertor which enables stable voltage of 4V for whole system. Since battery used for this system are Lithium Thionyl Chloride type with voltage of 3,6 V, value which is adequate for all electronic elements in system, there would be no need for additional electronics in power supply if voltage level for type of batteries was stable enough. Output voltage of this batteries significantly depends on temperature, and when external temperatures falls below zero, voltage that this battery gives is around 3V, which is not sufficient for GSM module, which input voltage is between 3.4V and 4.2V. For this reason step-up convertor U1 (ADP1612) and auxiliary passive components are used, with maximum current of 3A, needed for GSM module when transmitting. ADP1612 is selected because of its ability of working properly with minimum voltage difference in input and output voltage. High-quality capacitors C1,C3 and C9 were used to reduce voltage peaks from switching power supply, to minimize its influence on measurement signal.

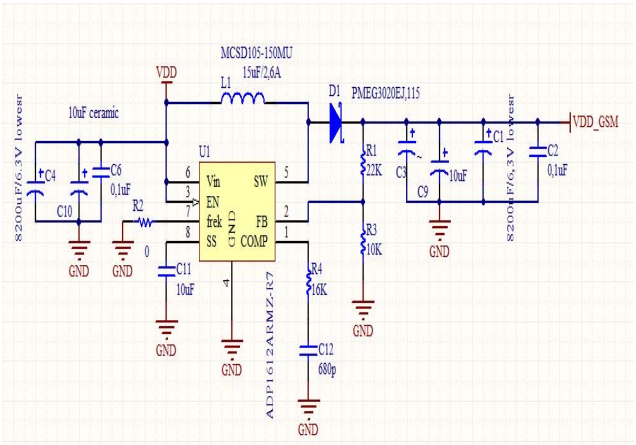


Figure 7. Schematic of power supply

VI. MEASUREMENT RESULTS

Data loggers were realized, tested in laboratory conditions and then placed in 25 locations in north Vojvodina. Currently they are in use for several months, measurement are made once a day and send via GSM network once a week, to extend the battery life. In **Fig**. 8. measurement results are shown with graph display on every location. Secure internet connection for the server is running data logger system software realized as web based application. This application serves as database for all measurement, and also as control interface for configuration of data logger parameters.

Monitoring system was tested in several wells, measurement error of the system was up to 0.7%, which is adequate for this application. Measurement resolution of 1 cm was obtained, with measurement range of 15 meters. Similar measurement error were obtained in [5], and industrial solutions [6] have slightly better results, but these

solutions are realized with ultrasonic measurement signals which cannot be used in relative small plastic pipe because signals would be reflected from walls of the pipe giving false measurements.

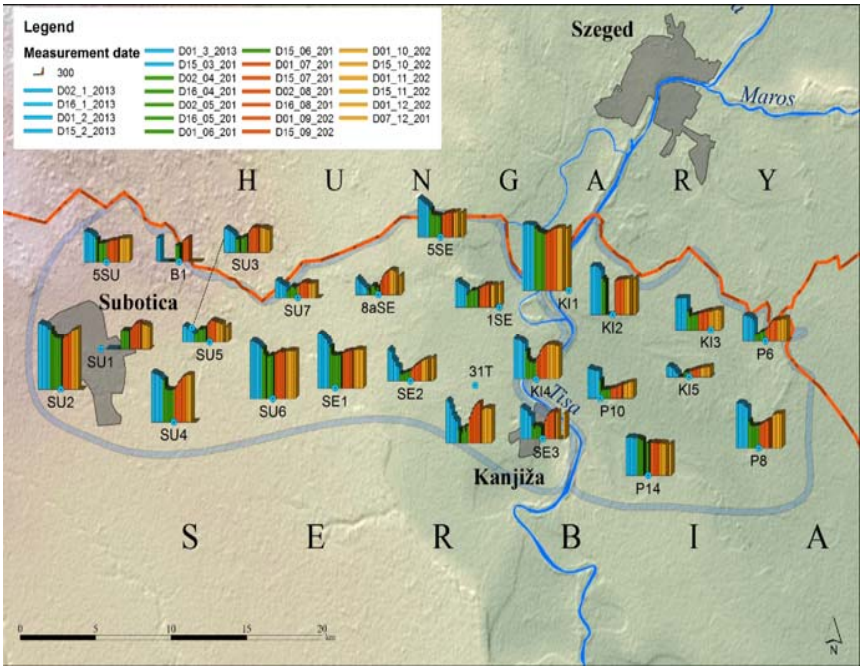


Figure 8. Measurement results

VII. CONCLUSION

In this paper hardware realization of data logger system is presented. Schematic of all elements of the system are shown and described in details. This system was tested in several wells, is realized as low-cost and robust system for measurement of level of groundwater. Using GSM network for sending data and battery for power supply, system is realized as stand alone, with autonomy of around one year, when batteries needed to be replaced.

Measurement result from this system are used for prediction of level of groundwater in region of northern Vojvodina [6].

This monitoring system can be easily adapted to be used in wide range of environmental and other measurements. With addition of solar power supply, it can be completely autonomous for several years on every location with GSM coverage.

ACKNOWLEDGMENTS

This work has been financed by EU IPA cross-border cooperation project “Measurement, monitoring, management and Risk assessment of inland excess water in South-East Hungary and North Serbia”, number: HUSRB/1002/121/088 , and supported by the Ministry of Education and Science, Republic of Serbia, under projects III43008

REFERENCES

[1] Pavić D., Mészáros M., Dolinaj D., Savić S., Obradović D., Brkić M., Živanov M., 2013, Inland excess water in Vojvodina, (Serbia) – innovative methods in cross border research for an old, common problem, In: Dr Lóki József (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában IV. Konferencia helye, ideje: Debrecen, Magyarország, 2013.05.23-2013.05.24., Debrecen: Debreceni Egyetemi Kiadó, pp. 89-96. (ISBN:978-963-318-334-2)

[2] Barta K., 2013: Inland Excess Water Projection based on Meteorological and Pedological Monitoring Data on a Study Area Located in the Southern Part of The Great Hungarian Plain. Journal of Environmental Geography 6 (3–4), 31–37.

[3] Miodrag Brkic, Djordje Obradovic, Viktor Dogan, Barta Károly, and Milos Zivanov: “Acoustic Measurement and Monitoring System for Level of Groundwater“,SENSOR LETTERS Vol. 11, Number 9, 1627–1631, 2013 ISSN: 1546-198X, American Scientific Publishers

[4] Si4465 datasheet - www.vishay.com/docs/73856/73856.pdf

[5] A New Ultrasonic-based device for Accurate Measurement of Oil, Emulsion, and Water Levels in Oil Tanks, MAHMOUD MERIBOUT,MOHAMED HABLI, AHMED AL-NAAMANYand KHAMIS AL-BUSAIDI, Instrumentation and Measurement Technology Conference, 2004. IMTC 04. Proceedings of the 21st IEEE (Volume:3), 1942 - 1947 Vol.3, ISSN :1091-5281, 18-20 May 2004

[6] http://www2.emersonprocess.com/enus/brands/rosemount/level/ultrasonic-transmitters/3100-series-ultrasonic/pages/index.aspx

[7] Van Leeuwen B., Henits L., Mészáros M., Tobak Z., Szatmári J., Pavić D., Savić S., Dolinaj D., 2013, Classification methods for inland excess water modelling, Journal of Environmental Geography 6:(1-2) pp. 1-11.

Application of Digital Stochastic Measurement of Definite Integral Product of Two or More Signals Using Two-Bit A/D Converter

B. Ličina, D. Pejić, P. Sovilj, V. Vujičić
Department of Power, Electronics and Communication Engineering
Faculty of Technical Sciences, University of Novi Sad
Trg Dositeja Obradovica 6, 21000, Novi Sad, Serbia
boris.licina@yahoo.com, pejicdra@uns.ac.rs, platon@uns.ac.rs, vujicicv@uns.ac.rs

***Abstract* — Commonly used strategy in discrete digital measurements is to capture digital value of signal's magnitude at a chosen time instant. The mathematics in the background of this strategy is algebra, while the applied theory is the Theory of discrete signals and systems. An alternative measurement strategy, named "measurement over an interval", has been researched in three challenging areas: (i) measurement of fast-changing signals, (ii) measurement of noisy signals, and (iii) measurement that requires high accuracy and linearity. Numerous simulations, experiments and developed measurement instruments have proven the engineering/metrological applicability of this "measurement over an interval" strategy. This paper presents application of digital stochastic measurement over interval of the finite integral product of two or more signals using two-bit A/D converter. Error of this method is shown through a large number of simulations.**

Keywords *– A/D conversion, Digital measurements, Probability, Stochastic processes.*

I. INTRODUCTION

Nowadays, the term „measurement" is mostly considered as discrete digital measurement (sampling method measurement). This commonly used strategy called "measurement in a point" is to capture digital value of signal's magnitude at a chosen time instant. Digital representation of the signal's parameters at that time instant has to be characterized with maximum accuracy (theoretically - with an ideal accuracy). In order to capture full information on the measured quantity, all conditions of the sampling theory must be satisfied. The mathematics that explains this approach is algebra, while the applied theory is the Theory of discrete signals and systems. This measurement strategy has been the backbone of the development in metrology, control, telecommunications, etc. In the conversion process, accuracy and speed are opposing requirements. Accurate measurements of low-level, noisy and distorted signals have been a challenging problem in the theory and practice of measurement science and technology. Since 1956 [1], the possibility for reliable operation of instruments with inherent random error has been researched. It has been shown that adding a random uniform dither to the input signal prior quantization, decouples the measurement error from the input signal [2].

This work was supported in part by the Provincial Secretariat for Science and Technological Development of Autonomous Province of Vojvodina (Republic of Serbia) under research grant No. 114-451-2723, and supported in part by the Ministry of Science and Technological Development of the Republic of Serbia under research grant No. TR32019

Measurement strategy "measurement over an interval" formulated in [3] has clear advantages over standard "measurement in a point" approach in three challenging areas:

(*i*) measurement of fast-changing signals,

(*ii*) measurement of noisy signals,

(*iii*) measurement that requires high accuracy and linearity.

The third case, which is of the greatest importance in metrology, is elaborated in [4]. Measurement over an interval entails a very simple hardware (flash A/D converter as the simplest and the fastest, but the least precise device) hence lowering the number of systematic error sources (it is well known that the flash A/D converter hardware is being doubled with each new bit of the resolution thus doubling a systematic error sources). The sampling method has two inherent sources of systematic measurement error: discretization in time and discretization in value. If the sampling theorem conditions are satisfied, discretization in time is eliminated as a source of measurement error. Discretization in value always causes a systematic measurement error – it cannot be eliminated. In [4] it is shown, under certain conditions, how to reduce it and keep within acceptable range.

Measurement over an interval is an integral approach to measure a signal and its parameters - a signal [5] or some of its parameters [4] are measured during a finite time interval of an arbitrary duration. It has been shown that a singular measurement in every instant does not need to be maximally accurate, while the measurement uncertainty is reduced by adding a random uniform dither. Mathematical tools utilized in this case belong to mathematical analysis, probability theory, statistics and sampling theory. From theoretical point of view, the problem is highly non-linear and stochastic, and therefore neither the standard linear Theory of discrete signals and systems nor the Theory of random processes can be applied. It was necessary to develop an alternative mathematical approach, and this approach has been developed based on Central limit theorem. The most commonly we use measurement devices in greed measurement: measurement of true RMS of current and voltage, as well as power and energy, also with some derived values is taking place (cos φ, distortion factor...).

II. PRINCIPLE OF MEASUREMENT

Fig. 1 shows the principle of adding a uniform random dither h to the measured signal y. The role of the dither is to decouple the quantization error of the uniform quantizer from the input signal [2], and thus enabling accurate measurements using a low-resolution converter [4]. Ψ is output of the uniform quantizer of the digital multiplier.

Figure 1. Block diagram of application of a uniform random dither h to the measured signal y.

For the subsequent discussion let us assume that the following conditions are satisfied:

$$|y| \le R,\ R = Z \cdot a,\ |h| \le \frac{a}{2},\ |y+h| \le R + \frac{a}{2} \qquad (1)$$

where range of the input quantity is labeled with R, number of positive quantization levels with Z and quantum of uniform quantizer with a. A probability density function (PDF) of uniform random distribution dither signal h is $p(h) = \frac{1}{a}$ for $|h| \le \frac{a}{2}$. We can interpret it graphically as in Fig. 2.

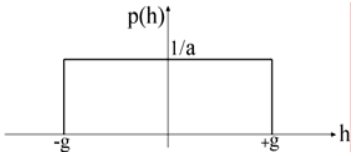


Figure 2. PDF of the uniform random dither signal h.

Voltage ranges and decision thresholds associated with process of measuring average input signal by uniform quantizer are represented graphically in Fig. 3:

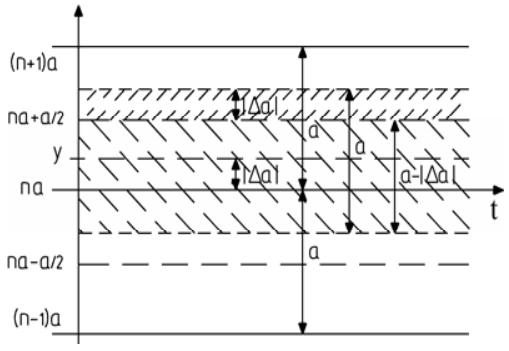


Figure 3. Voltage ranges and decision thresholds associated with process of measuring $\overline{\Psi}$.

For $Z = 1$, the A/D converter is very simple, with the minimal hardware structure as shown in Fig. 4.

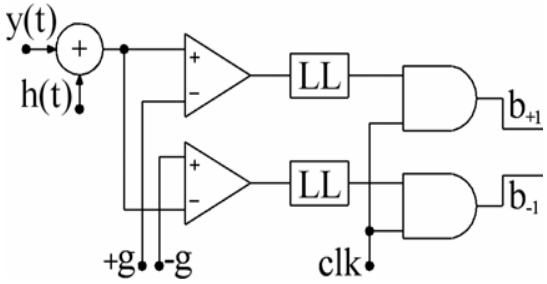


Figure 4. Two-bit flash A/D converter with an additive uniform dither.

For the device in Fig. 4 the quantum $a = 2g$ i.e.

$$g = \frac{a}{2},\quad R = a = 2g,\quad p(h) = \frac{1}{2g} \qquad (2)$$

Possible values of Ψ are $\Psi \in \{-2g, 0, 2g\}$, and the analytical term for Ψ is:

$$\Psi = 2g \cdot (b_1 - b_{-1}) \qquad (3)$$

where $b_1, b_{-1} \in \{0,1\}$ and $b_1 \cdot b_{-1} = 0$. It is never possible that b_1 and b_{-1} are equal to 1 simultaneously - it would have meant that $y \ge 0$ and $y \le 0$ simultaneously.

III. DIGITAL STOCHASTIC MEASUREMENT OVER INTERVAL OF THE FINITE INTEGRAL OF TWO SIGNALS PRODUCT

Device from Fig. 5 has 2 two-bit flash A/D converters shown on Fig. 4, with inputs $y_1 = f_1(t)$ and its uncorrelated signal h_1, as well as $y_2 = f_2(t)$ and its uncorrelated signal h_2, respectively. h_1 and h_2 are mutually uncorrelated random uniform dither signals. Outputs Ψ_1 and Ψ_2 are passed to a multiplier; the multiplier output is $\Psi = \Psi_1 \cdot \Psi_2$, and it can assume values: $\Psi \in \{-(2g)^2, 0, +(2g)^2\}$.

If, during one measurement interval, N A/D conversions are performed by each A/D converter, then the accumulator from Fig. 5 accumulates the sum of N subsequent multiplier's outputs: $\sum_{i=1}^{N} \Psi_1(i) \cdot \Psi_2(i)$. This accumulation can be simply used for calculation of the average value of the multiplier output $\overline{\Psi}$ over the measurement interval as:

$$\overline{\Psi} = \frac{1}{N} \cdot \sum_{i=1}^{N} (\Psi_1(i) \cdot \Psi_2(i)) \qquad (4)$$

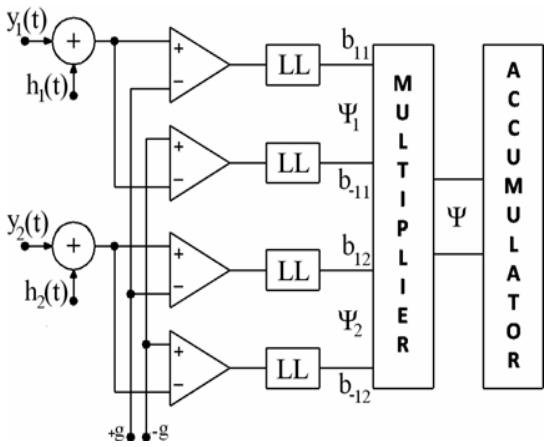


Figure 5. Structure of the device, based on two-bit flash A/D converters, for measurement of the finite integral. of two signals product

For Measurement purpose real input signals are being conditioned and thus reduced to appropriate voltage level ±5 V (U_{max} = +5 V, U_{min} = -5 V). On that input signal a random voltage signal is now superimposed. Random voltage signal is dieter with level ±2.5 V.

A. Application in the measurement of true RMS

To analyze measurement of signal true RMS we will observe two wave forms: simply periodic signal (sinusoidal signal with frequency 50 Hz) and typical complex periodic signal (jagged signal with same frequency).

Simulation is done with signal amplitude of 0.5 V, 1 V, 1.5 V, 2 V, 2.5 V, 3 V, 3.5 V, 4 V, 4.5 V and 5 V, respectively. For each amplitude of the input signal we perform a series of 100 measurements with different duration: 20 ms, 40 ms, 60 ms, 80 ms, 100 ms, 200 ms, 500 ms, 1 s, 2 s, 5 s, 10 s, 20 s, 30 s, 1 min, 2 min, 5 min, 10 min and 15 min respectively.

To calculate true RMS of individual measurement it is necessary to multiply U_{max} with the square root of the mean output. Now we can start calculating individual measurement errors as:

- absolute error = measured value - true value
- relative error = absolute error / true value · 100%
- errors relative to full scale (FS) = absolute error / ($U_{max} - U_{min}$) · 100%

After we do these calculations for each 100 measurements group we calculate mean and standard deviation for both relative error and error relative to FS (absolute error does not give us any important information about the accuracy nor precision of measurement).

Fig. 6 shows error diagrams for measurements of true RMS of simply periodic (sinusoidal) signal for error in relation to the FS.

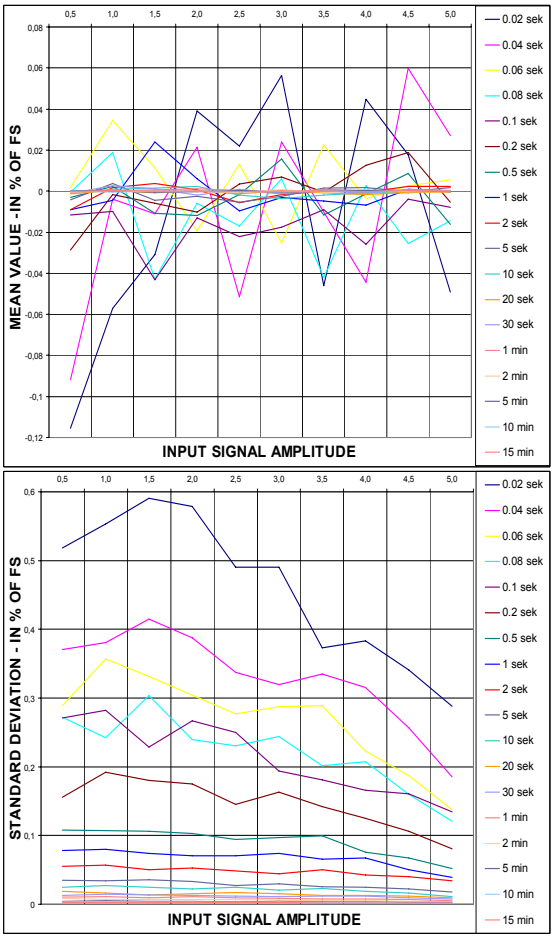


Figure 6. Measurement error of true RMS of simply periodic signal for error in relation of FS.

The theoretical error [6] is presented by expression:

$$\sigma_e^2 = \frac{(2g)^2}{t_2 - t_1} \cdot \int_{t_1}^{t_2} |f_1(t) f_2(t)| dt - \frac{1}{t_2 - t_1} \cdot \int_{t_1}^{t_2} f_1(t)^2 f_2(t)^2 dt \quad (5)$$

Simulation results confirm theoretically expected error. It is important to note that this method requires accurate voltages of two thresholds (+ 2.5V and -2,5V) and uniform dither signal distribution. If we compare this method with classical digital measurement approach, it can be calculated that 14-bit A/D converter should be used in digital measurement approach for obtaining the error of presented method.

Fig. 7 shows error diagrams for measurement of true RMS of complex periodic (jagged) signal for error in relation to the FS.

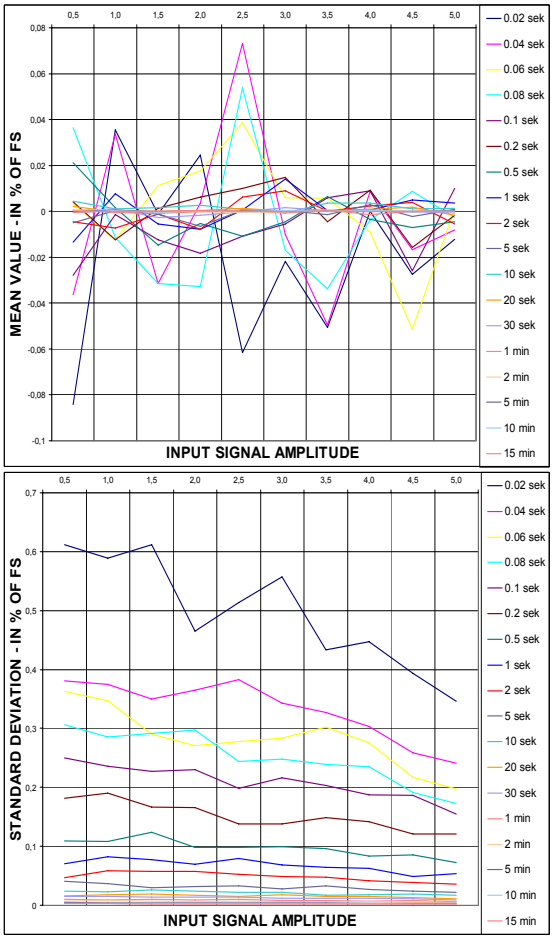


Figure 7. Measurement error of true RMS of complex periodic signal for error in relation of FS.

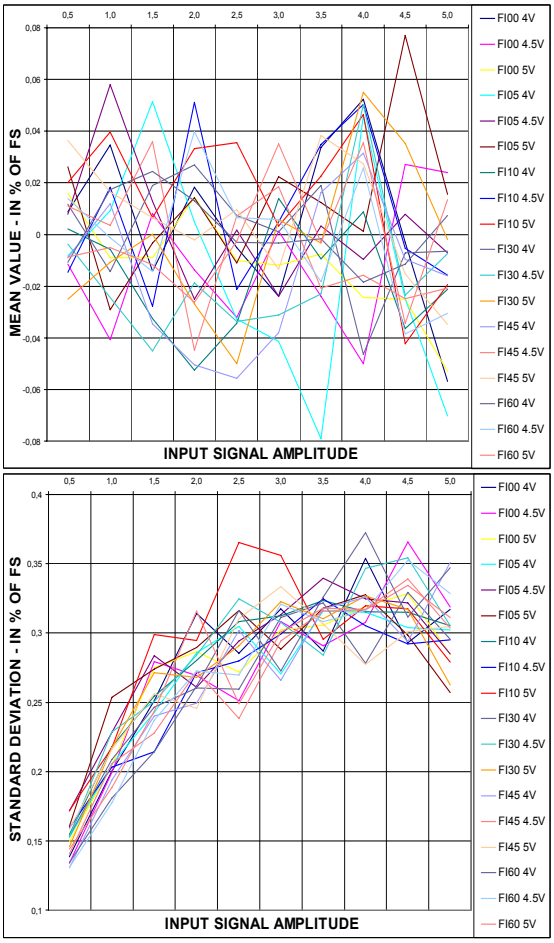


Figure 8. Measurement error of power of simply periodic signal relative to the FS.

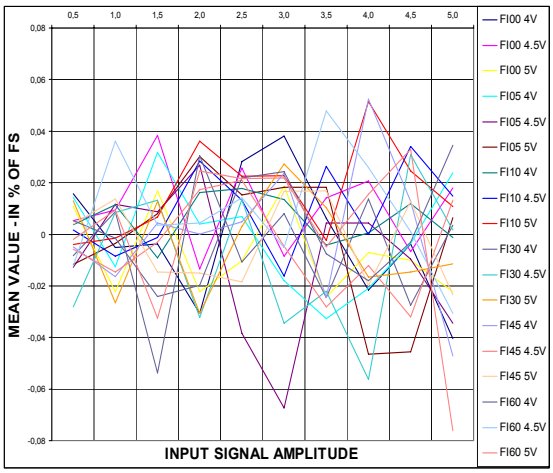
B. *Application in the measurement of power*

In the analysis of the signal strength measurement will observe the same waveforms as in the measurement of the true RMS. Unlike the measurement of the true RMS where we had one input, the strength measurement we do on two input signals. One of them represents the voltage signal, and the other one represents a current signal. In grid measurements voltage signal is approximately constant therefore we try to scale it in the "upper zone" in order to have less measurement error, so the simulation is also done for the amplitude of 4 V, 4.5 V, and 5 V. For each of these signals conditioned current signal takes a value of 0.5 V, 1 V, 1.5 V, 2 V, 2.5 V, 3 V, 3.5 V, 4 V, 4.5 V and 5 V respectively. In addition for each combination of power measurements we do simulation for phase shifts of $\pi/36$ (5°), $\pi/18$ (10°), $\pi/6$ (30°), $\pi/4$ (45°), $\pi/3$ (60°). The duration of each measurement is 1 s.

To calculate power of individual measurement it is necessary to multiply $U_{\max}^2$ with the mean output. Now we can start calculate individual measurement errors like we have done in the measurement of the true RMS.

Fig. 8 shows error diagrams for measurements of power of simply periodic (sinusoidal) signal for error relative to the FS.

Fig. 9 shows error diagrams for measurement of power of complex periodic (jagged) signal for error relative to the FS.



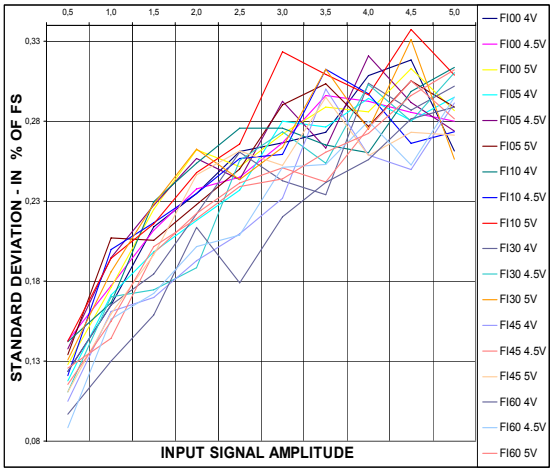


Figure 9. Measurement error of power of complex periodic signal relative to the FS.

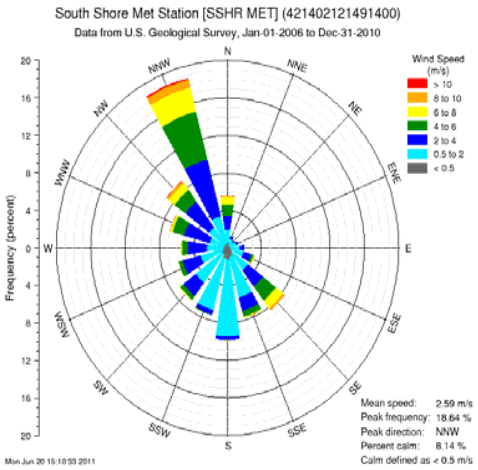


Figure 10. Wind rose [9].

IV. DIGITAL STOCHASTIC MEASUREMENT OVER INTERVAL OF THE FINITE INTEGRAL OF MORE THAN TWO SIGNALS PRODUCT

In all so far published articles, of which some are listed in references [3-6] it is processed application of devices on 2 signal product. Each of the articles includes the theoretical part. Some articles include simulation and description of the prototype device. However neither article describes application for product of more than two signals beside theoretical conclusions. To show a practical implementation of the method for the measurement of the product of more than 2 signals, that is until now only mentioned as a possibility for the generalization of the method, we give an example of measuring the power of the wind. Measurement of wind power [7] is calculated by the formula:

$$P_S = 0{,}5 \times \rho \times A \times v^3 \tag{6}$$

wherein:

ρ – density of air (1,25kg / m3)

A – surface area through which air flows

v - air velocity

This formula is in practice reduced to a formula in which besides the speed of wind flow exists only circle diameter (D) that windmill wings plot:

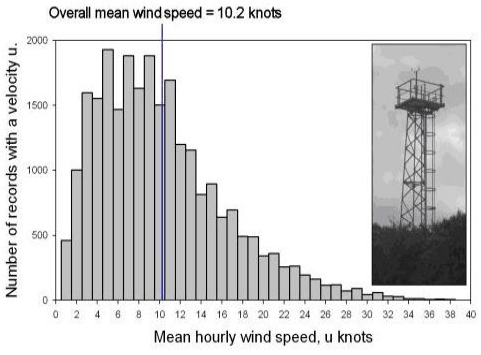
$$P_S = 0{,}291 \times D^2 \times v^3 \tag{7}$$

As you can see, wind power is a value proportional to the cube of its speed. Measurement of wind speed itself is done with an anemometer [8]. The measurements are performed on different locations in quite long period (up to one year). After that we determine the best place to set up a windmill and its optimal position for maximum efficiency at a chosen location.

Important parameters are the wind rose and the relative frequency of the wind histogram. Fig. 10 shows the direction of the wind in the period.

Fig. 11 shows the relative frequency histogram from where you can see how long each wind blows. These two parameters are important for determining the viability of a potential investment.

For this purpose, the device in Fig. 5 is being extend, as shown in Fig. 12. Device from Fig. 12 has 3 two-bit flash A/D converters from Fig. 4, with inputs $y_1 = f_1(t)$ and signal h_1, and $y_2 = f_2(t)$ and signal h_2, and $y_3 = f_3(t)$ and signal h_3, respectively. h_1, h_2 and h_3 are mutually uncorrelated random uniform dither signals. In this special case is $y_1 = y_2 = y_3 = v(t)$. Outputs Ψ_1, Ψ_2 and Ψ_3 are passed to a multiplier; the multiplier output is $\Psi = \Psi_1 \cdot \Psi_2 \cdot \Psi_3$, and it can assume values: $\Psi \in \left\{-(2g)^3, 0, +(2g)^3\right\}$.



Histogram of hourly wind speeds at Plymouth, Mountbatten.
(Years 2005 to 2007 - 25,203 valid data points)

Figure 11. Histogram of hourly wind speed [10].

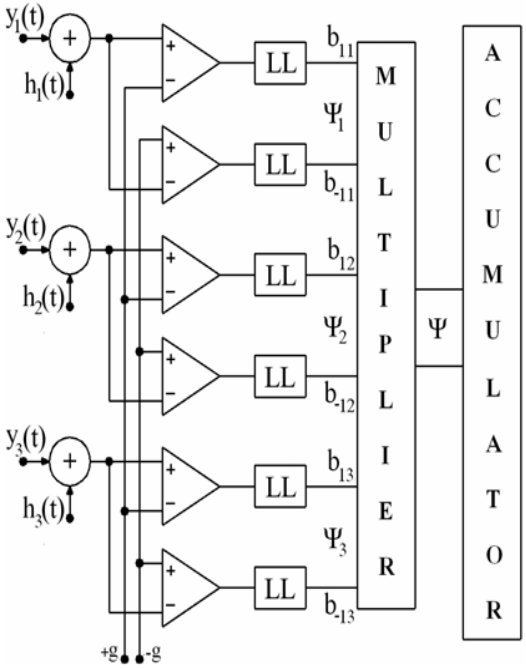


Figure 12. Structure of the device, based on two-bit flash A/D converters, for measurement of the finite integral product of three signals.

The corresponding simulation was performed to measure wind power with data usage from the website [11] for June, 11th 2011. The sampling frequency of 7.6 Hz was used, and data amount is 656104. Explanation of presented data is given on the website [12]. Maximum wind speeds at a given location are up to 25 m/s, while the particular sample maximum speed is 23.89 m/s. It is necessary to condition input signal (as well as the signals of current and voltage used in the previous sections) in order to obtain a voltage signal of an appropriate level (± 5 V).

To calculate wind energy of presented data it is necessary to multiply U_{max}^3 with the output value. To calculate wind power of presented data it is necessary to multiply U_{max}^3 with the mean output. Now we can start calculate individual measurement errors like we have done in both the measurement of the true RMS and the measurement of the power.

For a given set of data the measured energy by using that device is $148.5 \cdot D^2$ MJ, while the average power is $1720 \cdot D^2$ W. The corresponding errors are: absolute error = $3.83 \cdot D^2$, relative error = 0.223 % and errors relative to FS = 0.112 %

V. Conclusion

This paper presents the possibilities of digital stochastic measurement methods i.e. application of measurements of the definite integral product of two or more signals using two-bit flash A/D converters. Most important idea of this approach is to treat the time within the integration interval as an independent uniform random variable. In the developed method no assumptions are made regarding the waveform shape of measured signals. The two-bit flash A/D converter's design is rather simple, making it suitable for simple measurement of signal, and for parallel measurements without affecting conversion speed, measurement precision and accuracy.

Measurement cases of true RMS and power for both simple periodic (sinusoidal) and complex periodic wave form of signal we use for measuring of two signal product integral were examined trough simulation. In addition, the simulation was performed to measure three signal product, for which we show an example in measurement of wind power (wind power is proportional to the wind speed cube). Simulation samples are real life data with a sampling frequency of 7.6 Hz. Simulations confirm the theoretical expectations.

References

[1] J. von Neumann, "Probabilistic logic and the synthesis of reliable organisms from unreliable components," in *Automata Studies*, C. E. Shannon, Ed. Princeton, NJ: Princeton Univ. Press, 1956.

[2] M.F. Wagdy and W. Ng, "Validity of uniform quantization error model for sinusoidal signals without and with dither." *IEEE Transactions on Instrumentation and Measurement*, vol. 38, no. 3, June 1989. pp.718-722, doi: 10.1109/19.32180.

[3] V.V. Vujicic, I. Zupunski, Z. Mitrovic and M.A. Sokola, "Measurement in a point versus measurement over an interval." *Proc. of the IMEKO XIX World Congress*; Lisbon, Portugal. Sep. 2009. pp. 1128-1132 no.480.

[4] D. Pejic, M. Urekar, V. Vujicic and S. Avramov-Zamurovic, "Comparator offset error suppression in stochastic converters used in a watt-hour meter.", in *Proc. CPEM 2010, Proceedings*; Korea. June 2010.

[5] V. Pjevalica and V.V. Vujicic, "Further generalization of the low-frequency true-RMS Instrument." *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 3, March 2010. pp. 736-744.

[6] B. Ličina, P. Sovilj, Application of Digital Stochastic Measurement over an Interval in Time and Frequency Domain, pp. 297-302, International Conference ICIST 2014, Kopaonik, 09.03.–13.03.2014.

[7] M. Milosavljevic, M. Marjanovic, M. Obućina, "The estimation of wind energy potential for the production of electrical energy in Petrovac na Mlavi", in 56th *Proc. ETRAN Conference, Zlatibor, ML3.2-1-4, June 11-14, 2012.*

[8] M. Zlatanovic, "What does a cup anemometer measure?", in 56th *Proc. ETRAN Conference, Zlatibor, ML3.1-1-4, June 11-14*, 2014.

[9] http://nj.usgs.gov/grapher/tutorial/graphs/example_wind_rose.png.

[10] http://www.wind-power-program.com/wind_statistics.htm.

[11] https://code.google.com/p/google-rec-csp/downloads/detail?name=06-11-2011.txt.gz&can=2&q=&sort=-size.

[12] http://google.org/pdfs/google_heliostat_wind_data_collection.pdf.

Sessions TO-06 and TO-9

SIGNAL PROCESSING AND INFORMATION TECHNOLOGIES

Script Recognition by Statistical Analysis of the Image Texture

Invited Paper

Darko Brodić
Technical Faculty in Bor
University of Belgrade
Bor, Serbia
dbrodic@tf.bor.ac.rs

***Abstract*—The paper proposed an algorithm for the script identification using the statistical analysis of the texture obtained by script mapping. First, the algorithm models the script sign by the equivalent script type. The script type is determined by the position of the letter in the baseline area. Furthermore, the extraction of the features is performed. This step of the algorithm is based on the script type occurrence and co-occurrence pattern analysis. Then, the resultant features are compared. Their differences simplify the script feature classification. The algorithm is tested on the German and Slavic printed documents incorporating different scripts. The experiment gives the results that are promising.**

Keywords - Coding, Cultural heritage, Historical documents, Script recognition, Statistical analysis.

I. INTRODUCTION

Script recognition is a part of document image analysis [1]. Many techniques have been proposed for the script recognition. They are typically classified into global or local ones.

Global methods characterize the processing of the large image areas, which are subjected to the statistical or frequency-domain analysis [2]. However, the image area normalization is mandatory. [3]. On contrary, the local methods split up text into small pieces. They typically represents characters, words or lines. After that, the black pixel runs analysis is performed [4].

The proposed algorithm integrates the local and global approach. First, it extracts characters from the text. Then, it codes the characters according to their script type [5]. The coded text is obtained, which is an input to an occurrence (frequency analysis) and co-occurrence (statistical analysis) similarly as in global methods. As the results of aforementioned analysis, statistical measures of the gray-level co-occurrence matrix (GLCM) are extracted [6],[7]. To classify the results a linear discrimination function is proposed. It represents a key point in a decision-making process of the script discrimination.

The proposed approach incorporates the statistical analysis of the texture. Texture is suitable for extracting similarities and dissimilarities between images. However, the novelty of the proposed approach [8] is given by specific text modeling and 1-D texture analysis. During text modeling the number of variables is substantially reduced. Furthermore, the image is replaced with text. Hence, the image which represents a 2-D image is replaced with the text given by 1-D "image". All aforementioned contributes to the algorithm's speed.

The paper is organized as follows. Section 2 describes the proposed algorithm. Section 3 explains the experiment. Section 4 presents the results and gives the discussion. Section 5 makes conclusions.

II. THE METHODS

The proposed algorithm is a multi-stage method. It includes the following stages:

1. Coding,
2. Feature extraction,
3. Feature classification,
4. Script identification criteria.

Figure 1 illustrates the multi-stage method flow.

A. Coding

The first step of the algorithm represents a coding. It is established using into account the position of the letter in the text line. To explain it, the text line needs to be considered. It consists of three vertical zones [9]:

- Upper zone,
- Middle zone,
- Lower zone.

Figure 2 illustrates the vertical zones that belong to the text line.

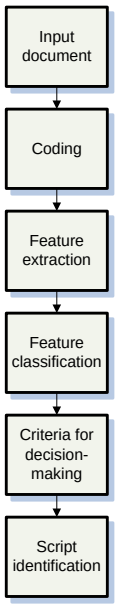


Figure 1. The multi-stage method flow

Figure 2. Vertical zones in text line

Using into account text line zones, one can distinguish four different script types [9]:

- Base letter (B),
- Ascender letter (A),
- Descendent letter (D),
- Full letter (F).

Base letters (B) like the letter **x**, occupy a middle zone only; ascender letters (A), like the letter **t**, spread over the middle and upper zones; while descendent letters (D), like the letter **p**, include the middle and lower zones. Few letters like the capital letter **Lj** (in Serbian or Croatian Latin alphabet) comprise all three zones. They are classified as a full letter (F). This way, all letters are coded according to their script type classification. To organize data, the following mapping is made [5,10]:

$$B \rightarrow 1,\ A \rightarrow 2,\ D \rightarrow 3,\ F \rightarrow 4 \qquad (1)$$

This way, the coded text is established (Appendix contains the alphabets and equivalent codes [10]).

B. Feature Extraction

Feature extraction is based on statistical analysis of the coded text. Figure 3 illustrates the text written in different Slavic scripts and their coded text, while Figure 4 shows German text written in Fraktur and Latin with their coded text.

Čuvaj uši svoje da slušaju samo svete i časne razgovore, a ne ružne i svjetovne, jer je napisano:

(a)

2 1 1 1 4 1 2 2 1 1 1 4 1 2 1 1 2 1 2 1 4 1 1 1 1 1 1 1
1 2 1 2 2 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 2 1
1 4 1 2 1 1 1 1 4 1 1 4 1 1 1 3 2 1 1 1 1

(b)

ⰘⰁⰃⰂⰓⰔ ⰁⰎⰅ ⰁⰂⰃⰁⰔⰓⰇ ⰊⰓⰓ ⰁⰂⰆⰁⰎⰓⰔⰁⰁ ⰁⰓⰓⰔ ⰁⰂⰃⰁⰂⰂⰇ ⰁⰓⰓⰔ ⰁⰂⰃⰁⰂⰂⰂⰇ Ⰵ ⰘⰓⰓⰁⰇ ⰆⰓⰓⰘⰊⰅⰁⰃⰁⰃⰆⰇ, ⰓⰓ ⰇⰇ ⰆⰁⰁⰘⰇ Ⰵ ⰁⰂⰃⰁⰔⰇⰂⰃⰁⰂⰂⰇ, ⰔⰔⰇⰆ ⰔⰔⰇ ⰇⰓⰘⰂⰅⰁⰓⰔⰇⰇ:

(c)

3 1 1 2 1 1 1 1 1 1 1 1 1 2 1 2 1 1 2 1 1 1 2 1 1 1 1
1 1 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 3 1 1 2 1 1

(d)

Чувај уши своје да слушају само свете и часне разговоре, а не ружне и свјетовне, јер је написано:

(e)

2 3 1 1 4 3 1 1 1 1 1 4 1 3 1 1 1 3 1 1 4 3 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 3 1 1 1 1 3 3 1 1 1 1 1
1 4 1 1 1 1 1 1 4 1 3 4 1 1 1 1 1 1 1 1 1

(f)

Figure 3. Same text given in different Slavic scripts: (a) Original text in Latin script, and (b) its coded counterpart; (c) Original text in Glagolitic script, and (d) its coded counterpart; (e) Original text in Cyrillic script, and (f) its coded counterpart.

Füllest wieder Busch und Tal Still mit Nebelglanz,
Lösest endlich auch einmal Meine Seele ganz;

(a)

2 2 2 2 1 1 1 1 1 1 2 1 1 2 1 1 1 2 1 1 2 2 1 2 2 2 1 2
2 1 1 2 2 1 2 1 2 3 2 1 1 1 2 2 1 1 1 2 1 1 2 2 1 1 2 1
1 1 2 1 1 1 1 1 2 2 1 1 1 1 2 1 1 2 1 3 1 1 1

(b)

Füllest wieder Busch und Tal Still mit Nebelglanz,
Lösest endlich auch einmal Meine Seele ganz;

(c)

4 2 2 2 1 4 2 1 1 1 2 1 1 2 1 4 1 4 1 1 2 2 1 2 2 2 1 2
2 1 2 2 2 1 2 1 2 3 2 1 1 3 2 2 4 1 4 2 1 1 2 2 1 1 4 1
1 1 4 1 1 1 1 1 2 2 1 1 1 1 2 1 1 2 1 3 1 1 3

(d)

Figure 4. Same text given in different German scripts: (a) Original text in Latin script, and (b) its coded counterpart; (c) Original text in Fraktur script, and (d) its coded counterpart

In the first step, the script type distribution of coded text is analyzed. As a result, four script features are extracted. After the script type distribution analysis, the coded text is subjected to the co-occurrence analysis [6,7]. This way, the texture features are calculated. They use the conditional joint probabilities of all pair wise combinations of grey levels in the window of interest (WOI). WOI is determined by the inter-pixel distance Δx and Δy shown in Figure 5.

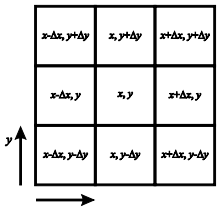


Figure 5. Window of interest (WOI).

The method starts from the top left corner and counts number of each reference pixel occurrences in respect to neighbour pixel relationship. At the end of this process, the element (i, j) gives the number of how many times the gray levels i and j appears as a sequence of two pixels located at Δx and Δy. This way, GLCM **P** for an image **I** with M rows and N columns is given as [6,7]:

$$P(i,j)=\sum_{x=1}^{M}\sum_{y=1}^{N}\begin{cases}1, \text{if } I(x,y)=i,\ I(x+\Delta x,y+\Delta y)=j\\ 0, \text{otherwise}\end{cases} \tag{2}$$

Furthermore, matrix **P** is normalized giving a matrix **C**:

$$C(i,j)=P(i,j)/\sum_{i,j}^{G}P(i,j) \tag{3}$$

In our case, the coded text is given as 1-D image, which leads to following: $\Delta x = \pm 1$, $\Delta y = 0$ [10]. Furthermore, the number of texture features can be extracted from the GLCM:

$$Energy=\sum_{i}^{G}\sum_{j}^{G}C(i,j)^2 \tag{4}$$

$$Entropy=\sum_{i}^{G}\sum_{j}^{G}C(i,j)\cdot\log C(i,j) \tag{5}$$

$$\text{Maximum}=\max\sum_{i}^{G}\sum_{j}^{G}C(i,j)\quad\forall i,j \tag{6}$$

$$Dissimilarity=\sum_{i}^{G}\sum_{j}^{G}C(i,j)\cdot|i-j| \tag{7}$$

$$Contrast=\sum_{i}^{G}\sum_{j}^{G}C(i,j)\cdot(i-j)^2 \tag{8}$$

$$Inverse\ Different\ Moment=\sum_{i}^{G}\sum_{j}^{G}C(i,j)/[1+(i-j)^2] \tag{9}$$

$$Homogeneity=\sum_{i}^{G}\sum_{j}^{G}C(i,j)/[1+(i-j)] \tag{10}$$

$$Correlation=\sum_{i}^{G}\sum_{j}^{G}(i-\mu_x)\cdot(j-\mu_y)\cdot C(i,j)/(\sigma_x\cdot\sigma_y) \tag{11}$$

C. Feature Classification

According to the aforementioned script type distributions and GLCM features, the text examples from Figures 3 and 4 are analyzed.

Table I shows the script type distributions, which are obtained from the same text written in different Slavic scripts [11].

TABLE I. SCRIPT TYPE DISTRIBUTIONS BETWEEN SLAVIC SCRIPTS

Script Type	**Different Slavic Scripts**		
	Cyrillic	***Glagolitic***	***Latin***
Base	0.7786	0.8015	0.6336
Ascender	0.0153	0.1679	0.2595
Descendent	0.1298	0.0305	0.0229
Full	0.0763	0.0000	0.0840

Similarly, the same analysis is carried out with German text. The results are shown in Table II [12].

TABLE II. SCRIPT TYPE DISTRIBUTIONS BETWEEN GERMAN SCRIPTS

Script Type	**Different German Scripts**	
	Fraktur	***Latin***
Base	0.5324	0.6115
Ascender	0.3237	0.3669
Descendent	0.0360	0.0216
Full	0.1079	0.0000

Table III shows typical GLCM features extracted from the Slavic documents. Differences in their features characterize each script [11].

At the end, Table IV shows typical GLCM features extracted from the German documents [12].

TABLE III. GLCM FEATURES BETWEEN SLAVIC SCRIPTS

GLCM Feature	Different Slavic Scripts		
	Cyrillic	***Glagolitic***	***Latin***
Energy	0.4140	0.4651	0.2159
Entropy	−1.3957	−1.1347	−1.8432
Maximum	0.6231	0.6538	0.3692
Dissimilarity	0.7615	0.6540	0.8846
Contrast	1.7923	1.3769	1.8077
Inverse different moment	0.7223	0.7454	0.6500
Homogeneity	0.7641	0.7859	0.6769
Correlation	0.0791	0.0742	−0.1291

TABLE IV. GLCM FEATURES BETWEEN GERMAN SCRIPTS

GLCM Feature	Different German Scripts	
	Fraktur	***Latin***
Energy	0.1596	0.2604
Entropy	-2.0675	-1.4985
Maximum	0.2754	0.3768
Dissimilarity	0.9855	1.0000
Contrast	2.0725	2.0725
Inverse different moment	0.6159	0.6072
Homogeneity	0.6504	0.6715
Correlation	-0.1403	-0.0149

D. Script Identification Criteria

All criteria obtained from occurrence and co-occurrence analysis are used as input criteria for decision-making. However, the comprehensive criteria will be established after applying the algorithm to the database of Slavic and German text documents. As a result, the statistical analysis will show the clear difference between scripts in document with the same content.

III. EXPERIMENTS

The algorithm is subjected to the experiment in order to investigate its efficiency and correctness. To perform the experiment, a custom-oriented database is created. First part of database includes 100 Slavic documents written in Glagolitic, Latina and Cyrillic script. Typical length of text is from approx. 500 to 4,000 characters. Texts are extracted from the book “Le château de virginité” (“The Castle of Virginity”) written in 1411 by George d’Esclavonie (Juraj Slovinac) [13]. Second part of database includes 100 German documents with the poems written by J. W. von Goethe written in Latin and Fraktur script. Typical length of text is from approx. 200 to 1,000 characters. The result of the experiment gives the percentage of the correct script recognition.

IV. RESULTS AND DISCUSSION

The results of experiment are given in the form of the script type distributions and the extended set of eight GLCM texture features. The script type distributions are used to extract four script features, which are used to characterize different scripts. To quantify the obtained results, we used the minimum and maximum values. Furthermore, the extended set of eight GLCM texture features is used as a basis to discriminate different scripts. To quantify the obtained results, we have used the minimum and maximum values. The texture features obtained from a statistical analysis of database texts written in Latin, Glagolitic and Cyrillic in the first place, and texts written in Latin and Fraktur in the second place.

It is very important to use only the measures with distinct difference in values for the different scripts. Establishing the ratio between these measures for different scripts gives their relation that can be utilized as a part of the identification criteria. These measures create the criteria for script discrimination.

The experiment with Slavic documents shows the results given in Tables V-VI [11,12].

TABLE V. SCRIPT TYPE DISTRIBUTIONS OF THE SLAVIC SCRIPTS

Script Type	Different Slavic Scripts					
	Cyrillic		***Glagolitic***		***Latin***	
	min	max	min	max	min	max
Base	0.48	0.62	0.68	0.79	0.68	0.85
Ascender	0.28	0.44	0.16	0.24	0.03	0.16
Descendent	0.03	0.07	0.03	0.17	0.07	0.16
Full	0.01	0.08	0.00	0.00	0.01	0.07

TABLE VI. GLCM FEATURES OF THE SLAVIC SCRIPTS

GLCM Feature	Different Slavic Scripts					
	Cyrillic		***Glagolitic***		***Latin***	
	min	max	min	max	min	max
Energy	0.167	0.214	0.309	0.432	0.325	0.507
Entropy	-2.026	-1.757	-1.499	-1.211	-1.647	-1.173
Maximum	0.237	0.355	0.497	0.626	0.539	0.701
Dissimilarity	0.757	0.986	0.684	0.981	0.595	0.871
Contrast	1.076	1.978	1.520	2.243	1.227	2.021
ID moment	0.604	0.679	0.636	0.742	0.679	0.780
Homogeneity	0.636	0.700	0.694	0.784	0.727	0.810
Correlation	-0.243	-0.159	-0.131	0.480	-0.118	0.075

Hence, the combination of the aforementioned results creates the final criteria for the script differentiation in the Slavic documents. It is given by the following pseudo code:

```
IF   [(Energy < 0.25) AND (Entropy < -1.7)
      AND (Maximum < 0.45) AND (Correlation < -0.15)
      AND (B < 0.65) AND (A > 0.26)]
            Writeln('Latin Text')
ELSEIF [((A < 0.16) AND (F > 0)]
            Writeln('Cyrillic Text')
ELSE
            Writeln(Glagolitic Text')
END
```

The experiment with German documents shows the results given in Tables VII-VIII [11,12].

TABLE VII. SCRIPT TYPE DISTRIBUTIONS OF THE GERMAN SCRIPTS

Script Type	**Different German Scripts**			
	Fraktur		***Latin***	
	min	max	min	max
Base	0.49	0.52	0.55	0.57
Ascender	0.37	0.39	0.41	0.43
Descendent	0.01	0.04	0.01	0.03
Full	0.08	0.10	0.00	0.01

TABLE VIII. GLCM FEATURES OF THE GERMAN SCRIPTS

GLCM Feature	**Different German Scripts**			
	Fraktur		***Latin***	
	min	max	min	max
Energy	0.152	0.175	0.236	0.248
Entropy	-2.108	-1.952	-1.572	-1.495
Maximum	0.219	0.241	0.2709	0.2871
Dissimilarity	0.921	1.037	0.596	1.199
Contrast	1.726	2.029	0.661	2.455
Inverse different moment	0.581	0.620	0.526	0.708
Homogeneity	0.611	0.641	0.604	0.711
Correlation	-0.186	-0.119	-0.181	-0.130

Accordingly, the combination of the aforementioned results creates the final criteria for the script differentiation in the German documents. It is given by the following pseudo code:

```
IF [(B < 0.53) AND (A < 0.4) AND (F > 0.05)
    AND (Energy < 0.2) AND (Entropy < -1.85)
    AND (Maximum < 0.25)]
            Writeln('Fraktur Text')
ELSE
            Writeln('Latin Text')
END
```

At the end, the speed testing of the proposed method shows that it is a computationally non-intensive. Its processing time is as low as 0.1 sec. per text that includes 2K characters.

V. CONCLUSIONS

The manuscript proposed the algorithm for the script characterization and identification. The algorithm includes the comprehensive statistical analysis of coded document, which is obtained by mapping the initial text document according to the script types of each character. Because the characteristics of both scripts are different, the statistical analysis shows significant diversity between them. Hence, the successful script identification can be conducted by creating joint criteria which is based on the obtained statistical features. The proposed technique is tested on the example of Slavic and German printed documents . The experiments gave encouraging results.

ACKNOWLEDGMENT

This work was supported by the Grant of the Ministry of Republic Serbia, as a part of the project TR33037.

REFERENCES

[1] D. Ghosh, T. Dube, A. P.Shivaprasad, “Script Recognition - A Review,” IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.32, no.12, 2010, pp. 2142–2161.

[2] G. D. Joshi, S. Garg, J. Sivaswamy, “A Generalised Framework for Script Identification,” International Journal of Document Analysis and Recognition, vol.10, no.2, 2007, pp.55–68.

[3] A. Busch, W. W. Boles, S. Sridharan, “Texture for Script Identification. IEEE Transaction on Pattern Analysis and Machine Intelligence,” vol.27, no.11, 2006, pp.1720–1732.

[4] U. Pal, B. B. Chaudhury, “Identification of Different Script Lines from Multi-Script Documents,” Image and Vision Computing, vol.20, no.13-14, 2002, pp.945–954.

[5] D. Brodić, Z. N. Milivojević, Č. A. Maluckov, “Recognition of the Script in Serbian Documents using Frequency Occurrence and Co-occurrence Analysis,” The Scientific World Journal, vol.2014, art.896328, 2013, pp.1–14. doi:10.1155/2013/896328

[6] R. Haralick, K. Shanmugam, I. Dinstein, “Textural Features for Image Classification,” IEEE Transactions on Systems, Man, and Cybernetics vol.3, no.6, 1973, pp.610–621.

[7] D. A. Clausi, “An Analysis of Co-occurrence Texture Statistics as a Function of Grey Level Quantization,” Canadian Journal of Remote Sensing, vol28, no.1, 2002, pp.45–62.

[8] S. Bourennane, J. Marot, C. Fossati, A. Bouridane, K. Spinnler, “Multidimensional Signal Processing and Applications,” The Scientific World Journal, vol.2014, art.365126, 2014, pp. 1–2.

[9] A. W. Zramdini, R. Ingold, “Optical font recognition using typographical features,” IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.20, no.8, 1998, pp.877–882.

[10] D. Brodić, Z. N. Milivojević, Č. A. Maluckov, “Script Characterization in the Old Slavic Documents,” Image and Signal Processing (ICISP 2014), LNCS 8509 (Eds. A. Elmoataz et al.), 2014, pp.230–238.

[11] D. Brodić, Z. N. Milivojević, Č. A. Maluckov, “An approach to the script discrimination in the Slavic documents,” Soft Computing, 2014, pp.1–11. doi: 10.1007/s00500-014-1435-1

[12] D. Brodić, Z. N. Milivojević, Č. A. Maluckov, “Identification of the Fraktur and Latin Script in German Printed Documents by Texture Analysis,” In review.

[13] http://www.croatianhistory.net/etf/juraj_slovinac_misli.html

APPENDIX

TABLE I. CODING OF SLAVIC ALPHABETS

Glagolitic	Coding	Latin	Coding	Cyrillic	Coding	Glagolitic	Coding	Latin	Coding	Cyrillic	Coding
ⰶⰻ	3	Lj	4	Љ	2	ⰾⰻ	2	lj	4	љ	1
ⰽⰻ	3	Nj	4	Њ	2	ⰴ	1	nj	4	њ	1
ⰵ	3	E	2	Е	2	ⰵ	1	e	1	е	1
ⰱ	3	R	2	Р	2	ⰱ	1	r	1	р	3
ⰸ	C	T	2	Т	2	ⰿⰿ	1	t	2	т	1
Ⰶⰾ	3	Z	2	З	2	Ⰶⰾ	2	z	1	з	1
ⰯⰯ	3	U	2	У	2	ⱏ	1	u	1	у	3
Ⱘ	3	I	2	И	2	Ⱘ	1	i	2	и	1
Ⱟⰻ	3	O	2	О	2	ⰹ	1	o	1	о	1
ⰻⰾ	3	P	2	П	2	ⰼⰾ	3	p	3	п	1
ⰻⰻ	3	Š	2	Ш	2	ⰻ	1	š	2	ш	1
ⰻⰽ	3	Đ	2	Ъ	2	ⰿⰽ	1	đ	2	ђ	4
ⰰⰽ	3	A	2	А	2	ⰰⰽ	2	a	1	а	1
ⰲ	3	S	2	С	2	ⰲ	1	s	1	с	1
ⰴⰼ	3	D	2	Д	2	ⰿⰼ	1	d	2	д	3
Ⱆⰾ	3	F	2	Ф	2	Ⱆⰾ	3	f	2	ф	4
ⰳⰾ	3	G	2	Г	2	ⰳⰾ	1	g	3	г	1
Ⱗⰾ	3	H	2	Х	2	Ⱗⰾ	1	h	2	х	1
ⰽⰽ	3	J	2	Ј	2	ⰽⰽ	1	j	4	ј	4
ⰷⰼ	3	K	2	К	2	ⰷⰼ	1	k	2	к	1
ⰶⰾ	3	L	2	Л	2	ⰾⰾ	2	l	2	л	1
ⰸⰸ	3	Č	2	Ч	2	ⰸⰸ	2	č	2	ч	1
ⰵⰽ	3	Ć	2	Ћ	2	ⰵⰽ	2	ć	2	ћ	2
ⰶⰶ	3	Ž	2	Ж	2	ⰾⰾ	2	ž	2	ж	1
ⰶⰻⰼ	4	Dž	2	Џ	4	ⰾⰻⰼ	2	dž	2	џ	3
ⰹⰯ	3	C	2	Ц	4	ⰹⰯ	1	c	1	ц	3
ⰹⰾ	3	V	2	В	2	ⰿⰿ	1	v	1	в	1
ⰱⰸ	3	B	2	Б	2	ⰱⰸ	1	b	2	б	2
ⰲ	3	N	2	Н	2	ⰲ	1	n	1	н	1
ⰿ	3	M	2	М	2	ⰿ	1	m	1	м	1
ⰰⰷ	3	Ja, (I)je	-	Ја, (И)је	-	ⰷ	1	ja, (i)je	-	ја, (и)је	-

TABLE II. CODING OF GERMAN DIACRITICS

Latin	Coding	Fraktur	Coding	Latin	Coding	Fraktur	Coding
Ä	2	𝔄̈	2	ä	2	𝔞̈	2
Ö	2	𝔒̈	2	ö	2	𝔬̈	2
Ü	2	𝔘̈	2	ü	2	𝔲̈	2
-		-		ß	2	𝔰𝔰	4

TABLE III. CODING OF GERMAN ALPHABETS

Latin	Coding	Fraktur	Coding	Latin	Coding	Fraktur	Coding
A	2	𝔄	2	a	1	𝔞	1
B	2	𝔅	2	b	2	𝔟	2
C	2	ℭ	2	c	1	𝔠	1
D	2	𝔇	2	d	2	𝔡	2
E	2	𝔈	2	e	1	𝔢	1
F	2	𝔉	4	f	2	𝔣	4
G	2	𝔊	2	g	3	𝔤	3
H	2	ℌ	4	h	2	𝔥	4
I	2	ℑ	2	i	2	𝔦	2
J	2	𝔍	4	j	4	𝔧	4
K	2	𝔎	2	k	2	𝔨	2
L	2	𝔏	2	l	2	𝔩	2
M	2	𝔐	2	m	1	𝔪	1
N	2	𝔑	2	n	1	𝔫	2
O	2	𝔒	2	o	1	𝔬	2
P	2	𝔓	4	p	3	𝔭	3
Q	2	𝔔	2	q	3	𝔮	3
R	2	ℜ	2	r	1	𝔯	1
S	2	𝔖	2	s	1	𝔰	1
T	2	𝔗	2	t	2	𝔱	1
U	2	𝔘	2	u	1	𝔲	1
V	2	𝔙	2	v	1	𝔳	1
W	2	𝔚	2	w	1	𝔴	1
X	2	𝔛	2	x	1	𝔵	3
Y	2	𝔜	4	y	3	𝔶	3
Z	2	ℨ	4	z	1	𝔷	3

Design of Two-Channel Analysis Part of Hybrid Filter Bank

Nikola Stojanović, Dragiša Milovanović
University of Niš
Faculty of Electronic Engineering
Niš, Serbia
{nikola.stojanović,dragisa.milovanovic}@elfak.ni.ac.rs

Negovan Stamenković
University of Priština
Faculty of Natural Science and Mathematics
K. Mitrovica, Serbia
negovan.stamenkovic@pr.ac.rs

***Abstract*—In the present paper a new design method for continuous-time power-symmetric active RC filters, which is suitable for two-channel hybrid filter bank realization, is proposed. Some theoretical properties of continuous-time power-symmetric filters in a more general perspective are studied. This includes the derivation of a new general analytical form, and a study of poles and zeros locations in s-plane. In the proposed design method the analytic solution of filter coefficients is solved in s-domain using only one nonlinear equation. Finally, the proposed approximation is compared to standard approximations. It was shown that attenuation and group delay characteristic of the proposed filter lie between Butterworth and elliptic characteristics.**

Keywords- Hybrid filter bank, power complementary filter pair, rational transfer function, allpass network, active RC filter.

I. Introduction

Many continuous-time signals have a low level nature, such as the output of sensors, then their processing often require multi-band decomposition for time-frequency analysis, manipulation, recognition of the signal, or storage. The hybrid filter bank (HFB) can be used for these applications and it is also suitable for high resolution conversion between analog and digital signals. Therefore, HFB is associated with analog to ADCs working at lower sample rate in comparison with the Nyquist sampling rate. Thus, the HFB is an unconventional class of the filter bank that employs both analog and digital filters [1]–[4].

The principle of a continuous-time linear hybrid two channel filter bank is shown in Fig. 1. The system consists of an continuous-time analysis two channel filter bank, uniform samplers, quantizers, and a discrete-time synthesis filter bank. The analysis filter bank consists of the low-pass filter $H_0(s)$ and high-pass filter $H_1(s)$. Both filters have the same passband edge and split the spectrum of the band limited input signal $x(t)$ by the factor of 2. The sampling and quantization takes place at the output of the analysis filters with the twice lower sampling frequency $1/(2T_s)$. The quantized signal goes into a linear discrete-time synthesis filter bank, which generate a single signal from two upsampled and interpolated signals. The up samplers are used to retain the desired Nyquist sampling rate $1/T_s$.

The continuous-time filters chosen to build the analysis filter bank play an important role in the performance of the hybrid filter bank. It is known that continuous-time filters of odd degree can be realized as the sum of two stable all-pass filters with real coefficients having no common poles [5]. As all-pass sums, such filter bank can be realized with low complexity structures that are robust to finite precision of components.

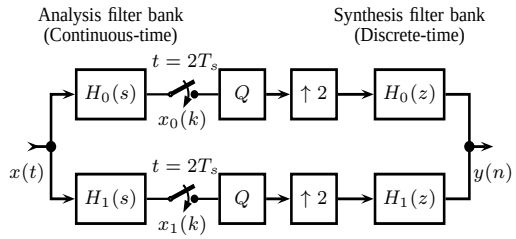


Fig. 1. Two channel hybrid continuous/discrte filter bank.

For the definition purposes, consider a continual prototype lowpass-highpass filter pair, denoted by $[H_0(s),\ H_1(s)]$, where $H_0(s)$ is transfer function of lowpass part of filter pair and $H_1(s) = H_0(s^{-1})$ is transfer function of highpass part filter pair. Normalized passband edge for both filters is equal to one.

A filter pair $[H_0(s),\ H_1(s)]$ is a *power-complementary* filter pair [6] if the sum of the squares of their magnitude responses satisfies

$$H_0(s)H_0(-s) + H_1(s)H_1(-s) = 1 \tag{1}$$

or at real frequency $s = j\,\omega$

$$|H_0(j\omega)|^2 + |H_1(j\omega)|^2 = 1 \tag{2}$$

For this pair, the angular frequency $\omega_c = 1$, where $|H_0(j)|^2 = |H_1(j)|^2 = 0.5$, is the *crossover angular frequency*. At this angular frequency, the gain responses of both filters are approximately 3 dB below their maximum values. Note, crossover angular frequency is 3 dB passband edge for both lowpass and highpass part of filter pair. Thus, $H_1(s) = H_0(s^{-1})$.

A filter pair $[H_0(s),\ H_1(s)]$ is an *all-pass-complementary* filter pair [7] if the sum and difference of $H_0(s)$ and $H_1(s)$ satisfies

$$\begin{aligned} H_0(s) + H_1(s) &= A_1(s) \\ H_0(s) - H_1(s) &= A_2(s) \end{aligned} \tag{3}$$

where $A_1(s)$ and $A_2(s)$ are all-pass transfer functions.

Transfer function sets which are simultaneously all-pass complementary and power complementary are termed *double complementary*. All double complementary filter pair can

be expressed as the sum of stable allpass filters such as the Butterworth and the elliptic filters. The other classical approximation cannot form a double complementary pair.

New efficient approximation of the doubly complementary filter pair is proposed in this paper. The realization, based on the continuous-time all-pass filters, which validate this approach is also presented.

II. Approximation

Necessary and sufficient conditions for the transfer function to be suitable for the realization of the continuous-time two channel filter banks are given in this section. In general, the squared magnitude characteristic of the lowpass prototype in the s-plane is expressed in the form

$$H_0(s)H_0(-s) = \frac{1}{1+K(s)K(-s)} \tag{4}$$

where filter characteristic function $K(s)$ is rational, and the polynomial in denominator contains only even or odd power of s, but, the polynomial in nominator contains only even power of s. For power symmetric filter design, at normalized passband edge frequency $s = \pm j$ the characteristic function is equal to one, then the insertion loss of filter at this frequency is 3.0103 dB. In fact, Butterworth, Chebyshev1, Chebyshev2, and elliptic filters are introduced in this form, and the filter properties are governed in a way where $K(s)$ is chosen [8].

Lemma 1: The rational transfer functions $H_0(s)$ and $H_1(s)$ in Eq. (1) satisfied power symmetric in z [9] and in s-domain if

$$K(s)K(-s)K(s^{-1})K(-s^{-1}) = 1 \tag{5}$$

Proof: For Eq. (4) we have

$$\begin{aligned} &H_0(s)H_0(-s) + H_0(s^{-1})H_0(-s^{-1}) \\ &\quad = \frac{1}{1+K(s)K(-s)} + \frac{1}{1+K(s^{-1})K(-s^{-1})} \\ &\quad = \frac{1+R(s)}{R(s)+K(s)K(-s)K(s^{-1})K(-s^{-1})} \end{aligned} \tag{6}$$

where $R(s) = 1 + K(s)K(-s) + K(s^{-1})K(-s^{-1})$. Clearly, this is equal to one if and only if

$$K(s)K(-s)K(s^{-1})K(-s^{-1}) = 1$$

This lemma is proved.

Based on the preceding result it is possible to develop a general analytic form for $K(s)$ which is suitable for continuous-time power symmetric filter design.

Lemma 2: A rational filter transfer function (4) satisfied power symmetric [10] in s-domain (5) if and only if characteristic function has the form

$$K(s) = s^k \prod_{m=1}^{M} \left(\frac{s^2+\omega_m^2}{\omega_m^2 s^2 + 1} \right)^{l_m} \tag{7}$$

with $\omega_m < 1$ and arbitrary integer l_m, for $m = 1, 2, \ldots, M$. Filter order is $N = k + 2\sum_{m=1}^{M} l_m$.

This condition can be expressed equivalently by

$$K(-s)K(s^{-1}) = (-1)^k \tag{8}$$

for all s.

Proof: This comes from the following facts:

1) $K(-s) = (-1)^k K(s)$
2) $K(s^{-1}) = \dfrac{1}{K(s)}$

This lemma is proved.

Note, for $k > 1$ and $l_1 = l_2 = \cdots = l_M = 0$ we have Butterworth filter which is power-symmetric. For $k = 0$ or 1 and $l_1 = l_2 = \cdots = l_M = 1$ we have Elliptic filters which are also power-symmetric. Chebyshev filters cannot be power-symmetric because they have ripples only in the passband or stopband.

Lemma 3: Let $H_0(s)$ be a rational with real coefficients power symmetric filter function, then all poles of it are restricted to be on the unit circle.

Proof: Since $K(s)$ has the form (7), its poles are restricted to be on the imaginary axis, then power symmetric $H_0(s)$ implies (8) and equation (4) can be rewritten as

$$H_0(s)H_0(-s) = \frac{1}{1+(-1)^k \dfrac{K(s)}{K(s^{-1})}} \tag{9}$$

At pole frequencies of $H(s)$ the denominator of the expression (9) is zero, that is

$$\frac{K(s)}{K(s^{-1})} = (-1)^{k+1} \tag{10}$$

In view of the real-coefficient assumption at $s = e^{-j\theta}$ we have $K(e^{-j\theta}) = K^*(e^{j\theta})$. On the unit circle of the s-plane we therefore have

$$\left| \frac{K(e^{j\theta})}{K(e^{-j\theta})} \right| = 1 \tag{11}$$

So, the quantity $K(s)/K(s^{-1})$ has unit-magnitude on the unit circle, then all poles of $H(s)H(-s)$ are on the unit circle.

Lemma 4: A filter pair $[H_0(s), H_1(s)]$ is an all-pass-complementary filter pair of transfer functions, i.e., a par satisfied

$$H_0(s) + H_1(s) = A_1(s) \tag{12}$$

where $A_1(s)$ is an stable all-pass transfer function. Then, the following equation is also automatically satisfied

$$H_0(s) - H_1(s) = A_2(s) \tag{13}$$

Proof: Since $A_1(s)$ and $A_2(s)$ are the allpass transfer functions, then $A_1(s)A_1(-s) = 1$ and $A_2(s)A_2(-s) = 1$. Further, the squared magnitude characteristic of left side of equation (12) is

$$\begin{aligned} G(s) =& [H_0(s) + H_1(s)][H_0(-s) + H_1(-s)] \\ =& 1 + H_0(s)H_1(-s) + H_0(-s)H_1(s) \end{aligned} \tag{14}$$

because filter pair $[H_0(s),\ H_1(s)]$ is power complementary.

Let N be odd. Transfer function $H_0(s)$ and $H_1(s)$ have the form

$$H_0(s) = \frac{\prod_{m=1}^{M}(\omega_m^2 s^2+1)^{l_m}}{(s+1)(s^{N-1}+\alpha_1 s^{N-2}+\cdots+\alpha_1 s+1)}$$
$$H_1(s) = \frac{s^{N-2\nu}\prod_{m=1}^{M}(s^2+\omega_m^2)^{l_m}}{(s+1)(s^{N-1}+\alpha_1 s^{N-2}+\cdots+\alpha_1 s+1)} \quad (15)$$

where $\nu = \sum_{m=1}^{M} l_m$ and $N = k + 2\nu$. Since $H_0(s)$ is a ratio of even and odd polynomials then $H_0(-s) = -H_0(s)$. On the other hand, $H_1(s)$ is a ratio of odd polynomials then $H_1(-s) = H_1(s)$. When this is substituted into equation (14), we obtain $G(s) = 1$, i.e., squared magnitude characteristic of $H_0(s) + H_1(s)$ is equal to one. Thus, $H_0(s) + H_1(s)$ is an allpass function.

III. The two Channel Filter Bank

Two-channel power complementary filter bank [11] is shown in Fig. 2 is considered in this section, where $A_1(s)$ and $A_2(s)$ are two continuous time stable all-pass filters with real coefficients having no common poles.

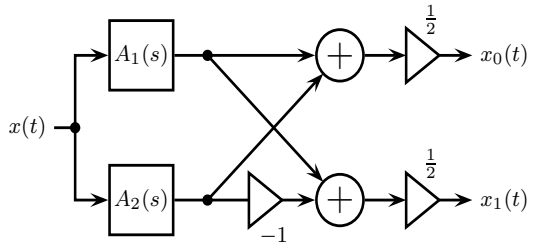


Fig. 2. The system of two-channel power complementary filter bank.

It is interesting to note that only continuous-time filters of odd degree can be realized as all-pass sums.

$$H_0(s) = \frac{X_0(s)}{X(s)} = \frac{1}{2}[A_1(s) + A_2(s)] \quad (16)$$

and

$$H_1(s) = \frac{X_1(s)}{X(s)} = \frac{1}{2}[A_1(s) - A_2(s)] \quad (17)$$

The transfer functions $H_0(s)$ and $H_1(s)$ can be implemented simply by implementing all-pass networks $A_1(s)$ and $A_2(s)$.

A. Approximation

The squared magnitude of the transfer function of the proposed n-th degree with single pair of zeros ($M = 1$ and $l_1 = 1$) at $\pm j\omega_1$ is:

$$|H(j\omega)|^2 = \frac{1}{1+\omega^{2k}\left(\frac{\omega_1^2-\omega^2}{1-\omega^2\omega_1^2}\right)^2} \quad (18)$$

where $\omega_1 < 1$ determined minimum stop-band attenuation.

Performing analytic continuation $\omega = -js$, equation (18) gets form

$$H(s)H(-s) = \frac{(1+s^2\omega_1^2)^2}{(1+\omega_1^2 s^2)^2 + (-1)^k s^{2k}(s^2+\omega_1^2)^2}, \quad (19)$$

or in simpler form, $H(s)$ can be written as

$$H(s)H(-s) = \frac{(\omega_1^2 s^2+1)^2}{s^{2N} + d_1^{2(N-1)} + \cdots + d_1 s + 1} \quad (20)$$

where

$$d_i = \begin{cases} (-1)^k\binom{2}{i}(\omega_1^2)^i, & i = 0,\ldots,k-1, \\ \binom{2}{i}(\omega_1^2)^i + \binom{2}{i-k}(\omega_1^2)^{2-i+k}, & i = k,\ldots,2, \\ \binom{2}{i-k}(\omega_1^2)^{2-i+k}, & i = 3,\ldots,N \end{cases}$$

If $k \geq 4$, then $d_i = 0$, for $i = 4,\ldots,k$. The poles of $H(s)$ are the poles of $H(-s)$, reflected about origin. Since the desired filter function must have all poles in the left half of the s-plane, we must associate the left half plane poles of $H(s)H(-s)$ with $H(s)$. Unknown parameter ω_1 to be determined so that the minimum attenuation in the stop-band has specified value R_s. This can be done by solving a single nonlinear equation in one unknown.

B. An example

For example, by setting $k = 3$, $M = 1$, $l_1 = 1$ and $\omega_1 = 0.551017$ the order of the filter is $N = 5$, and minimum stopband attenuation is $R_s = 40$ dB. The factored form of transfer functions of the two channel filter bank in the s-plane, which is designed by using proposed method, is

$$H_0(s) = \frac{0.303617s^2+1}{(s+1)(s^2+0.472822s+1)(s^2+1.472822s+1)}$$

and the analogue highpass prototype is

$$H_1(s) = \frac{s^3(s^2+0.303617)}{(s+1)(s^2+0.472822s+1)(s^2+1.472822s+1)}$$

These two prototypes are all-pass complementary

$$A_1(s) = \frac{s^2-1.472821\,s+1}{s^2+1.472821\,s+1}$$
$$A_2(s) = \frac{(-s+1)(s^2-0.472822s+1)}{(s+1)(s^2+0.472822s+1)} \quad (21)$$

and power complementary, i.e., double complementary. In terms of the pole frequency ω_p and the pole quality factor q_p, we recognize in biquad transfer function $\omega_p = 1$ (as in the case of the Butterworth filter) and first degree coefficient is $1/q_p$.

Pole-zero plots of these two allpass functions are shown in Fig. 3. In $A_2(s)$ are included the outermost pole pair, the third outermost pole pair, and so on (see the pole pair tagged with $\times$ on the Fig. 3). The remaining pole pairs belong to $A_1(s)$.

Fig. 4 gives a comparison of new filter frequency responses ($R_s = 40$ dB and $k = 5$) with Butterworth filter frequency responses. Both filters are 7th degree.

A sample of some normalized proposed transfer functions for 3 dB maximum passband attenuation and 40 dB minimum stopband attenuation, in factored form, can be found in Table I. In this table the frequencies are normalized to passband edge frequency $\omega_c = 1$. The numerator is normalized to that the dc gain of thee system is equal to unity. The passband ripple can be calculated by using (2). Since stopband ripple is 40 dB (0.01 times) then passband ripple is 4.3432×10^{-4} dB.

TABLE I
DOUBLE COMPLEMENTARY APPROXIMATION FUNCTION FOR $R_p = 40$ dB

n	Denominator of $H(s)$	Numerator of $H(s)$
3	$(s+1)(s^2+0.9124264s+1)$	$0.0875736s^2+1$
5	$(s+1)(s^2+1.4728206s+1)(s^2+0.4728206s+1)$	$0.3036200s^2+1$
7	$(s+1)(s^2+1.7136030s+1)(s^2+1.0198032s+1)(s^2+0.3062001s+1)$	$0.4649038s^2+1$
9	$(s+1)(s^2+1.8305027s+1)(s^2+1.3690518s+1)(s^2+0.7627434s+1)(s^2+0.2241943s+1)$	$0.5714580s^2+1$

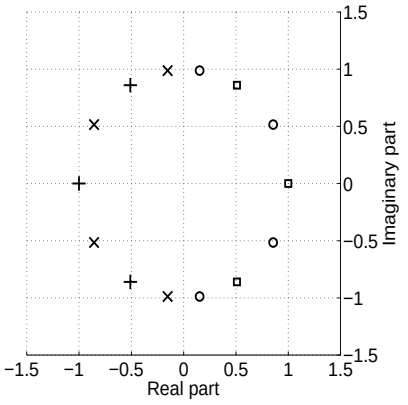


Fig. 3. Pole-zero plot of allpass functions in s-plane of the 7th-order prototype for transfer function with $R_s = 40$ dB, $M = 1$ and $l_1 = 1$. Poles and zeros of $A_1(s)$ are tagged with $\times$ and $\bigcirc$ respectively; but poles and zeros of $A_2(s)$ are tagged with $+$ and $\square$ respectively.

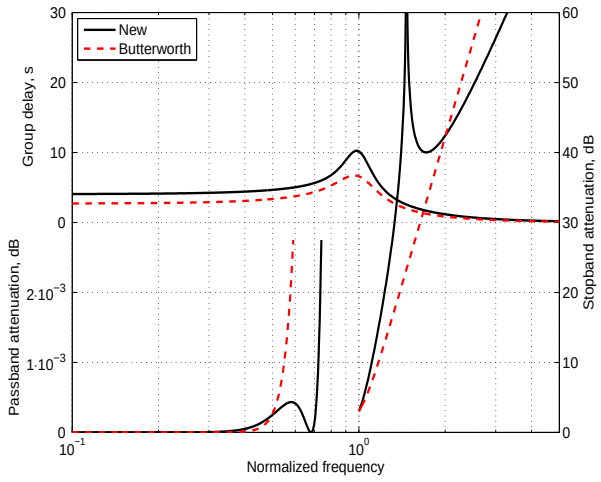


Fig. 4. Comparison of 7th degree new filter frequency responses ($R_s = 40$ dB and $k = 5$) with 7th degree Butterworth filter frequency responses.

IV. IMPLEMENTATION

In this section we discuss the design of continuous-time part of two channel hybrid filter bank based on all-pass active RC structure. All-pass transfer functions are non-minimum phase transfer function i.e., they have zeros in the right half of the complex-frequency plane.

Basically, there are two implementation manners of realizing such continuous-time allpass filter circuits. One alternative is the passive implementation, consisting of only passive components like capacitors and inductors. A number of passive circuit topologies exist, which can be used for this purpose, for instance the Lattice or T-section filters. The other alternative is the active implementation, consisting of active devices like operational amplifiers as well. Through the application of active components, it is possible to omit the bulky and costly inductor components, as well as providing more freedom in the shaping of the filter characteristic.

The active RC realization by cascading first and second order section (biquad) is proposed in this paper. The biquad can be realized with single or more operational amplifiers. The single amplifier lowpass and highpass filter will be discussed.

The transfer function of an even degree is not suitable for complementary decomposition because its allpass transfer functions have complex coefficients.

The single amplifier all-pass network first degree and the Delyiannis second degree all-pass circuit are given on the figure 5(a) and 5(b), respectively [12, chap. 4]. It is assumed that the ideal operational amplifiers are used.

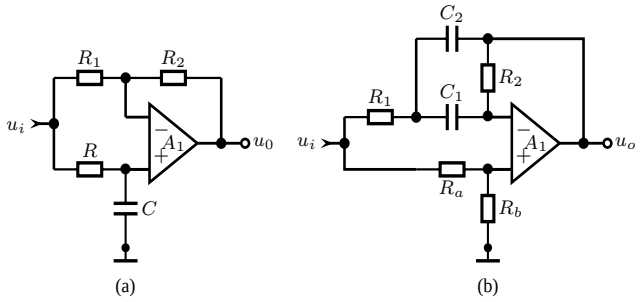


Fig. 5. (a) The single amplifier all-pass network first degree and (b) the all-pass network second degree.

The circuit on the figure 5(a) is referred to as the grounded capacitor allpass network. For $R_1 = R_2$ transfer function of first degree allpass network is

$$H(s) = -\frac{s - \dfrac{1}{RC}}{s + \dfrac{1}{RC}}$$

Transfer function of allpass network second degree, assuming ideal operational amplifier, $C_1 = C_2 = C$ is as follows

$$H(s) = h_o \frac{s^2 - \left[\dfrac{R_a}{R_b}\dfrac{1}{R_1 C} - \dfrac{2}{R_2 C}\right]s + \dfrac{1}{R_1 R_2 C^2}}{s^3 + \dfrac{2}{R_2 C}s + \dfrac{1}{R_1 R_2 C^2}} \qquad (22)$$

where $h_o = R_b/(R_a + R_b)$. Equating coefficients of equal powers of s in Eqs. (21) and (22) we can obtain the following component values

$$R_1 = \frac{1}{2Cq_p}, \quad R_2 = \frac{2q_p}{C}, \quad \frac{R_a}{R_b} = 4\frac{R_1}{R_2} \qquad (23)$$

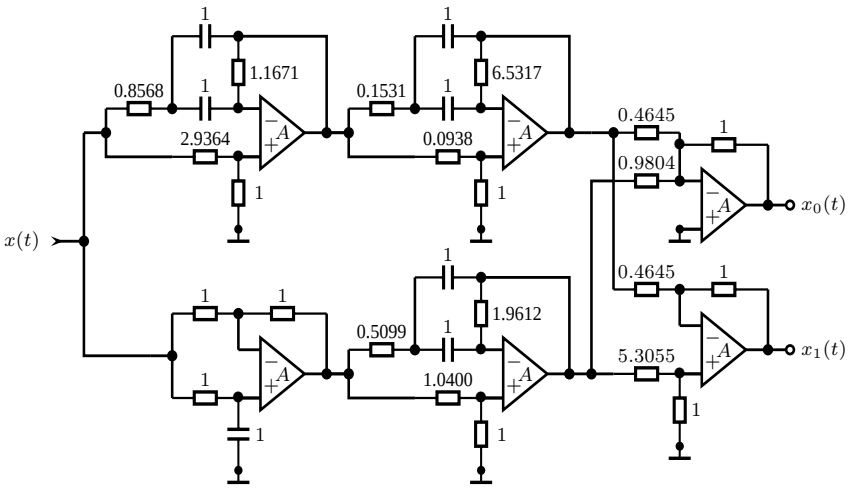


Fig. 6. Implementation of seven degree continuous-time part of two channel hybrid filter bank.

From Fig. 5, we have the implementation of the two channel continuous-time filter bank with component values shown in Fig. 6. Pole-zero plot is on the Fig. 3 given. The first all-pass filter $A_1(s)$ is realized by two second order sections shown in figure 5(b) and placed in cascade. The second all-pass filter $A_2(s)$ is realized by one first order section shown in figure 5(a) and one second order section placed in cascade.

The standard inverting summing amplifier is used for combining two signals. The differential amplifier circuit is used as subtractor. In these circuits, input signals are scaled to the desired values by selecting appropriate values for the resistors.

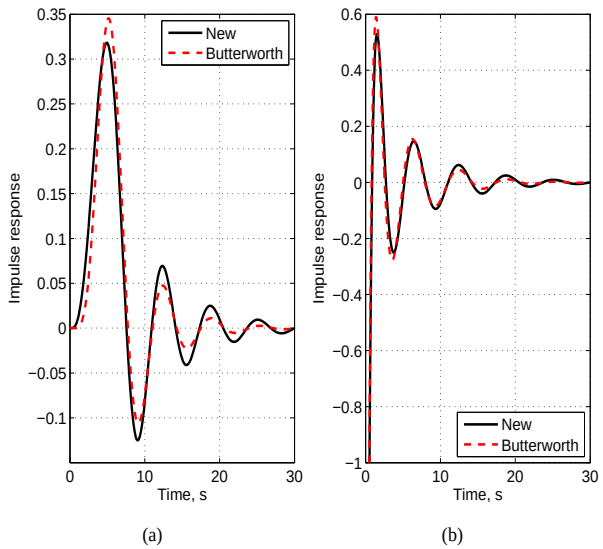


Fig. 7. Impulse responses for $N = 7$. (a) Lowpass filter, (b) highpass filter

The lowpass and highpass impulse responses for a seventh-order conventional Butterworth and proposed filter can be seen in Fig. 7. As can be seen both impulse response are very similar. In connection with Figure 4 new filter has better frequency responses in comparison with the Butterworth filter.

V. Conclusion

A new class of continuous-time filter structures has been presented, which can be used for efficient implementation of the hybrid filter bank. The conditions required to be satisfied by the transfer functions, so as to be implemented as a parallel connection of two all-pass filters, are listed. If filter degree is odd, allpas functions have real coefficients, but for even filter degree the allpass function involves in the implementation the complex coefficients. The efficiency of the proposed design has been demonstrated by means of an example.

Acknowledgment

This work is supported by Serbian Ministry of Science and Technologies, Project No. 32009TR.

References

[1] S. R. Velazquez, T. Q. Nguyen, and S. R. Broadstone, "Design of hybrid filter banks for analog/digital conversion," *IEEE Trans. on Signal Processing*, vol. 46, no. 4, pp. 956–967, Apr. 1998.

[2] B. Rumberg and D. Graham, "A low-power and high-precision programmable analog filter bank," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 59, no. 4, pp. 234–238, Apr. 2012.

[3] A. Kammoun, C. Lelandais-Perrault, and M. Debbah, "SNR efficient approach for the design of hybrid filter bank A/D converters," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, May 2013, pp. 2716–2720.

[4] N. Stamenković and V. Stojanović, "The design of two channel IIR QMF bank directly from analog prototype," *Digital Signal Processing*, vol. 98, no. 7, pp. 961–972, July 2011.

[5] I. W. Selesnick, "Lowpass filters realizable as allpass sums: design via a new at delay filter," *IEEE Trans. on Circuits and Systems Part II*, vol. 46, no. 1, pp. 40–50, Jan. 1999.

[6] S. Koshita, M. Abe, and M. Kawamata, "State-space analysis of power complementary analog filters," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, May 2007, pp. 2814–2817.

[7] J.-H. Lee and Y.-H. Yang, "Design of two-channel linear-phase QMF banks based on real IIR all-pass filters," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 150, no. 5, pp. 331–8–, Oct. 2003.

[8] M. D. Lutovac, D. V. Tosic, and B. L. Evans, *Filter Design for Signal Processing using Matlab and Matematica*. New Jersey, USA: Prentice Hall, 2001.

[9] P. Vaidyanathan, "Some properties of iir power-symmetric filters," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, Apr. 2007, pp. III–1449–III–1452.

[10] D. Živaljević and V. Stojanović, "Transitional Butter-Elliptic filter suitable for complementary decomposition," *Inverse Problems in Science and Engineering*, vol. 20, no. 1, pp. 117–125, Jan. 2012.

[11] H. Lollmann and P. Vary, "Design of iir qmf banks with near-perfect reconstruction and low complexity," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, Mar. 2008, pp. 3521–3524.

[12] T. Deliyannis, Y. Sun, and J. K. Fidler, *Continuous-Time Active Filter Design*. Boca Raton: CRC Press LLC, 1999.

Constrained Lexicon Speaker Dependent Recognition of Whispered Speech

Jovan Galić[1,2]
[1]Telecommunications Department
School of Electrical Engineering
Belgrade, Serbia
jgalic@etfbl.net

[2]Department of Electronics and Telecommunications
Faculty of Electrical Engineering, University of Banja Luka
Bosnia and Herzegovina

Slobodan T. Jovičić[1,3], Đorđe Grozdić[1,3]
[3]Life Activities Advancement Center, Laboratory for
Psychoacoustics and Speech Perception
Belgrade, Serbia
jovicic@etf.rs, djordjegrozdic@gmail.com

Branko Marković
Čačak Technical Colledge
Computing and Information Technology Department
Čačak, Serbia
brankomarko@yahoo.com

***Abstract*— In this paper we present results on automatic speech recognition of isolated words with part of Whi-Spe database with female speakers, in speaker dependent fashion and constrained lexicon (50 words). Word recognition rate is calculated for four train/test scenarios, with modeling of context independent monophones, context dependent triphones and whole words. As feature vectors, we used Perceptual Linear Prediction Coefficients and Mel Frequency Cepstral Coefficients. The Hidden Markov Model Toolkit was used to implement isolated word recognizer. Further improvement is achieved with reduction in number of monophone units used for modeling. Due to very high deviation in performance among different speakers, influence of Signal to Noise Ratio of tested recordings on performance of recognizer is examined in particular.**

Keywords- HTK, speech recognition, Whi-Spe database

I. INTRODUCTION

Whispering is a specific mode of speech often used in everyday life, esspecially by cellular phones. People whisper for a number of reasons, for example, in environments where normally phonated speech is not appropiate or concealment of some confidental information from the others ears. Beside conscious production of whisper, whispering may happen due to health problems which appear after rhinitis and laryngitis [1], [2]. The whisper has different characteristics compared to normally phonated speech. Due to the absence of the glottal vibrations, whispering lacks the fundamental frequency of the voice and much prosodic information. In addition, whispered speech has a significantly lower energy as compared to the normal speech [2], and the slope of the spectrum being much flatter than in the normal speech [3]. The duration of whispered speech is slightly longer [4], and the formant frequencies for whispered vowels is substantially higher than for the normal voice [4]. The amount of shift is higher for vowels with low formant frequencies [5]. Figure 1 shows waveform and Fig. 2 shows spectrogram of sentence "govor šapata" ("whispered speech" in English), uttered in normally phonated speech followed by whispered speech. Pictures are supported with phonetic transcription for both the normal (capital letters) and whispered speech (small letters). Because of the lack of sonority, difference in amplitude intensities could be observed, especially for vowels. Also, spectrogram shows that the harmonic structure of vowels in whispered speech is completely lost, which is separately presented for vowel /o/ in Fig. 3. However, spectral characteristics of unvoiced consonants (for example fricative /š/) are not significantly changed. Similar shape of spectrum of phoneme /r/ in Serbian is observed.

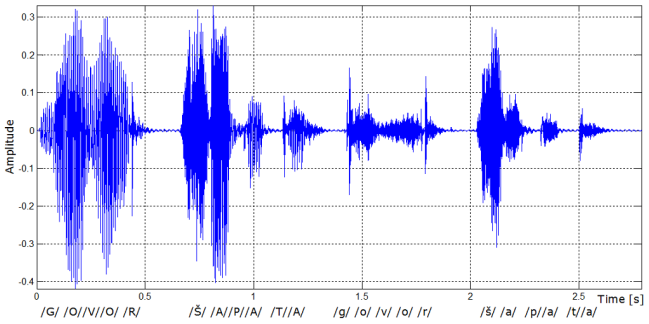


Figure 1. Waveform of sentence "Govor šapata" in normal phonation (capital letters) and whispered phonation (small letters)

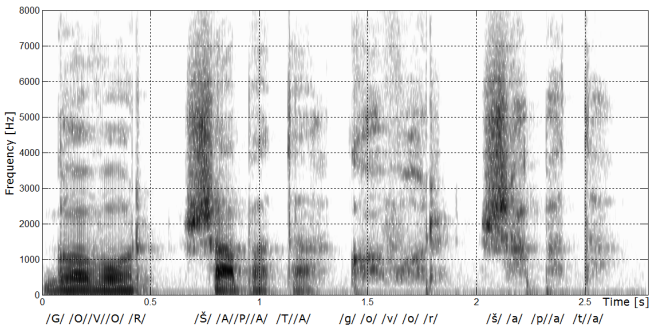


Figure 2. Spectrogram of sentence "Govor šapata" in normal phonation (capital letters) and whispered phonation (small letters)

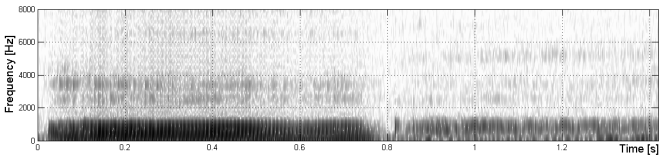


Figure 3. Spectrogram of vowel /o/ in sustained phonation (normal and whispered).

For all mentioned differences, automatic recognition of whispered speech is much more difficult than of normal speech. From [6], we know that speech can be generally classified into five categories based on vocal effort differences: whispered, soft, neutral, loud and shouted speech. The speaker identification performance significantly degrade with a change in vocal effort ranging from whisper through shouted, where whispered speech have the most serious loss in performance [7].

Nevertheless, despite of increased efforts in perception, this type of speech is perfectly understandable [8]. There are different approaches, techniques and methods of speech recognition. These techniques are usually based on algorithms of the HMM (Hidden Markov Model), the DTW (Dynamic Time Warping), the ANN (Artificial Neural Network) and their hybrid solutions [9]. This paper presents results on investigation of recognition of isolated words from part of Whi-Spe database [10] with female speakers, using a software toolkit HTK (Hidden Markov Model Toolkit). The HTK is a widely used software for ASR (Automatic Speech Recognition), that was originally developed at the Machine Intelligence Laboratory of the Cambridge University Engineering Department [11].

The reminder of this paper is organized as follows. In Section 2 the brief description of database Whi-Spe is given. In Section 3 we give the system overview used in experiments. The experimental results, as well as its discussion, are given in Section 4, while concluding remarks and further directions are stated in Section 5.

II. WHI-SPE DATABASE

The Whi-Spe database contains two parts: the first one contains speech patterns of a whispered speech, while the second one contains speech patterns of the normal speech. All patterns were collected from the five female and five male speakers. During the session of recording, each speaker read 50 isolated words of Serbian. The words were divided in three sub-corpora: basic colors (6 words), numbers (14 words) and phonetically balanced words (30 words). Balanced words were taken from the Serbian emotional speech database GEES [12], which satisfies the basic linguistic criteria of Serbian language. Sessions were repeated ten times, with a pause of a few days between recordings. Finally, the database collection grew to 10.000 utterances, half in the whispered speech and half in the normal speech. The speakers of ages between twenty and thirty were Serbian native volunteers from Čačak Technical College.

The speech was digitized by using the sampling frequency of 22.050 Hz, with 16 bits per sample, and stored in the form of Windows PCM (Pulse-Code Modulation) wave files. In this experiment, all samples from female speakers were used. Specific details about database concerning content, recording process and quality control could be find in [10].

III. SYSTEM OVERVIEW

In this work, all experiments were conducted on the latest version of HTK, 3.4.1 [13]. The toolkit was ported to Windows 7, and all experiments were done under this operating system. As a feature vectors, we used Mel Frequency Cepstral Coefficients (MFCC) and Perceptual Linear Prediction (PLP) coefficients, as widely used features in speech recognition [14], [15]. For obtaining a feature vector, Hamming window with preemphasis coefficient of 0.97 was used. The window size was set to 24 ms, and frame shift to 8 ms. Also, cepstral coefficient C_0, delta and acceleration coefficients were appended and cepstral mean subtraction was performed. These auxiliary features and modification techniques significantly improves recognition rate [16]. Number of filterbank channels was set to 26, and number of output cepstral coefficients per frame was set to 12. For better performance in these experiments, in filterbank analysis power was used instead of magnitude and normalization of energy was not included. The other parameters were set to default values.

The model topology is a continuous density HMM with one Gaussian mixture component and diagonal covariance matrix. There were 5 states in total, 3 of which are emitting. As an acoustic modeling units, we used context independent (CI) monophones, context dependent (CD) triphones and whole word (WW) models. Despite more frequent use of sub-word modeling (CI and CD) in ASR systems, there are still some applications where whole word modeling presents optimal solution, especially in recognition of isolated and connected words from constrained lexicon. The WW models consisted of the same number of states as their CI and CD counterparts and followed the same transition matrix structure, that is strictly left-right, with no skips. For CI models, phonetic transcription was done manually. Stops and affricates are labeled as pairs of semi-phones that consist of occlusion and explosion parts. Phoneme /ə/ (schwa) is marked separately when phoneme /r/ is found in the consonant environment [17]. The model of silence is added at the start and the end of every utterance. Initial model parameters were estimated using the flat-start method , since training data is not time labeled. In the training phase, location of word boundaries were estimated using forced alignment. At last, in the testing phase the Viterbi algorithm was applied to determine the most likely model that best matched each test utterance.

Our goal was to compare the performance of different acoustic models in four train/test scenarios:

1. Normal/Normal (Nor/Nor) - the system is trained on normally phonated speech and tested on the speech of same mode;
2. Normal/Whisper (Nor/Whi) - the system is trained on normally phonated speech and tested against whispered speech;
3. Whisper/Normal (Whi/Nor) - the system is trained on whispered speech and tested against normally phonated speech;
4. Whisper/Whisper (Whi/Whi) - the system is trained on whispered speech and tested on the speech of same mode.

The scenarios where training and testing is in the same mode of speech are denoted as match, and in the opposite case scenarios are denoted as mismatch.

IV. RESULTS AND DISCUSSION

In match scenarios, 80% utterances were in the part for training, while the other 20% were in the part for testing. The training and test set were rotated, which gave 4 additional tests. Recognition rate is calculated as mean value of 5 tests. In mismatch scenarios, all utterances of one mode were in the part for training, while in testing part were all utterances of the other mode of speech. The results (word recognition rate - WRR) are shown in Tables I-III for modeling monophones, triphones and whole words, respectively. The speakers are labeled from G1 up to G5. The last column presents average WRR for respected scenario and feature vector. For better clarity, results are integrated, and depicted in Fig. 4 (PLP feature vector) and Fig. 5 (MFCC feature vector), where average recognition rates of all five speakers are graphically presented in dependence of scenario and modeling units. From results presented in Tables I-III, and Figures 4-5, we can observe that CD models contribute to higher scores in match scenarios, compared to CI and WW models. These results are expected, since CD models are more specialized and superior in highly matched conditions, which is the case with Whi-Spe corpus. Performance of CI and WW models could be improved (WRR above 99.5%) by increasing number of mixture components [18].

TABLE I. WORD RECOGNITION RATE FOR MONOPHONE MODELS

Speaker/ Scenario-Feature		G1	G2	G3	G4	G5	Avg
Nor/Nor	PLP	98.6	98.0	98.8	98.0	98.0	98.28
	MFCC	98.2	99.2	98.0	98.0	98.2	98.32
Nor/Whi	PLP	89.2	52.6	76.0	55.4	61.0	66.84
	MFCC	85.4	47.8	70.4	47.2	58.6	61.88
Whi/Nor	PLP	87.6	57.6	75.6	80.4	72.0	74.64
	MFCC	87.2	56.8	70.8	81.6	72.0	73.68
Whi/Whi	PLP	96.6	94.2	97.8	97.2	90.6	95.28
	MFCC	96.4	94.2	98.0	97.0	90.4	95.20

TABLE II. WORD RECOGNITION RATE FOR TRIPHONE MODELS

Speaker/ Scenario-Feature		G1	G2	G3	G4	G5	Avg
Nor/Nor	PLP	99.8	100.0	100.0	99.8	99.8	99.88
	MFCC	100.0	100.0	100.0	99.8	99.8	99.92
Nor/Whi	PLP	50.0	12.6	35.6	20.0	21.4	27.92
	MFCC	41.6	12.0	31.8	14.8	12.4	22.52
Whi/Nor	PLP	73.8	34.0	53.2	56.8	60.8	55.72
	MFCC	73.4	33.8	51.0	61.4	60.4	56.00
Whi/Whi	PLP	100.0	100.0	100.0	100.0	99.8	99.96
	MFCC	100.0	100.0	100.0	100.0	99.8	99.96

TABLE III. WORD RECOGNITION RATE FOR WHOLE WORD MODELS

Speaker/ Scenario-Feature		G1	G2	G3	G4	G5	Avg
Nor/Nor	PLP	99.2	99.2	99.8	99.8	99.2	99.44
	MFCC	98.0	99.4	99.2	100.0	99.6	99.24
Nor/Whi	PLP	57.2	22.6	41.0	35.0	37.4	38.64
	MFCC	49.2	22.8	44.6	34.2	33.2	36.80
Whi/Nor	PLP	59.8	29.8	38.0	42.0	43.4	42.60
	MFCC	57.2	25.8	40.8	49.8	41.8	43.08
Whi/Whi	PLP	98.2	98.8	98.8	97.2	93.8	97.36
	MFCC	98.2	98.6	99.0	98.0	93.2	97.40

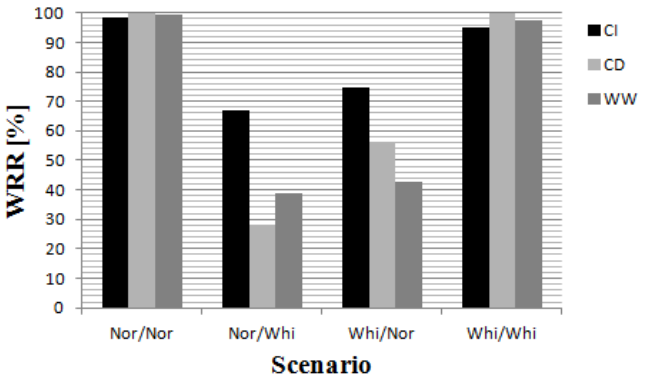


Figure 4. Average word recognition rates with PLP feature vector for context independent (CI), context dependent (CD) and whole word (WW) models and different scenario

Also, except CD modeling, in match conditions recognition of whispered speech is a few percents poorer. Both feature vectors give very similar results, and because of the "ceiling effect" it is hard to determine which is better. In mismatch scenarios, the most robust are CI models, with the average WRR of 66.84% in Nor/Whi scenario, and 74.64% in Whi/Nor scenario.

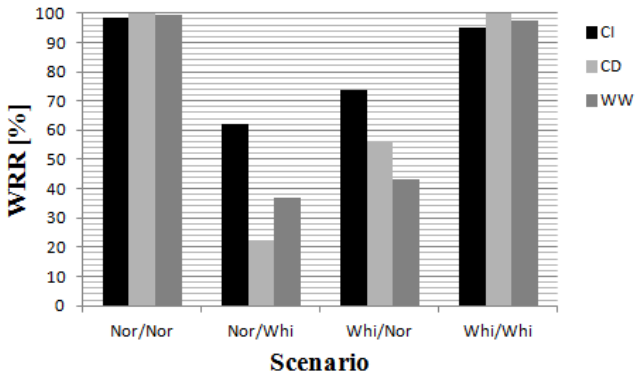


Figure 5. Average word recognition rates with MFCC feature vector for context independent (CI), context dependent (CD) and whole word (WW) models and different scenario

Utilization of PLP feature vector give an absolute improvement in average WRR of 5%, compared to recognition with MFCC feature vector. In experiments with same database and speaker independent fashion, PLP had also shown better performance, compared to MFCC feature vector [19]. The phenomenon of asymmetric performance in Nor/Whi and Whi/Nor scenario, in favor of Whi/Nor scenario, was examined in experiments with neural networks and same database [20]. Same study had shown that the sonority in speech stimuli is the main cause of difference in word recognition scores in mismatch scenarios, and that the most of whisper features are contained in normal speech, which is not the opposite case. The CD and WW give much lower recognition scores in mismatch scenarios, for both feature vectors.

Since the most of ASR systems are primarily trained on normally phonated speech, from the point of view of ASR recognition of whispered speech the greatest importance have the Nor/Whi scenario. The greatest effort in research area is devoted to maximizing performance in that scenario, because that concept does not need adaptation to whisper, of any kind. The best recognition scores for whispered speech with constrained lexicon (160 words) for recognition of English [21] are over 80% (in speaker independent fashion), so there is optimistic expectation that those scores are not far away for Whi-Spe database, for recognition with statistical ASR framework. First step is made in reduction in number of HMM models. The general approach in transcription of monophones, which includes separate modeling of inclusion/explosion parts in stops and fricative, and stressed/unstressed vowels, requires large databases for training, of several hours. That is not the case with Whi-Spe database. The unification of occlusion and explosion parts in stops and affricates, as well as stressed and unstressed vowels, has lead to reduction of HMM models to 32 monophones (30 phonemes in Serbian, schwa and silence). The results and absolute improvement with reduced number of HMM models are presented in Table IV.

TABLE IV. WORD RECOGNITION RATE IN NOR/WHI SCENARIO WITH REDUCED NUMBER OF HMM MODELS

Speaker/ Feature	G1	G2	G3	G4	G5	Avg	Improve-ment
PLP	93.6	62.8	81.4	64.0	71.0	74.56	7.72
MFCC	92.8	61.8	78.8	58.6	70.8	72.56	10.68

Results presented in Table IV clearly show greater robustness of models with reduced number of HMM models, compared to generalized case. Absolute improvement of average WRR is 7.72% for PLP feature vector, and 10.68% for MFCC feature vector. Beside better improvement of average WRR with MFCC feature vector, average WRR with PLP feature vector is again higher. It is important to note that performance of recognizer are not much degraded in match scenarios with reduced number of monophone units. The average WRR with recognition of normally phonated speech is 98.08% with PLP feature vector, and 98.04% with MFCC feature vector. Also, in recognition of whispered speech, recognizer gives the WRR score of 96.16% with PLP feature vector, and 96.24% with MFCC feature vector. The recognition of whispered speech is even higher, by comparing results in Table I.

From results in Table IV, very high difference in performance among different speakers could be observed. The recognition rate varies from poor (speakers G2 and G4) to excellent (speaker G1). Similar observation is found in whispered speaker identification with neutral trained HMM models [22], where it was stated that the degradation is concentrated for a certain number of speakers, while other speakers displayed consistent performance to that seen in neutral speech. One of the reasons for that deviation is signal to noise ratio (SNR) of tested utterances. Also, driven by that observation we examined correlation between SNR and WRR of all five speakers. In Table V are shown average SNR of all recordings in whispered speech, for all female speakers. Because of the way the manual segmentation of recordings is done, last 100 samples are taken into account for calculating power of noise.

TABLE V. AVERAGE SNR FOR RECORDINGS OF FEMALE SPEAKERS

Speaker	G1	G2	G3	G4	G5
SNR [dB]	15.3	8.0	13.0	11.4	9.7

From results in Tables IV-V, in this experiment the highest SNR leads to highest performance, and vice versa, the lowest SNR leads to lowest WRR (for PLP feature vector). In order to quantify the degree of correlation between average SNR of tested utterances and corresponding WRR, we determined the coefficient of correlation. For two random variables X (with mean μ_X and standard deviation σ_X) and Y (with mean μ_Y and standard deviation σ_Y), the correlation coefficient and covariance are defined according equations 1 and 2, respectively. In equation 2, E denotes expectation of random variable.

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \quad (1)$$

$$Cov(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right] \quad (2)$$

Also, corresponding p-value is calculated for testing the hypothesis of no correlation. For SNR values given in Table V and WRR in Table IV (for PLP feature vector), correlation coefficient of 0.89 and corresponding p-value of 0.04 are obtained. Besides relatively small number of tested speakers, obtained values unambiguously show high and significant correlation.

V. CONCLUSSION AND FUTURE WORK

Whispered speech, as an alternative mode of speech production is not seldom used in everyday life. Inspired by the fact that whisper is well understandable in human to human communication, performance of statistical ASR approach based on HMM is examined in this paper, for whispered speech recognition and speaker dependent fashion. A part of Whi-Spe database with female speakers is used in this study. In match train/test conditions, recognizer had the best performance with context dependent models, where recognition of normally phonated speech, as well as whispered speech was nearly 100%. In mismatch scenarios, the monophone models had shown best results with average word recognition rate of 66.84% in Nor/Whi scenario, and 74.64% in Whi/Nor scenario (both with PLP feature vector). The greatest attention was paid to performance improvement of whispered speech recognition with models trained on normal speech. Reducing numbers of HMM models had lead to significant absolute improvement of word recognition rate of 7.72% for PLP, and 10.68% for MFCC feature vector. Due to high difference in performance among speakers, the hypothesis of correlation between tested average SNR and obtained WRR is tested. The results of hypothesis induce significant correlation.

Future work will examine performance of recognizer in multi-condition training, where training corpus is composed of normal and whispered utterances. It would be interesting to examine if the performance could reach separately trained conditions and the amount of whisper data added to training corpus for a satisfactory recognition. Since any part of data in training process had not been labeled, flat-start method in initialization of HMM models was used. Thus, to bootstrap a set of HMM models, part of utterances in training process is to be manually labeled. That activity is in a progress and preliminary results for completed speakers show noticeable improvement in recognition rate. Using alternative feature vectors will be examined, especially those which are reported to be very robust in highly mismatch condition. From the perspective of application of whispered speech in ASR, the greatest challenge will be optimizing performance without adaptation to whisper, in speaker independent fashion. Due to very high influence of amount of data in training to performance, speaker independent ASR recognizer is expected to have better performance, compared to speaker dependent recognizer. Also, the future research should include comparative analysis of word recognition efficiency using different algorithms such as DTW, HMM and ANN, using the same feature set and each speaker from database.

REFERENCES

[1] T. Ito, K. Takeda, F. Itakura, "Analysis and recognition of whispered speech," Speech Communication 45, 2005, pp. 139–152.

[2] S.T. Jovičić, Z. M. Šarić, "Acoustic analysis of consonants in whispered speech," Journal of Voice, Vol. 22, No. 3, 2008, pp. 263-274.

[3] S.T. Jovičić, "Formant feature differences between whispered and voiced sustained vowels," Acta Acustica, 84(4), 1998, pp.739-743.

[4] Y. Swerdlin, J. Smith, and J. Wolfe, "The effect of whisper and creak vocal mechanisms on vocal tract resonances," Journal of Acoustical Society of America 127, 2010, pp. 2590–2598

[5] X. Fan, J.H.L. Hansen, "Speaker identification with whispered speech based on modified LFCC parameters and feature mapping," ICASSP 2009, pp. 4553-4556, 2009.

[6] C. Zhang, J.H.L. Hansen, "Advancements in whisper-island detection using the linear predictive residual," ICASSP 2010, pp. 5170-5173, 2010.

[7] C. Zhang, J.H.L. Hansen, "Analysis and classification of speech mode: whisper through shouted," Interspeech 2007, pp. 2289–2292, 2007.

[8] Đ.T. Grozdić, B. Marković, J. Galić, S. T. Jovičić, "Application of neural networks in whispered speech recognition," Telfor Journal, Vol. 5, No. 2, pp. 103-106 , 2013.

[9] J. Holms, W. Holms, Speech synthesis and recognition. Taylor & Francis, London, 2001.

[10] B. Marković, S.T. Jovičić, J. Galić, Đ.T. Grozdić, "Whispered speech database design, processing and application," 16th International Conference TSD 2013, pp. 591-598, Pilsen, Czech Republic , 2013.

[11] S. Young, G. Evermann, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason,D. Povey,V. Valtchev, P. Woodland, The HTK Book v.3.2.1, Cambridge University Engineering Department, 2002

[12] S. T. Jovičić, Z. Kašić, M. Đorđević, M. Rajković, "Serbian emotional speech database: design, processing and evaluation," Proceedings of SPECOM-2004, St. Petersburg, Russia, pp. 77-81 , 2004.

[13] The Hidden Markov Model Toolkit, http://htk.eng.cam.ac.uk/

[14] S.B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 28, no. 4, pp. 357–366, 1980.

[15] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech," The Journal of the Acoustical Society of America, vol. 87, no. 4, pp. 1738–1752, 1990.

[16] J. Kacur, G. Rozinaj, "Practical issues of building robust HMM models using HTK and SPHINX systems," Speech Recognition, France Mihelic and Janez Zibert (Ed.), ISBN: 978-953-7619-29-9

[17] S. Sovilj-Nikić, V. Delić, I. Sovilj-Nikić, M. Marković, "Tree-based phone duration modeling of the serbian language," Electronics and Electrical Engineering (Elektronika ir Elektrotechnika), Kaunas University of Technology, Vol. 20, No. 3, pp. 1392-1215, 2014.

[18] J. Galić, S.T. Jovičić, Đ. Grozdić, B. Marković,"HTK-based recognition of whispered speech," SPECOM 2014, pp. 251-258, in press.

[19] J. Galić, S.T. Jovičić, Đ. Grozdić, B. Marković, "The influence of feature vector selection on performance on performance of automatic recognition of whispered speech," Speech and Language 2013, pp. 258-264 , 2013.

[20] Đ.T. Grozdić, S.T. Jovičić, D. Šumarac-Pavlović, B. Marković, J. Galić, "Whispered speech recognition with neural networks," unpublished.

[21] S. Ghaffarzadegan, H. Boril, J. H. L. Hansen, "UT-VOCAL EFFORT II: Analysis and constrained-lexicon recognition of whispered speech," ICASSP 2014, pp. 2544-2548, 2014.

[22] Fan, X., Hansen, J. H. L, "Speaker identification within whispered speech audio stream," IEEE Transactions on Audio, Speech and Language Processing, 19(5), 1408-1421, 2011.

Simple Arcade Game from Hardware Side using MicroBlaze

Milan Tucic, Ivan Kastelan
University of Novi Sad, Faculty of Technical Sciences
Novi Sad, Serbia
milan.tucic@rt-rk.com

Dragan Topalovic, Milos Nikolic
RT-RK Institute for Computer Based Systems
Novi Sad, Serbia

***Abstract*— This paper proposes reconstruction of the simple arcade game Pong using MicroBlaze soft core and explains hardware side of this process. Reconstruction is split into two time phases. In the first phase hardware is developed with functionalities which, as they are, are sufficient for game implementation. This means hardware part of system is made of simple graphic controller and paddle movement controller. The second phase covers system analysis and hardware made as a result of this analysis. All upgrades are done considering MicroBlaze's flexibility. This kind of approach and organization leads to better balance between hardware and software implementation, and demonstrates advantages of designing a system in which it is possible to monitor, analyze and influence the further evolution of the system.**

Keywords-game; hardware; microblaze; pong; vhdl

I. INTRODUCTION

FPGA configured by some Hardware Description Language (HDL) provides simple and fast development of complex digital circuits and systems. Developers have accepted this approach, which has led us to a lot of examples and solutions that we have today. Using them, development is easier, more flexible and it is giving additional time for research, rather than creating system from a scratch.

Good example of available solutions today is the Xilinx Embedded Development Kit (EDK), including Intellectual Property Catalog (IP Cores) and the MicroBlaze embedded processor soft core (MicroBlaze) with a reduced instruction set computer (RISC) architecture optimized for implementation in Xilinx Field Programmable Gate Arrays (FPGAs) [1].

Ke, Eric and Winston [2] were using this solution, while working on a student project of reconstructing Pac-Man game. They split the project into software and hardware programming, and with little effort on software side they got reduced complexity on hardware side. Game logic includes collision detection, scoring and enemy artificial intelligence (AI) and it is implemented using the MicroBlaze. Hardware configuration is described in Very Large Scale Integration HDL (VHDL), and it implements graphics engine with sprites and paddle control over the MicroBlaze UART IP Core.

As one of the earliest arcade video games, Pong game is used as referent simple arcade game in this paper. It has simple game logic and two-dimensional graphics. While other arcade video games, such as Computer Space, came before it, Pong was one of the first video games to reach mainstream popularity. It is a multiplayer tennis like game, where the goal is to defeat an opponent by earning a higher score. The game was originally manufactured by Atari Incorporated (Atari), who released it in 1972.

Armandas [3] has done full reconstruction of the Pong game in VHDL. He used Nintendo Controllers as an input device for paddle control. Ball and paddle objects are represented by graphics stored in memory, separated of text symbols which are used for text generation. Logic for game, which includes ball movement and paddle movement has also been implemented in VHDL.

This paper proposes an approach to accelerate drawing of graphics with hardware-based drawing using VHDL, MicroBlaze and IP Cores, suitable for reconstruction of the simple arcade game Pong [4] as one of the earliest arcade video games. VGA Controller handles communication with VGA screen and it is combined with modules for text generation and graphics memory. Game logic (scoring, ball and paddle movement) is handled by MicroBlaze. Paddle control has been done initially over the UART IP Core, but has changed (after system analysis) to less robust Joy Peripheral. This kind of approach brought additional time for modifications to graphics memory and this system module was upgraded with hardware sprites.

The rest of the paper is organized as follows: section 2 gives the overview of the proposed system. Section 3 explains graphics modules and first approach for paddle movement. Section 4 explains modifications done after system analysis. Section 5 gives results of the proposed approach implemented on Pong game. Finally, section 6 gives some concluding remarks and plans for future research.

II. SYSTEM OVERVIEW

This paper recommends Embedded Engineering Learning Platform (E2LP) [5] as a platform for research projects based on FPGA. The E2LP provides an advanced hardware platform that consists of a low cost Spartan-6 Platform FPGA surrounded by a comprehensive collection of peripheral components that can be used to create a complex embedded system. Additionally, software IDE is developed to support usage of the board. In this paper Spartan-6 is used for VHDL configuration, movement control is done over joy buttons and

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, under grant no. TR32030, year 2011-2015

VGA controller is using VGA connector for connection to a VGA monitor.

The central figure of the proposed system implemented on the E2LP is the MicroBlaze - soft 32-bit RISC processor. MicroBlaze [1] is implemented with Harvard memory architecture. Separate address spaces for instruction and data accesses are used. Both instruction and data interfaces of MicroBlaze are 32 bits. There is no separate access to I/O and memory (it uses memory mapped I/O). This system for Block RAM (BRAM) memory accesses uses Local Memory Bus (LMB) and Processor Local Bus (PLB).

Software programming is not described in this paper, but it is done in "C" programing language on MicroBlaze. Every peripheral has driver with functions that are important to them. Some functions were just abstraction of peripheral's native functionalities, but there were also functions with extra algorithms like object drawing. For accesses to peripherals AXI4-Lite IP Interface (AXI Lite) is used (Fig. 1).

III. PHASE I

This time phase of system evolution presents hardware with functionalities which, by themselves, are sufficient for game implementation. That means in this phase hardware is made of a graphics controller and a controller for paddle movement. The rest of the system, implemented in this phase, including scoring, ball and paddles drawing is implemented on MicroBlaze. Everything done here, will serve as a referent point for the phase II.

A. VGA Controller

VGA standard represents an easy way of forming images on the screen. An image is formed by placing successive pixels of the corresponding color, which are always placed from the relative position (0, 0) until the end of the starting line (in the case of 640x480 screen, it is pixel at position (0, 639)), when switching to the next line and so on until the end (in the case of 640x480 screen, it is pixel at position (479, 639)) (Fig. 2).

VGA interface defines 4 digital and 3 analog signals. Digital signals are used for synchronization and positioning. The end of setting a single screen line is signalized by a horizontal synchronization signal (hsync). Vertical synchronization signal (vsync) is signalizing the completion of setting the entire image on the screen. Signals for positioning are indicating which row and column are going to be set on the screen. Analog signals are used to determine the color which is

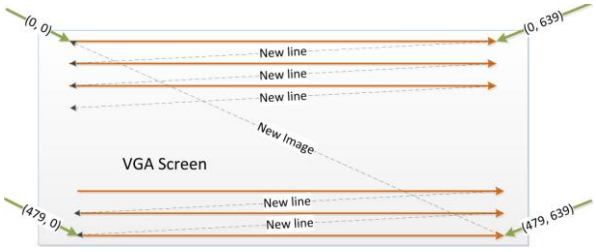


Figure 2. Image Process on VGA Screen

going to be displayed at a given position. The color is formed by combining the illumination intensity of three RGB format components. The intensity of each color is represented by 8 bits; hence each pixel on the screen is represented by 24 bits.

Screen content control relies on modules for text printing, graphics memory and set of registers. Using "pure" VGA controller directly with signals described above is complicated and it is not recommended. Because of that, these three modules were implemented to abstract VGA controller, and give simplified control for screen content.

1) Text Printing: Printing text on the screen is based on addressing a specific symbol from the BRAM memory. The BRAM memory contains predefined values as symbol representation for each symbol from a set of given symbols. Symbols are defined as 8x8 binary values. Giving symbol address, current row and column, module generates screen output value (Fig. 3). This module organizes screen in partitions. One partition is representing 8x8 pixels from the screen. Memory word stores address of a symbol and it maps one screen partition. This means it is only possible to set symbol to fit one of these partitions, not to a random position.

2) Graphics Memory: Graphics memory is used as an extra layer for printing desired values on the screen. The words in this memory are 32 bits wide and represent 32 successive pixels on the screen. Successive locations in memory are mapped to successive positions on the screen. At the end of each line of the screen, the memory locations wrap to the beginning of the next line. Graphics memory is accessed every time when value for the current position on screen is requested (also as text memory). Each of bits indicates which of the two predefined colors will be chosen for current point position. This kind of organization leads to easy screen content control, and everything is executed from MicroBlaze. Also, VGA controller generates interrupts which are processed

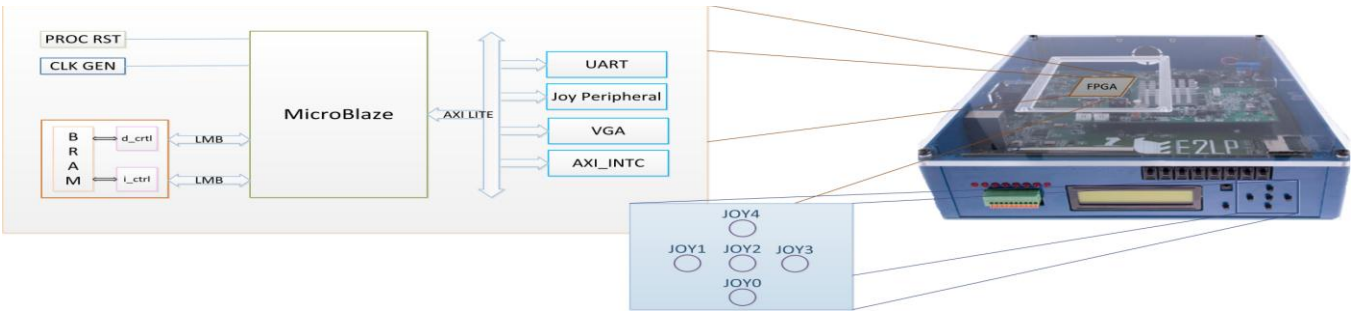


Figure 1. E2LP with MicroBlaze System and Joy Buttons

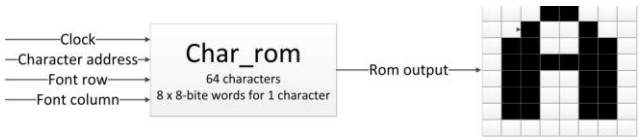


Figure 3. Text Module

by LogiCORE™ IP AXI Interrupt Controller (AXI INTC). These interrupts could be used as the clock for screen content refreshment (Fig. 4).

3) Registers: Some of the values for graphics modules are stored here. They define graphical mode, border line and colors (foreground, background and border). Registers for graphical mode determine whether the memory for both modules forms a picture, or the memory for just one module. Border line refers to the whole screen border and it can be set or not. This module is also accessible from MicroBlaze, and register content can be controlled from there.

B. *Paddle Movement Controller (PMC)*

The system uses the LogiCORE™ IP AXI Universal Asynchronous Receiver Transmitter Lite (UART Core) [6] to redirect commands from keyboard to the MicroBlaze. UART Core provides interface between UART signals and the Advanced Microcontroller Bus Architecture (AMBA) and also provides a controller interface for asynchronous serial data transfer. UART Core is enabling usage of keyboard buttons, but they must be filtered for reaction on specific buttons. In this case "S" and "K" keyboard buttons are used to move paddles up, and "X" and "M" buttons are used to move paddles down. This process is handled from MicroBlaze.

IV. PHASE II

The first phase concentrated on implementation of the basic Pong game. In the second phase, which is the main contribution of this paper, Pong was reconstructed using MicroBlaze and IP Cores as representative solutions for systems based on FPGA with the goal of accelerating the graphic processing and a broader usage of E2LP platform to make playing more user-friendly. Functionalities are implemented from both software and hardware sides. Phase II upgrades the system (Fig. 6), based on comparison to the similar systems and directly comparing software with hardware from phase I (Fig. 4).

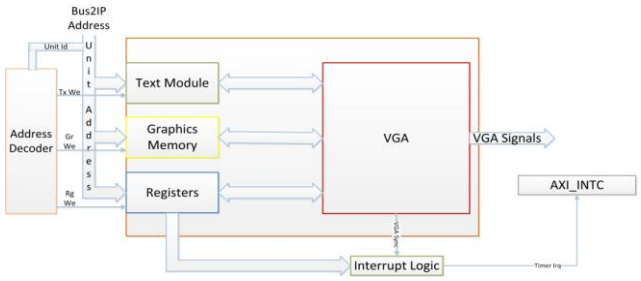


Figure 4. VGA Peripheral – Phase I

Figure 5. Project Results after Phase I

A. *Joy Peripheral*

Joy Peripheral (Fig. 7) holds PMC functionality and logic that is easy to handle. It controls signals from E2LP Joy buttons (Fig. 1). For purpose of this system implementation two instances of Joy Peripheral are used, controlling signals from buttons JOY1 and JOY4 (for up paddles direction) and JOY0 and JOY3 (for down paddles direction). Paddle control is based on Joy Peripheral interrupts handled from MicroBlaze. Interrupts are triggered at the moment when peripheral state is changed, and this depends on signals coming from JOY buttons.

Three states are defined, and here is an example for one instance:

1) UP: Peripheral state UP is set when button JOY1 is pressed, but JOY0 is not pressed.

2) DOWN: Peripheral state DOWN is set when button JOY0 is pressed, but JOY1 is not pressed.

3) IDLE: Peripheral state IDLE is set in cases which are not defined in 1 and 2.

After interrupt happens, defined state register holds current peripheral's state. Accessing to this register from MicroBlaze, paddle position on the screen can be updated.

B. *Hardware Sprites*

In the phase I object drawing is done by setting values that represent objects to graphics memory; every time when object changes its position graphics memory must be updated. This kind of approach has complex processing which must be provided by the MicroBlaze, particularly because memory architecture (32-bit words) for graphics memory is not ideal for setting object of any size to random position. Also, simply setting two objects to the same memory word will result in overwriting the first of these two. Because of this behavior, software side must handle this relatively complex situation.

In early video gaming, hardware sprites were a method of compositing separate bitmaps so that they appear to be part of a single image on a screen. Simple arcade games often work with predefined graphics objects, which fully meet the definition of hardware sprites. It enables much easier screen content control and that is the reason why they are implemented here. In contrast to the full control of an object drawing, which must be provided from the MicroBlaze, this kind of approach dramatically reduces effort on software side. All that is necessary is to fill the memory with sprite bite representation and then control them by setting starting pixel's position. This process is similar to text printing described in section 3, only it has better resolution defined with 4x4 pixels partitions and it uses blocks to represent one sprite. Block is defined by number of cells in row and column. For example, block for ball sprite contains 4x4 partitions (16x16 pixels). Sprite memory is

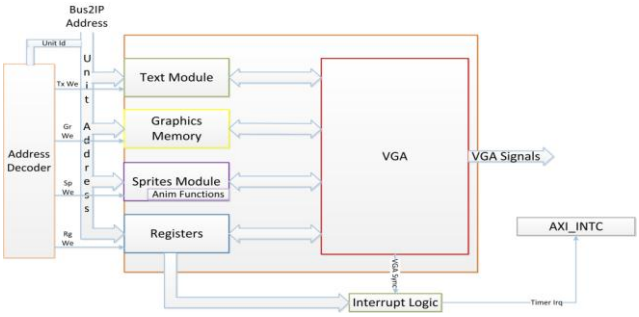


Figure 6. VGA Peripheral – Phase II

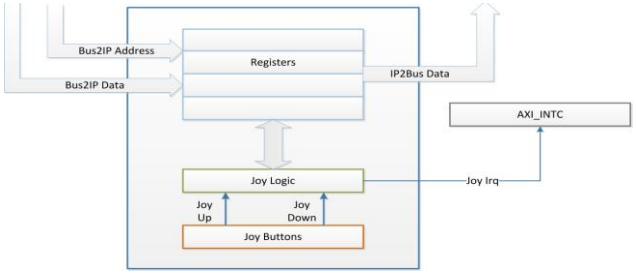


Figure 7. Joy Peripheral

representing sprites on the screen. While writing to this memory, it is necessary to provide starting address from BRAM memory for current sprite and number of rows and columns (cells) occupied with this sprite.

Sprites implemented here are not static. Sprite can be rotated and shifted to create simple animations effect. This procedure uses sprite bite representation and modifies output depending of function demanded.

Supported functions are:

1) Rotation: Using original sprite, simulates rotation for 90, 180 or 270 degrees.

2) Shifting: Shifts original sprite for the given number of bits.

V. RESULTS

PMC over UART Core, described in section 2 (also proposed by Ke, Eric and Winston [2]) is large overhead for this kind of systems. Paddle movement is controlled with 4 buttons and it is unnecessary to include whole UART Core for that. After comparing PMC from section 2 with Armandas's [3] Nintendo Controllers, it was clear that it is better to have simple peripheral for PMC. Also, comparing to Armandas, including peripheral to MicroBlaze is much easier than creating new peripheral for system proposed by Armandas. Very important fact is that peripheral in that kind of systems must be considered from the beginning. Adding features like this in the middle of project progress can be very complex (or impossible), which is not case with system that proposes this paper.

Graphics memory seems to be sufficient from the beginning, but after increased effort needed on software side it was clear that extra abstraction layer for graphics functionalities is necessary. Hardware Sprites were good solution, and this is confirmed by results of measuring execution time needed for drawing objects in phase I and phase II (Table I). Drawing object, in phase I, depends of object position, so execution time is not always the same. For this case, time in table I represents average value from 1000 samples of the same object drawing. In phase II there are no variations like in phase I; every sample within 1000 lasts equally. Phase II extremely reduced execution time compared to phase I in object drawing.

Extra features are also included to hardware sprites. It is possible to create simple animations, gaining user experience with rotation and shifting functionalities. Approach proposed by Armandas uses multiple sprites to represent different states of an animation (also possible here). Comparing to this approach, there is no need to store manually every image necessary for an animation, but they can be used only for simple animations.

TABLE I. OBJECTS DRAWING COMPARISON

	Execution time	
	Phase I	***Phase II***
Ball	6804 ns	7 ns
Paddle	7785 ns	7 ns

VI. CONCLUSIONS

The advantage of using MicroBlaze as a representative solution is the design flexibility during the progress of the system development. System's modules which require a lot of effort to be done by processor (software part of a system) can be reduced with hardware implementation of certain functionalities. Also, hardware modules of a system can be modified and their evolution brings better experience and easier development to the rest of the system. Choosing platform for research is also important. E2LP has wide range of peripherals and it is possible to use them to improve a system.

In the future research, system will be further improved to become a framework to support simple game generally, not to be bound to Pong game. Rotation will be considered to support any degree set. Also, more functionality for hardware sprites will be considered.

REFERENCES

[1] Xilinx, Inc., MicroBlaze Processor Reference, Guide UG984 (v2014.1), April 2, 2014

[2] Ke Xu, Eric Li, Winston Chao, "Pac-man final project report", University of Columbia, July 4, 2014

[3] Armandas Jarusauskas, "FPGA based VGA driver and arcade game", University of Sussex, 2009/2010

[4] Pong game, www.ponggame.org (accessed: August 10, 2014)

[5] I. Kastelan, N. Teslic, M. Temerinac: "E2LP: An Embedded Engineering Learning Platform", *IT System Conference 2013*, pp. 1-4

[6] Xilinx, Inc., LogiCORE IP AXI UART Lite v2.0 Product Guide PG142, April 2, 2014

[7] Pong P. Chu, "FPGA prototyping by VHDL examples", John Wiley & Sons, Inc., 2008

System for automatic testing of Android based digital TV receivers

Marko Kovacevic, Branimir Kovacevic
University of Novi Sad
Faculty of technical sciences
Novi Sad, Serbia
kovacevic.mare@gmail.com, brankomir@gmail.com

Dejan Stefanovic, Vukota Pekovic
RT-RK DOO
Novi Sad, Serbia
dejan.stefanovic@rt-rk.com, vukota.pekovic@rt-rk.com

***Abstract*—With constant increasing of digital TV receivers complexity and constant need for reducing the product time-to-market - the development of reliable and effective system for automatic testing of digital TV receivers becomes highly desirable. This paper proposes system for automatic testing of Android based digital TV receivers which is able to cover the entire process of product testing – from requirement definition, through verification test plan creation, automatic test cases execution and testing reports generation. System consists of web tool responsible for requirement and test management and stand-alone application responsible for automatic tests execution based on Black Box Testing methodology.**

Keywords-digital tv receiver, android, black box testing, set-top box, automated testing

I. INTRODUCTION

Modern receivers of digital TV signal offer a lot of different functionalities to end users, with a tendency to increase these functionalities with each new generation [1]. Standard set of TV features such as live TV channels playback (DVB-T, DVB-C, DVB-S), electronic program guide (EPG), Teletext, Subtitles, MHEG, personal video recorder (PVR), picture-in-picture (PiP) is extended with introduction of the Internet in the sphere of digital television. New features such as IPTV channels playback, hybrid broadcast broadband TV (HbbTV), video on demand (VOD), catch-up TV emerged.

With increasing hardware capabilities modern receivers become capable of executing complex applications written in Java programing language and besides traditional TV receivers based on Linux operating systems, new generation of TV receivers based on Android operating system appears.

With introduction of Android to digital television [2][3] - set of TV features is expanded with new features supported within Android OS. Besides TV functionalities Android OS offers a lot of interactive and non-interactive services and applications that can be installed and executed on the digital TV receivers.

With such increased complexity functional testing of TV receivers became very challenging and time demanding task. A minimal set of test cases required for testing of such complex system is 2.000. Some real examples (Android based Set-top box) showed that it takes approximately 5-10 minutes to carry out one test case manually. So execution of test suite that contains 2.000 test cases requires more than 20 working days. If it is assumed that in development phase of the project software release is typically launched each two weeks it is obvious that even in this stage manual testing is not acceptable. Manual testing is time-consuming and requires a lot of human resources. In later stages of project - software releases are more frequent and it is required that tests are executed frequently and repetitively so the human effort must be even more increased.

Therefore, manual testing of Android based digital TV receiver as a time consuming, human effort consuming and prone to errors due to manual mistakes, needs to be replaced with effective and reliable automated test system.

Currently available systems for testing of digital TV receivers are handling only test management process which facilitates manual testing or are capable to automatize small set of test cases needed for testing complex digital TV receivers.

One of the attempts for creating system which is able to cover all segments of verification and testing process of TV receivers is presented in [4]. This solution handles test management, test execution and reporting process during TV receivers' verification and testing.

This paper presents improvement of previously mentioned solution in the section of test management, reporting process and extension of test execution part with support for Android-specific feature testing. Test management and reporting are integral part of web tool so access to tests and results of test execution is possible from anywhere, at any time. Solution is able to cover the entire process of digital TV receiver testing – from initial requirement definition, through creation of verification test plan, automatic test cases execution and generating reports of testing. It is capable to automatize large part of the tests needed for reliable and effective testing of digital TV receivers based on Android OS.

The rest of the paper is organized as follows: Section II gives general overview of the proposed system for automatic testing. In Section III web-tool responsible for requirement and test management, test plan management and reporting is

presented. Section IV gives detailed description of module for automatic execution of test cases using Black Box Testing methodology (BBT). The conclusions are drawn in Section V.

II. THE PROPOSED SYSTEM FOR AUTOMATIC TESTING

General overview of the proposed system is depicted in Fig. 1. Central part of the system is server which is accessed by clients that may require different services, such as requirement management, test management, test plan management, automatic test execution and report generation. Web based client application is used to provide access to all services of the server except automatic test execution using REST (Representational state transfer) services exposed by server. For automatic test execution, stand-alone Executor application is used together with appropriate hardware for test execution. Application uses test plans made through web based client application to autonomously execute tests. After tests are executed, results of execution can be obtained through web based client application. Executor application uses SOAP (Simple Object Access protocol) web service exposed by server to retrieve test plan information and submit execution results.

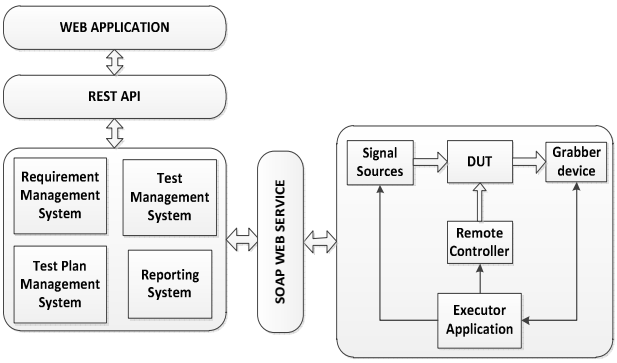


Figure 1. General system overview

III. REQUIREMENT AND TEST MANAGEMENT

Requirement and test management is implemented as web application which consists of four main sections (Fig 2.):

- Requirement management
- Test management
- Test planning
- Reporting

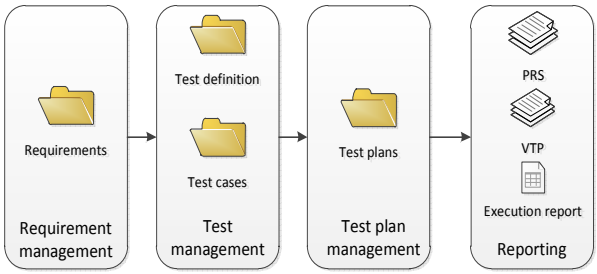


Figure 2. Requirement and test management application

Each section is represented as a top level node in a project tree. Project tree is the main interaction point with user, with all important application functions exposed in form of context menu options applicable on tree nodes.

Requirement management section is responsible for creation, modification and maintenance of project requirements specifications and project requirements. It provides options such as creating/editing/deleting of requirements specifications and requirements, assigning test definition to requirement, assigning documents to requirements and requirements specifications.

Test management section is responsible for creation, modification and maintenance of test suites, test definitions and test cases. It provides options such as creating/editing/deleting of test suites, test definitions and test cases, establishing relation between requirement and test definition, assigning documents to test suites and test definitions. For automatic and semi-automatic test cases test scripts can be added and modified in embedded Python script editor. Referent picture, video, audio and text files needed for automatic and semi-automatic test cases can be assigning to each test case.

Test plan management section is responsible for creation, modification and maintenance of execution test plans. This module enables composing of test plans by simply dropping test suites from test section into test plan, assigning users to test suites or test cases, assigning configuration for testing to test cases which defines devices and modules necessary for test execution (device under test, signal sources, grabber devices, remote control devices, algorithms, etc.). Execution of manual test cases is also possible within this section. It provides detailed description of test with the name of the test case, list of equipment necessary for test execution, description of test case, instructions how to setup test environment, list of inputs for test, test steps that needs to be executed. Test execution time is measured for each test execution. Outcome of each test may create/update issue in issue tracking systems such as Redmine if required.

Reporting section provides export of all data necessary for Project Requirement Specification (PRS) and Verification Test Plan (VTP) as well as reports of test plan executions. There are two types of test plan execution reports. Report for one test run where summary status of one execution is given, together with following information for each test case: test case name, description, execution result, bug opened in issue tracking system if test is failed, execution comment, duration of test execution, name of the tester/station that executed the test. Another type of report is test plan comparison report where execution status of each test through different test runs can be monitored. This report makes it easier to track project progress and to monitor what features are fixed and what are broken for each software release. It gives summary status of execution results through releases as shown on Fig 3. which facilitates project progress tracking and gives overview of quality and reliability of the device under testing. All results are available through web application or can be exported to different types of documents.

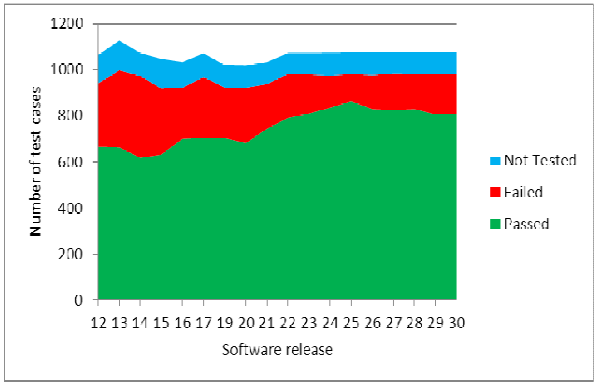


Figure 3. Test execution results

IV. MODULE FOR AUTOMATIC TEST EXECUTION

Executor is stand-alone application used to run automatic tests written in Python scripting language. Application uses test plans made in web application to execute tests. Application uses SOAP web service exposed by server to retrieve test plans and execution resources, as well as to submit execution results.

Black Box Testing (BBT) approach is used for automatic verification of digital TV receivers [5] that is described in this paper. TV receiver is considered as black box with unknown software/hardware architecture. Based on the test inputs, TV receiver produces outputs which are compared to the expected (referent) outputs in order to determine if STB is correctly functioning.

Executor application is able to control a large set of hardware and software modules needed for testing. Minimal set of hardware modules (controlled by Executor) required for testing of Android based digital TV receiver (Fig 4.) consists of:

1) *PC with Executor application*
2) *DVB-T, DVB-C and DVB-S stream modulators*
3) *Realtime audio video capturing device*
4) *Remote controller emulator*
5) *Power switch devices*

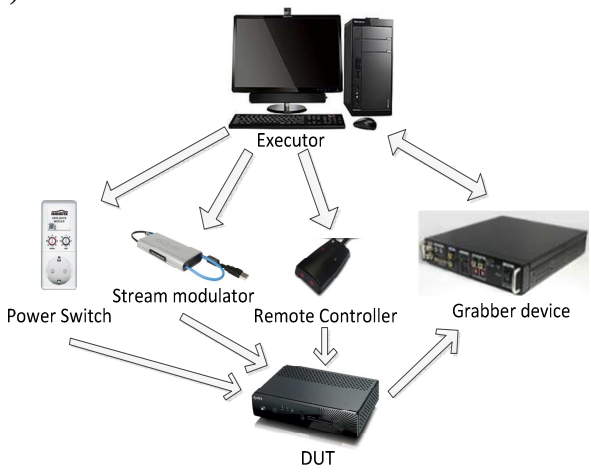


Figure 4. System for automatic testing

DVB-T, DVB-C and DVB-S stream modulators are used as signal generators which provide known input signals. Real-time audio/video capturing device [6] is used for image, video and audio capturing from different output interfaces which receiver might have (HDMI, S/PDIF, CVBS, SCART, COMPONENT, etc.). Device communicates with PC through network interface and transfers grabbed data to PC for further analysis. Remote controller emulator [7] is used to bring the Device Under Testing (DUT) to desired state. Power switch devices are intended for electrical power off/on of the DUT.

Minimal set of software modules (controlled by Executor) required for testing of Android based digital TV receiver consists of:

1) *Picture Block Compare module*
2) *Optical Character Recognition module*
3) *Video Quality Assessment module*
4) *Audio Quality Assessment module*
5) *Audio Compare module*
6) *Android Debug Bridge module*

Picture Block Compare algorithm [8] is used to compare picture captured from DUT output to referent picture. Result of algorithm is percent of matching of two pictures. It supports comparing of full pictures or only regions of the interest. It is commonly used for graphical user interface testing in Android applications, or for assess of image quality in digital television systems [9].

Optical Character Recognition module is used for extracting text from captured images. It is suitable in cases when we need to confirm that certain text appears on the image. It is commonly used for checking if EPG data is displayed correctly, if service list contains correct data, if UI is correctly displayed, if Teletext and Subtitles data are displayed correctly, etc.

Video Quality Assessment module contains a set of video analysis algorithms for analyzing video content in real time. Some of the available algorithms: blocking detection, freezing detection [10], black screen detection, packet loss, blurring, ringing. It is used for video decoding testing, video quality assessment of live and multimedia playback, testing of PVR, VoD or YouTube video playback.

Audio Quality Assessment module contains a set of audio analysis algorithms for analyzing audio content in real time. Some of the available algorithms: audio absence detection, audio level measurement, audio discontinuities (click), audio clipping. It is used for audio decoding testing, audio quality assessment of live and multimedia playback, PVR, VoD, YouTube audio testing.

Audio compare module allows comparison between two audio streams. Actual audio output of the DUT recorded with real-time audio/video capturing device is compared with the referent (expected) output. Signals are compared in frequency domain. It is used for audio decoding testing when expected audio output is known.

Android Debug Bridge (adb) module [11] is used for controlling of adb tool that is available on Android TV receivers. Adb provides many functions that can be used

during testing of Android based STB. Some of the functions are: obtaining log data from DUT which makes debugging process easier, performing of factory reset procedure on DUT, checking if DUT has internet connection, performing reboot of DUT, sending key events to DUT, starting of different applications that need to be tested, testing of functionalities such as video, audio and image decoding by starting appropriate application (*VideoView*, *ImageView*) and checking displayed content, pulling the desired files from the board for easier debugging, taking snapshots of user interface , etc.

V. CONCLUSION

This paper proposes system for automatic testing of Android based digital TV receivers using black box testing approach. Solution covers entire process of digital TV receiver testing – requirement definition, verification test plan creation, automatic test cases execution and report generation.

Presented system for automatic testing can be used for functional testing, regression testing, performance testing and stress testing of TV receivers. It improves accuracy, saves time and human effort in comparison to manual testing. It also enables automatic execution of tests that cannot be executed manually (audio/video synchronization tests, testing of playback trick modes such as fast forward and rewind, frame drops detection, measurement of GUI response).

For each executed test case system provides its execution status, captured images, recorded video and audio data and log data obtained during test execution. For failed test cases notification which test step failed and how to reproduce bug is available together with bug opened in issue tracking system. System facilitates project progress tracking and gives overview of quality and reliability of the device under testing – which is used for assessment when is DUT ready for market.

ACKNOWLEDGMENT

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant TR 32029.

REFERENCES

[1] Y. Wu, S. J. Hirakawa, and U. H. Reimers, "Overview of Digital Television Development Worldwide", Proceedings of the IEEE, Vol. 94, No.1, Jan. 2006, pp.8 - 21.

[2] M. Vidakovic, N. Teslic, T. Maruna, and V. Mihic, “Android4TV: a proposition for integration of DTV in Android devices”, 30th IEEE International Conference on Consumer Electronics (ICCE), January 2012.

[3] N. Kuzmanovic, T. Maruna, M. Savic,G. Miljkovic, and D. Isailovic, “Google's Android as an application environment for DTV decoder system”, 14th IEEE International Symposium on Consumer Electronics (ISCE), June 2010, pp. 1-5.

[4] Vukota Pekovic, Nikola Teslic, Ivan Resetar, Tarkan Tekcan, “Test Management and Test Execution System for Automated Verification of Digital Television Systems”, 2010 IEEE 14th International Symposium on Consumer Electronics.

[5] Dusica Marijan, Vladimir Zlokolica, Nikola Teslic, Vukota Pekovic, Tarkan Teckan, “Automatic Functional TV Set Failure Detection System”, IEEE Transactions on Consumer Electronics, Vol. 56, No. 1, FEBRUARY 2010.

[6] “Network Attached Video Capturing device – RT-AV100”, http://bbt.rt-rk.com/en/hardware/rt-av100

[7] V. Marinkovic, I. Kastelan, N. Vranic, Z. Marceta, V. Pekovic, “Protocol independent remote controller emulator for automated testing of multiple DTV/STB”, TELFOR 2012, pp. 1064 - 1067..

[8] D. Marijan, V. Zlokolica, N. Teslic and V. Pekovic, “Quality Assessment Of Digital Television Picture Based On Local Feature Matching”, Digital Signal Processing, 2009 16th International Conference on Volume , Issue , 5-7 July 2009 Page(s):1 – 6

[9] Vukota Pekovic, Nemanja Lukic, Ivan Kastelan, Nikola Teslic, “Platform For STB Video Output Integrity Verification Based On Full Reference Picture Comparison”, 2011 IEEE International Conference on Consumer Electronics - Berlin (ICCE-Berlin)

[10] Vladimir Zlokolica, Vukota Pekovic, Nikola Teslic, Tarkan Tekcan Miodrag Temerinac, Video Freezing Detection System for End-user Devices, 2011 IEEE International Conference on Consumer Electronics (ICCE)

[11] Android Debug Bridge,http://developer.android.com/tools/help/adb.html

Session TO-7

MODELING, IDENTIFICATION AND PROCESS CONTROL

Design and Tuning of PID Override Control System Based on Signal Filtering

Aleksandar I. Ribić
Institute Mihajlo Pupin,
Belgrade, Serbia
e-mail: aleksandar.ribic@pupin.rs

Miroslav R. Mataušek
Faculty of Electrical Engineering,
University of Belgrade
Belgrade, Serbia
e-mail: matausek@etf.rs

Abstract— **The proposed override control system consists of two anti-reset windup controllers, common actuator and a limiter, surrounded by pre and post biproper filters. Pre-filter is inverse of the post-filter. The desired limit on the override variable and its set-point are defined by output of the limiter. Signal selector is not applied in the proposed structure. Other characteristic feature of the proposed solution, compared to standard one, is that the override variable response is obtained practically without overshoot. Procedure for adjusting parameters of the post-filter are defined and illustrated in detail. Simulation results are used to demonstrate the basic ideas. Experiment on a laboratory thermal plant with noisy measurements is used to confirm validity of the proposed solution.**

Keywords-PID control; Override control; Constraints; Dead-time compensation; Tuning

I. INTRODUCTION

In many cases the number of variables to be controlled is higher then the number of manipulated variables. Override and cascade control are commonly applied to solve this control problem at the regulatory control level. In the present paper a new, effective solution for override control system design and tuning is proposed and experimentally verified.

In override control, one variable is a primary controlled variable y_p and have to be maintained at a given reference value r_p (or close as much as possible), but in such way that the override variable y_o had to be limited to a value r_o, defined in [1] as the soft constraint.

Block diagram of a standard override controller is presented in Fig. 1. Actuator output w is used as external reset feedback, common to both anti-reset windup controllers C_p and C_o. Two controller outputs are connected to a signal selector (min or max). The controller demanding higher or lower actuator output w will override the other. In the normal operating regime, this controller is the primary one, C_p. Stability of this standard override control system is considered in [2].

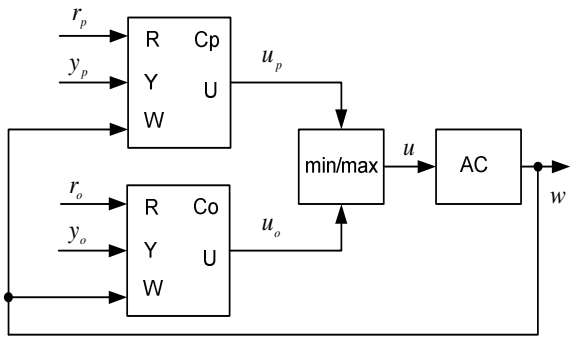


Figure 1. Block diagram of a standard override control system: primary controller C_p and override controller C_o connected to the common actuator AC trough the min/max selector.

Suppose that primary y_p and override y_o variables are defined by $y_p=G_p w$ and $y_o=G_o w$, where G_p and G_o are some process transfer functions. It is adopted that the selector in Fig. 1 is of a min type. In this case, when override variable y_o approaches closely to its limit r_o, then output signal u_o decreases. When it drops below the primary controller's output u_p, it is selected as actuator input u. This means that the controller C_o overrides the primary controller C_p, as demonstrated in Fig. 2 for the example with process transfer functions G_p and G_o defined by:

$$G_p(s)=\frac{e^{-0.2s}}{s},\ G_o(s)=\frac{10e^{-s}}{10s+1}, \tag{1}$$

and anti-reset windup PI controller used for the primary controller C_p and override controller C_o, defined in Section 2.

The basic problem with the override control in Fig. 1 is the overshoot in y_o response following the set-point r_p change. This is demonstrated in Fig. 3, for the above example defined by (1) and Fig. 1.

Appearance of overshoot in Fig. 3 reduces performance of override control. This conclusion follows from the fact that the override set-point r_o (soft constraint) has to be set sufficiently far from the interlock trip point, denoted in [1] as the hard constraint.

In the present paper a new structure of the override control system is proposed in Section II. It enables to solve efficiently the above overshoot problem in the case of set-point r_p change as well as in the presence of disturbances.

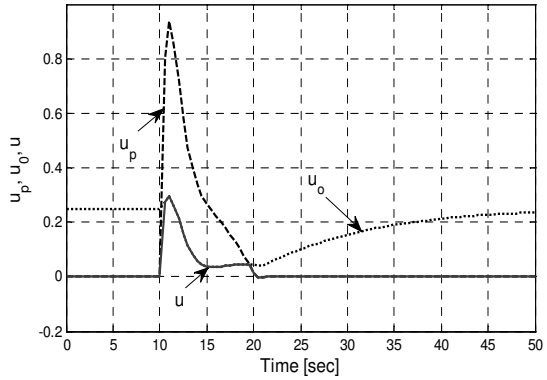


Figure 2. Control signals u_p (dashed), u_o (dotted) and resulting u (solid) of override control system in Fig. 1, with selector of a min type.

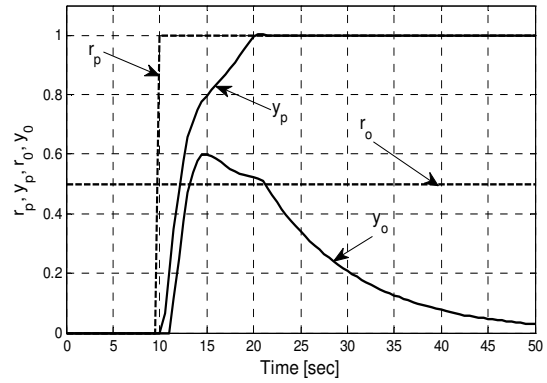


Figure 3. Set-point (dashed) and process output (solid) responses of override control system in Fig. 1. Standard override controller C_o can not prevent appearance of overshoot in the override variable y_o.

Design and tuning of the proposed override control system is presented in Section III. Then, for a laboratory thermal plant with noisy measurements, in Section IV the basic ideas are illustrated by simulation, and experimental results are used for final verification.

II. PROPOSED STRUCTURE

The proposed override control system is presented in Fig. 4. The characteristic feature of the proposed structure is the limiter surrounded with a biproper post-filter and its inverse, a pre-filter. Polynomial $A_2(s)$ is given by $A_2(s)=a_2s^2+a_1s+a_0$, $F_n(s)=1/(T_is+1)^n$ is an n^{th} order low-pass filter, $n\geq 1$, and L is a design parameter [3,4]. The same structure is used for the primary controller, with $F_r(s)\equiv F_n(s)$. Obviously, first-order filter $F_n(s)$ can be applied for $a_2=0$, as in Table I.

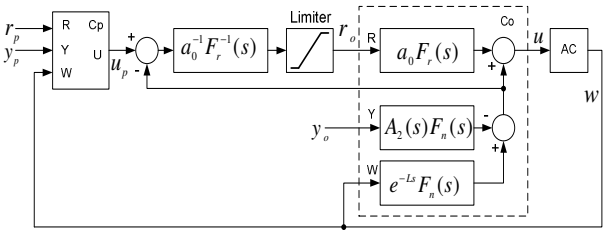


Figure 4. Proposed override control system: primary controller C_p and override controller C_o connected to the common actuator AC. Inside limits $u(t)\equiv u_p(t)$ and override controller C_o is inactive. $F_r(s)$ is a biproper post-filter with $F_r(0)=1$.

The proposed mechanism is simple and obvious. As the pre-filter is used to be inverse of biproper post-filter, inside limits one obtains $u(t)\equiv u_p(t)$ since the action of the override controller C_o in Fig. 4 is cancelled. When the set-point $r_o(t)$ reaches the limit, the connection relating primary controller to actuator is broken and override controller C_o is active with set-point defined by a desired limit.

The proposed structure is applied to the above example (1). To demonstrate advantages of the proposed override control system, the same PI controllers ($a_2=0$, $L=0$, $n=1$) are used in the standard override control system, with results presented in Figs. 2-3, and in the proposed structure, used to obtain results in Fig. 5. Parameters of PI controllers are given in Table I.

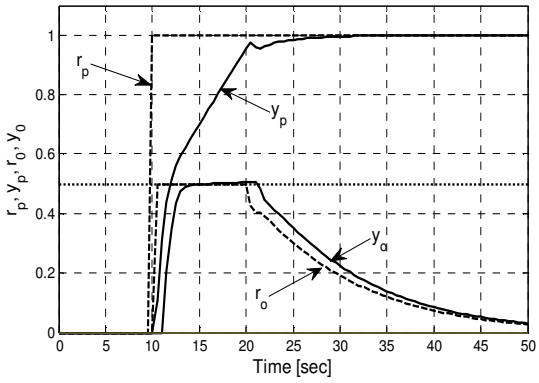


Figure 5. Set-point (dashed) and process output (solid) responses of the proposed override control system in Fig. 4. Dot line presents the desired limit for y_o.

TABLE I. PARAMETERS OF APPLIED PI CONTROLLERS

Controller	a_0	a_1	T_i
Override	0.5	2.5	5
Primary	1.25	1.25	0.5

Results presented in Fig. 5 confirm that the desired limit is almost strictly satisfied, by using the proposed override control system and the post-filter tuning derived in the next section.

III. DESIGN OF THE POST-FILTER

Assume that both controllers in Fig. 4 are tuned to satisfy some desired performance/robustness trade-off, for a given dynamic process characterization $y_p=G_pw$ and $y_o=G_ow$ defined by transfer functions G_p and G_o. This means that the post-filter $F_r(s)$ in Fig. 4 is designed assuming that G_o and controllers C_p and C_o are known.

Set-point response of y_o is defined by transfer function

$$G_{ro}(s)=\frac{Y_o(s)}{R_o(s)}=\frac{a_0G_o(s)}{1+F_n(s)\left(A_2(s)G_o(s)-e^{-Ls}\right)}, \quad (2)$$

for $F_r(s)\equiv 1$. To obtain fast set-point response of y_o without the overshoot, the ideal form of the post-filter is obtained as the inverse of $G_{ro}(s)$. However, this implementation is not always possible, and it is proposed to use a rational function approximation of $G_{ro}(s)$. Since $G_{ro}(0)=1$, this approximation is given by

$$G_{ro}(s)\approx\frac{1+C(s)}{1+D(s)}, C(s)=\sum_{i=1}^{p}c_is^i, D(s)=\sum_{i=1}^{q}d_is^i, p\le q\cdot \quad (3)$$

Coefficients of polynomials $C(s)$ and $D(s)$ are determined by minimizing the following criteria

$$\min_{C_m,D_m}\sum_{k=1}^{K}\left|\frac{1+D_m(i\omega_k)}{1+D_{m-1}(i\omega_k)}G_{ro}(i\omega_k)-\frac{1+C_m(i\omega_k)}{1+D_{m-1}(i\omega_k)}\right|^2, m=1,2,...,N, \quad (4)$$

performed through iterations, for $D_0(i\omega_k)\equiv 0$ and $\omega_k\in\Omega$, $\Omega=\{\omega_1,\omega_2,...,\omega_K\}$, where Ω is a set of frequencies in the desired range. It should be observed that from (4) one obtains $G_{ro}(i\omega_k)\approx(1+C_N(i\omega_k))/(1+D_N(i\omega_k))$ for $\omega_k\in\Omega$, as required by (3), since for $m=N$ it follows $D_{m-1}(i\omega_k)\approx D_m(i\omega_k)$.

Poles and zeros of $(1+C_N(s))/(1+D_N(s))$ in the left half s-plane, close to the imaginary axes, are used to approximate dominant dynamics of $G_{ro}(s)$. Right Half s-Plane (RHP) zeros are excluded. Then, the biproper post-filter $F_r(s)$ is designed as the inverse of dominant dynamics of $G_{ro}(s)$. To obtain the biproper post-filter $F_r(s)$ in the form

$$F_r(s)=\frac{1+B_r(s)}{1+A_r(s)}, B_r(s)=\sum_{i=1}^{m}b_is^i, A_r(s)=\sum_{i=1}^{m}a_is^i, \quad (5)$$

some additional zeros $(T_zs+1)^g$, defined by time constant T_z and order g, have to be included in polynomial $1+B_r(s)$, to obtain the same order m of $B_r(s)$ and $A_r(s)$ in (5). Further details will be demonstrated here and in the next section.

For the example (1), the following approximation of $G_{ro}(s)$ is obtained from (4):

$$\frac{1+C(s)}{1+D(s)}=\frac{(1-1.0977s)\ (4.1271s+1)}{(2.7988s+1)\ (1.2291s^2+1.7552s+1)}\cdot\frac{(0.0395s^2-0.1692s+1)}{(0.0308s^2+0.1725s+1)}.$$

The biproper post-filter $F_r(s)$, with time constant $T_z=0.5$ s, adopted as a trade-off between the performance and the high-frequency gain $F_r(\infty)$, is given by

$$F_r(s)=\frac{(2.7988s+1)\ (1.2291s^2+1.7552s+1)}{(4.1271s+1)(0.5s+1)^2}.$$

IV. SIMULATION AND EXPERIMENTAL RESULTS

A laboratory thermal plant, with noisy measurements, is used to demonstrate in detail the basic ideas and properties of the proposed solution and to verify experimentally the validity of the override control system in Fig. 4. In this section, DTC-PID controllers [3,4] are used as primary and override controller in the proposed override control system.

The plant is presented in Fig. 6. The temperature $T(x,t)$ of aluminum plate, long l=0.1 m and wide h=0.03 m, is measured by precision sensors LM35 (TO92) at positions x=0 and $x=l$. The plate is heated by terminal adjustable regulator LM317 (TO 220) at x=0. The input to the heater is the manipulated variable $w(t)$, obtained from the saturation element with limits l_{low}=0% and l_{high}=100% [4]:

$$w(t)=\begin{cases}u(t), l_{low}<u(t)<l_{high}\\ l_{low}, u(t)\le l_{low}\\ l_{high}, u(t)\ge l_{high}\end{cases}. \quad (6)$$

The primary controlled variable is $y_p(t)=T(l,t)$ [°C], while measurement at the position x=0 of the heater is used as the override variable $y_o(t)=T(0,t)$, to keep the temperature $T(0,t)$ below 59 °C.

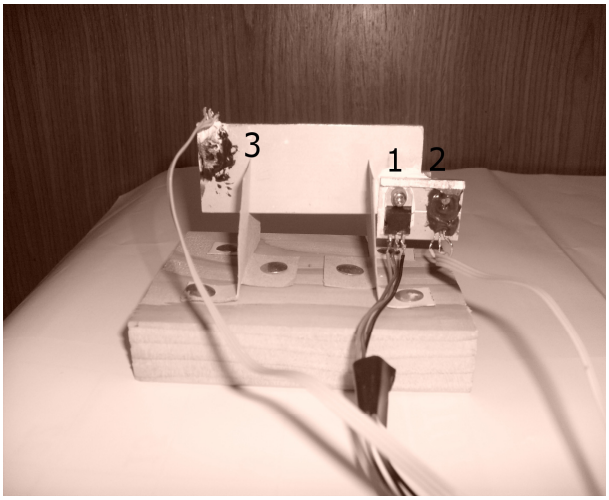


Figure 6. Laboratory thermal plant: 1- heater, 2- sensor at x=0, 3- sensor at $x=l$.

For the nominal regime, defined by the $T(l,t)\approx 50$ °C, the open-loop system responses of both outputs $y_p(t)$ and $y_o(t)$, obtained by applying a PRBS signal $w(t)$, are used to determine 100th-order ARX models, approximated then with the following models $y_p=G_pw$ and $y_o=G_ow$:

$$G_p(s) = \frac{1.375\,(3.627\,s+1)(1-1.295\,s)}{(580.05\,s+1)(4.85\,s+1)(3.24\,s+1)(201.13\,s^2+25.52\,s+1)},$$

$$G_o(s) = \frac{1.609\,(53.079\,s+1)(2.064\,s+1)}{(503\,s+1)(26.62\,s+1)(7.21\,s+1)(1.88\,s+1)(1.83\,s+1)}. \quad (7)$$

Both DTC-PID controllers are defined by the structure of the override controller in Fig. 4, with $F_r(s) \equiv F_n(s)$ used in the primary controller, as in [4]. In both controllers fourth-order filter $F_n(s)=1/(T_i s+1)^4$ is used. Parameters of DTC-PID controllers in Table II are obtained by optimization under constraints on the robustness and sensitivity to measurement noise [4].

TABLE II. PARAMETERS OF APPLIED DTC-PID CONTROLLERS

Controller	a_0	a_1	a_2	T_i	L
Override	14.276	265.983	1649.6	2.2191	0.1493
Primary	10.906	730.446	12785	5.0806	17.569

The following approximation of $G_{ro}(s)$ is obtained from (4):

$$\frac{1+C(s)}{1+D(s)} = -0.0002295\ \frac{(s-4.676)\ (s+0.01946)}{(s+2.167)\ (s+0.02023)} \cdot \frac{(s^2+0.4557s+0.06945)\ (s^2-0.737s+14.15)}{(s^2+0.1209s+0.008719)\ (s^2+0.2194s+0.0537)}. \quad (8)$$

Frequency responses of $G_{ro}(s)$ in (2) are used to define the desired range of frequencies $\omega_k \in \Omega$. It is proposed to use set of frequencies where amplitude characteristic of $G_{ro}(s)$ is greater than -60dB. Frequency responses of $G_{ro}(s)$ in (2), and its approximation (8), are presented in Fig. 7.

Filter $F_r(s)$ is obtained by applying design procedure (4)-(5) and taking into account only dominant poles and zeros of approximation (8). RHP zero s=4.576 and non-dominated pole s=–2.167 are neglected. Then, RHP zeros defined by $(s^2-0.737s+14.15)$=0 are neglected, and this term is replaced with $(7s+1)^2$ to obtain biproper $F_r(s)$ defined by:

$$F_r(s) = \frac{1+B(s)}{1+A(s)} = \frac{(49.4315s+1)}{(51.3875s+1)} \cdot \frac{(114.6921s^2+13.8663s+1)\ (18.622s^2+4.0857s+1)}{(14.398s^2+6.5616s+1)\ (7s+1)^2}. \quad (9)$$

Time constant T_z=7 s is adopted as a trade-off between the performance and high-frequency gain, to obtain $F_r(\infty) \approx 3$ as in the previous example with responses presented in Fig. 5.

Simulation of the laboratory thermal plant is performed first, by applying models $G_p(s)$ and $G_o(s)$ in (7) in the loop with the proposed override control system defined by Fig. 4, Table II and post-filter (9).

Results of this simulation, presented in Fig. 8, are obtained as follows. Limit for $y_o(t)$ is set to r_{omax}=6.5 °C. A band-limited white noise is added to both variables, $y_p(t)$ and $y_o(t)$. It is obtained from a band-limited white noise generator, with power PSD=0.0015 and cut-off frequency $\omega_C=\pi/0.5$ rad/s.

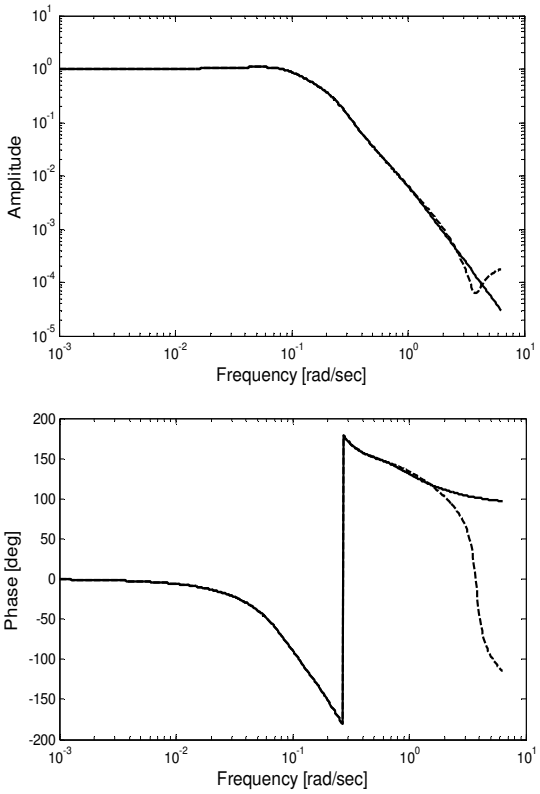


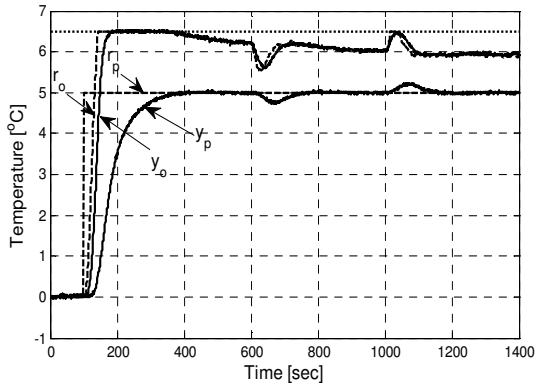
Figure 7. Frequency responses of $G_{ro}(s)$ (solid) and its approximation $(1+C(s))/(1+D(s))$ (dashed), obtained for $p=q=6$ in (3).

A step set-point $r_p(t)$ of amplitude 5 °C is activated at t=100 s. The set-point $r_o(t)$ is generated, followed by the override variable $y_o(t)$, regardless the fact that the override controller is inactive until $r_o(t)$ becomes equal to the desired limit. When $r_o(t)$ becomes equal to the desired limit, action of the override controller provides that the desired limit on override variable is strictly satisfied, as demonstrated in Fig. 8a.

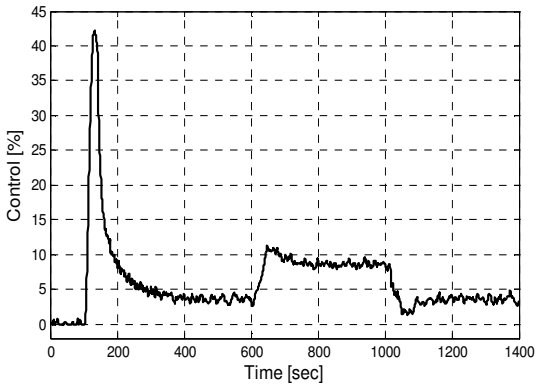
To illustrate performance of the proposed override control system in the presence of an unmeasurable load disturbance, a −5% step change of the control variable is inserted at time t=600 s and deactivated at t=1000 s.

Results of experimental verification of the proposed override control system in the loop with the real laboratory thermal plant are presented in Fig. 9. As in the simulation, set-point change of the primary variable equals 5 °C, from the nominal value of 45 °C.

In this case, limit on $y_o(t)$, given by r_{omax}=59 °C, and the new operating point are adjusted to demonstrate performance of the proposed override control in the vicinity of the desired limit. In the presence of unmeasurable disturbances, the desired override limit is not exceeded due to fast reaction of the override controller, as demonstrated in Fig. 9a.

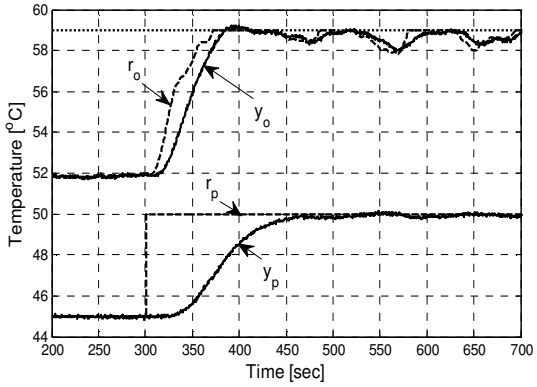


a)

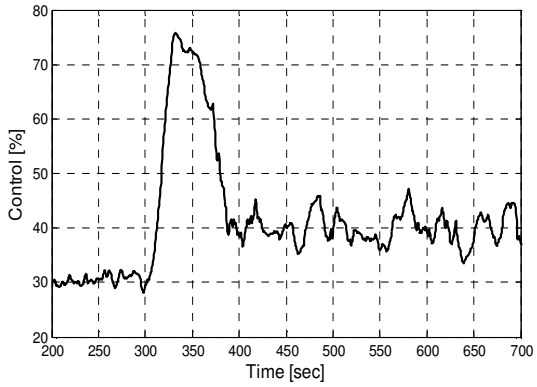


b)

Figure 8. Closed-loop simulation of the laboratory thermal plant, models $G_p(s)$ and $G_o(s)$ in (7) in the loop with the proposed override control system in Fig. 4: a) Set-points $r_o(t)$ and $r_p(t)$ (dashed), $y_o(t)$ and $y_p(t)$ (solid), and desired limit for $y_o(t)$ (dotted); b) control $w(t)$. A step change of control is activated at t=600 s and deactivated at t=1000 s.



a)



b)

Figure 9. Responses of the laboratory thermal plant in the loop with the proposed override controller. At t=300 s the set-point $r_p(t)$ is changed from 45 ^{0}C to 50 ^{0}C: a) Set-points $r_o(t)$ and $r_p(t)$ (dashed), $y_o(t)$ and $y_p(t)$ (solid), and desired limit for $y_o(t)$ (dotted); b) control $w(t)$. The desired override limit is not exceeded in the presence of unmeasurable disturbances.

V. CONCLUSIONS

Advances in PID like control algorithms are still of great importance, since, as demonstrated in [5] and confirmed recently in [6], PID controller still predominates on the regulatory control level. The proposed override controller offers an effective mechanism to deal with constraints in actuator and constraints on the process outputs at the regulatory control level. Tuning of this override controller is defined by the proposed tuning of the filter $F_r(s)$, taking into account that tuning procedures for PID and DTC-PID controllers are defined in the available literature.

Set-point filters are mostly applied to avoid abrupt changes of controlled and control variables. However, there are examples when a fast reaction on the set-point obtained from higher control levels is required and constraints on controlled and control variables allow such implementation. In these cases, a desired set-point following performance can be obtained if the procedure proposed to design and tune filter $F_r(s)$ is applied to define set-point filters for controllers on the regulatory control level.

Finally, in cascade control systems, a limiter is frequently applied on the inner loop set-point to keep the secondary output below the safety limit. In this case, the procedure proposed to design and tune filter $F_r(s)$ can be directly applied to avoid the overshoot. This is important to avoid action of safety devices, very costly to plant operation.

ACKNOWLEDGMENT

A.I. Ribić gratefully acknowledges the financial support from Serbian Ministry of Science and Technology (Project TR33022 and TR35003).

REFERENCES

[1] F.G. Shinskey, J.P. Shunta, J.E. Jamison, and A. Rohr, "Selective, Override, and Limit Controls", in Process Control and Optimization, 4th ed., vol. II, Béla Lipták Ed, CRC Press, Boca Raton, 2006, pp.336-344.

[2] A.H. Glattfelder, W. Schaufelberger, "Stability of discrete override and cascade-limiter single-loop control systems", IEEE Trans. on Automatic Control, pp. 532-540, 1998.

[3] A.I. Ribić, "Development of a new industrial controller: structure, models, tuning and implementation on thermal power plants", Ph.D.Thesis (in Serbian), University of Belgrade, Faculty of Electrical Engineering, Serbia, 2010.

[4] A.I. Ribić, M.R. Mataušek, “A dead-time compensating PID controller structure and robust tuning”, Journal of Process Control pp.1340-1349, 2012.

[5] L. Desborough, R. Miller, “Increasing customer value of industrial control performance monitoring-Honeywell’s experience”, 6th. Int. Conf. on Chemical Process Control, 2002, AIChE Symp. Series Number 326, vol. 98.

[6] M. Kano, M. Ogawa, 2010. “The state of the art in chemical process control in Japan: Good practice and questionnaire survey”, Journal of Process Control pp 969-982, 2010.

Rational approximations to design controllers for unstable processes, including dead-time

Marko Bošković, student
Faculty of Electrical Engineering
University of East Sarajevo
East Sarajevo, Bosnia and Herzegovina
boskovic.marko.etf@gmail.com

Tomislav B. Šekara
Faculty of Electrical Engineering
University of Belgrade
Belgrade, Serbia
tomi@etf.rs

Milan R. Rapaić
Boris Jakovljević
Faculty of Technical Sciences,
University of Novi Sad,
Novi Sad, Serbia
rapaja@uns.ac.rs
bjakov@uns.ac.rs

***Abstract* – This paper addresses the problem of designing complex controllers for unstable industrial processes with transport delay under constraints on robustness and performance. The solution to the control design problem is obtained in a non-rational form which is rationalized by using various methods. The paper also presents a comparative analysis of different approximation techniques. By means of numerical simulations it has been shown that the proposed methods lead to adequate performance and robustness indices.**

Keywords – Rational approximations, Complex controller, Robustness, Frequency domain.

I. INTRODUCTION

The paper considers the problem of designing complex controllers of unstable industrial processes including transport delays, observing the restrictions on performance and robustness. When designing complex controllers for unstable processes with transport delays, a time-delay may also appear within the transfer function of the controller $C(s)$. Thus, after the design phase, the controller transfer function is not rational and one often encounters internal instability. This internal instability in the controller appears as one or more unstable dipoles, i.e. pairs of poles and zeros at the same points at the right half-plane of the s-plane. Elimination of the intrnal instability by means of suitable rational approximations of the controller transfer function is an intrinsic step in solving the control problem under consideration, and thus represents a key motivator for the present paper. Several methods of rational approximation of $C(s)$ are considered in the present paper.

It is well known that nearly 94% of feedbacks in industry is realized by PI/PID controllers [1], while this percentage is over 97% in petrochemical industry [2, 3]. Owing to this high significance of PI/PID, the efficient and simple procedures for tuning of parameters of industrial controllers have been developed [4, 5], as well as the optimization procedures [6-19] for designing PI/PID with minimum *IAE*, observing the restrictions on robustness, which meets the criterion presented in [22].

In addition to the previously mentioned methods, there are methods for designing PID controllers derived from IMC controller [23-25]. For IMC method of designing controllers there is one adjustable parameter λ which, for a narrow class of processes, directly influences the time constant of the closed loop system. Response to a Heaviside disturbance of the process controlled by a controller obtained by IMC method is dependent on the dominant dynamics of the process. E.g. if the process has dominant oscillatory dynamics, responces to any disturbance will be oscillatory.

The problem of control of complex processes (multiple instabilities, multiple astatisms, dominant time delay) can not be solved adequately by applying PID controller, the basic reason for developing methods of designing complex controllers. For the purpose of accomplishing adequate indices of robustness and performance for a wide class of stable and unstable processes new methods [26-29] have been developed for determination of complex controllers based on modified IMC structure. However, the rules of design of complex controllers by applying these methods have not been defined for the general form of process transfer function $G_p(s)$, but they are defined for specific classes of processes $G_p(s)$ [26-29]. Controller $C(s)$ and its rational approximation defined in [31] is designed for the general form of the process transfer function $G_p(s)=H(s)\exp(-\tau s)/Q(s)$ under restrictions on robustness and sensitivity to measurment noise.

Adjustable parameters of a complex controller $C(s)$ are time constant λ and factor of relative damping ζ of dominant poles of the process in closed loop with controller $C(s)$, as in [6]. Adjusting of parameter ζ allows achieving compromise between indices of robustness and performance, which is not possible for complex controllers designed by IMC [23-25] or by modified IMC [26-29].

Through a series of simulations of a wide class of industrial processes a comparison of different methods of rational approximation of $C(s)$ in order to achieve an adequate index performance/robustness and internal stability of the controller has been obtained.

II. COMPARATIVE ANALYSIS OF SEVERAL METHODS OF RATIONAL APPROXIMATION OF INTERNALLY UNSTABLE CONTROLLERS

The control structure with $C(s)$ controller is presented in Fig. 1. $G_{ff}(s)$ describes the feed forward from the set point y_{sp} to control signal u and will not be considered here. For a wide class of transfer functions of industrial processes is $G_p(s) = H(s)e^{-\tau s} / Q(s)$, where $Q(s)$ and $H(s)$ are polynomials

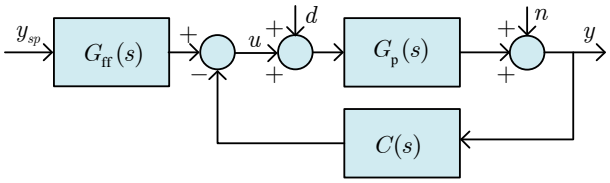


Figure 1. Control structure with controller C(s)

of the order $\deg Q(s) = n \geq \deg H(s) = 0$ and $H(0)\neq 0$. Complementary sensitivity function of the controlled process $G_p(s)$ of Fig. 1 is given by relation $T_p(s) = L(s)/(1+L(s))$ [32], the function of feedback transfer being of the form $L(s) = C(s)G_p(s)$. Let the desired complementary sensitivity function $T(s)$ is given by

$$T_p = T(s) = \frac{N(s)e^{-\tau s}}{P(s)}, \tag{1}$$

where: $N(s) = 1 + \sum_{j=1}^{n} \eta_j s^j$, $P(s) = (\lambda^2 s^2 + 2\zeta\lambda s + 1)^n$, $\zeta \in O(1)$, $n \in \mathbb{N}$ and adjustable parameteris $\lambda > 0$, $\eta_j \in \mathbb{R}$, $j = \overline{1,n}$, determined on the basis of desired peerformance of the closed loop system. From relation (1) controller $C(s)$ for process transfer function $G_p(s)$ for maximum attenuation of disturbance d or n is defined as

$$C(s) = \frac{1}{G_p(s)} \frac{T(s)}{1-T(s)} = \frac{1}{H(s)} \frac{N(s)Q(s)}{F(s)}, \tag{2}$$

where $F(s) = P(s) - e^{-\tau s}N(s)$, [30].

In general, parameters $\overline{\eta_1, \eta_n}$ are determined to achieve cancelling of poles of process $G_p(s)$ and zeros of function $F(s)$, [29], where for unstable processes an internal instability arises in complex controller (2).

Free parameters of complex controller (2) are time constant $\lambda > 0$ and factor of relative damping $\zeta > 0$ of the closed loop system, as in [6,30]. Damping factor introduced in the complex controller design plays a significant role in accomplishing a compromise between the performance and robustness indeces. It should be mentioned that parameter ζ affects sensitivity to the measurement noise at high frequencies M_n,

$$M_n = \lim_{\omega \to \infty} \left| \frac{C(i\omega)}{1 + C(i\omega)G_p(i\omega)} \right|, \tag{3}$$

In order to accomplish a compromise between the desired performance IAE and $M_s = \max_{\omega} |1/(1+L(i\omega))|$, time constant λ should fullfil condition

$$\max_{\omega,\lambda} \left| 1/(1 + C(i\omega)G_p(i\omega)) \right| = M_s . \tag{4}$$

Given ζ and M_s (4), time constant λ is determined by solving two nonlinear algebraic equations

$$\left| 1 + C(i\omega)G_p(i\omega) \right|^2 - 1/M_s^2 = 0 , \tag{5}$$

$$\partial\left(\left| 1 + C(i\omega)G_p(i\omega) \right|^2 \right)/\partial\omega = 0 , \tag{6}$$

as in [6, 30]. The initial choice of parameter ζ should be ζ=1 and parameter λ from the vicinity of the estimated transport delay. By determining time constant λ for different values of parameter ζ, one achieves compromise between values IAE, M_n, and $M_p = \max_{\omega} |L(i\omega)/(1+L(i\omega))|$, under condition that in the case of an unstable process $G_p(s)$ the unstable dipole in controller $C(s)$ is removed.

In order to remove the unstable dipole, if it exists in the controller, several methods found in the literature reduce to application of Padé approximation.

A. Padé approximation of controller C(s)

Padé approximation is one od the most frequent rational approximations met in the control systems and wider. It can be calculated as follows [26,31]

$$C(s) \approx \hat{C}^{LZG}_{N/N}(s) = \frac{\sum_{j=0}^{N} d_j s^j}{s\sum_{i=0}^{N-1} c_i s^i}, \tag{7}$$

where N is the user-specified controller order to achieve the desirable performance specification for the load disturbance rejection, and c_i and d_j are determined by the following two linear matrix equations.

$$\begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_N \end{bmatrix} = \begin{bmatrix} b_0 & 0 & 0 & \cdots & 0 \\ b_1 & b_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ b_N & b_{N-1} & b_{N-2} & \cdots & b_1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix}, \tag{8}$$

$$\begin{bmatrix} b_N & b_{N-1} & b_{N-2} & \cdots & b_2 \\ b_{N+1} & b_N & b_{N-1} & \cdots & b_3 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ b_{2N-2} & b_{2N-3} & b_{2N-4} & \cdots & b_N \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{N-1} \end{bmatrix} = -\begin{bmatrix} b_{N+1} \\ b_{N+2} \\ \vdots \\ b_{2N-1} \end{bmatrix}, \tag{9}$$

where $b_i = f^{(i)}(0)/i!, i = \overline{1,2N-1}$ are the Maclaurin coefficients of $f(s) = sC(s)$ and c_0 should be taken as $c_0 = \begin{cases} 1, & \text{if } c_i \geq 0 \\ -1, & \text{if } c_i < 0 \end{cases}$. In this way, under certain conditions on parameters λ, ζ, and N, the internal instability of controller $C(s)$ can be removed. However, the obtained controller $\hat{C}^{LZG}_{N/N}(s)$ is not always an adequate approximation of $C(s)$ thus index robustness/performace could be impaired. For this reason in [29, 30] other methods of rational approximation are met.

The method of rational approximation of $C(s)$ applied in [29] in essence also makes use of Padé approximation in the

following way. The obtained controller $C(s)$ from (2) is first decomposed in the form

$$C(s)=\frac{N(s)Q^{-}(s)}{H(s)P(s)}\frac{Q^{+}(s)}{\left(1-\dfrac{N(s)e^{-\tau s}}{P(s)}\right)}, \qquad (10)$$

where: $Q^{+}(s)$ and $Q^{-}(s)$ are polynomials whose roots represent poles of process $G_p(s)$ in the right and left half planes of the s-plane respectively. Then, Padé approximation is applied only to a part of the preceeding expresion $D(s)=Q^{+}(s)/\left(1-\dfrac{N(s)e^{-\tau s}}{P(s)}\right)$, taking care that the obtained controller is causal, ie.

$$\hat{C}^{\mathrm{LG}}(s)=\frac{N(s)Q^{-}(s)}{H(s)P(s)}\hat{D}_{N/M}(s),\ \hat{D}_{N/M}(s)=\frac{\sum_{j=0}^{N}b_j s^j}{s\sum_{i=0}^{M}a_i s^i}. \qquad (11)$$

Similarly in [30], in order to avoid calculation of Padé approximation of complex functions, the use is made of the known Padé approximation of function $e^{-\tau s}$ in controller $C(s)$ followed by factorization and elimination of unstable dipoles to remove internal instability of the controller. Since Padé approximation of function $e^{-\tau s}$ in the vicinity of $s=0$ is given in the form

$$e^{-\tau s}\approx\frac{B_M(s)}{A_N(s)}=\frac{\sum_{k=0}^{M}\dfrac{(M+N-k)!M!}{(M+N)!k!(M-k)!}(-\tau s)^k}{\sum_{k=0}^{N}\dfrac{(M+N-k)!N!}{(M+N)!k!(N-k)!}(\tau s)^k}, \qquad (12)$$

and on the basis of relation (2) it follows

$$\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{L/L}(s)=\underbrace{\left(\frac{1}{H(s)}\frac{A_N(s)N(s)Q(s)}{(P(s)A_N(s)-N(s)B_M(s))}\right)}_{\text{factorization and elimination of unstable dipoles}}, \qquad (13)$$

therefore in the obtained controller $\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{L/L}(s)$ there are no unstabile dipoles.

B. Simulation analysis

Comparison of the proposed methods for removal of internal instability in controller $C(s)$, (2), will be analyzed through four representative typical dynamic characteristics of unstable industrial processes including transport delay:

$$G_{\mathrm{p1}}=\frac{4e^{-2s}}{4s-1},\qquad G_{\mathrm{p2}}=\frac{e^{-1.2s}}{(s-1)(0.5s+1)},$$

$$G_{\mathrm{p3}}=\frac{e^{-0.2s}}{s(s-1)},\qquad G_{\mathrm{p4}}=\frac{2e^{-0.3s}}{(3s-1)(s-1)}.$$

The obtained controllers are:

Process G_{p1}.

$$\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{4/4}(s)=\frac{2.0488(s^2+3.6778s+6.4594)(s+2.3222)(s+0.0711)}{(s^2+0.3444s+6.8416)(s+8.7177)},\ \hat{C}^{\mathrm{LZG}}_{4/4}(s)=\frac{28.7970(s+2.7760)(s+0.0711)(s^2+4.5789s+9.0689)}{s(s+220.299)(s^2+0.1638s+6.3867)}$$

$$\hat{C}^{\mathrm{LG}}_{6/6}(s)=\frac{9.6209(s+2.6822)(s+0.3818)(s+0.3815)(s^2+4.3998s+8.6936)(s+0.0711)}{s(s+0.3817)^2(s+68.06)(s^2+0.1873s+6.3972)}$$

Process G_{p2}.

$$\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{4/4}(s)=\frac{6.4835(s+0.0154)(s^2+5s+8.3333)(s+2)}{s(s^2+0.0370s+12.6343)(s+7.7176)},\ \hat{C}^{\mathrm{LZG}}_{4/4}(s)=\frac{92.2837(s+1.9890)(s+0.0155)(s^2+7.2907s+17.0647)}{s(s+279.8)(s^2-0.4117s+10.099)}$$

$$\hat{C}^{\mathrm{LG}}_{7/7}(s)=\frac{242.8705(s^2+1.1839s+0.3505)(s^2+7.4912s+17.5525)(s+0.5807)(s+2)(s+0.0155)}{s(s+0.5882)^3(s+760.46719)(s^2-0.4209s+10.1131)}$$

Process G_{p3}.

$$\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{3/3}(s)=\frac{31.1548(s+10)(s^2+0.5542s+0.2037)}{s(s^2+17.3492s+127.1133)},\ \hat{C}^{\mathrm{LZG}}_{2/2}(s)=\frac{0.19065(2911.0054s^2+2309.2295s+1108.7320)}{s(s+190.6494)},\ \hat{C}^{\mathrm{LG}}_{N/N}(s)\ \text{none}$$

Process G_{p4}.

$$\hat{C}^{\check{\mathrm{S}}\mathrm{TG}}_{4/4}(s)=\frac{100.7255(s^2+0.4604s+0.2583)(s^2+20s+133.3333)}{s(s+23.0066)(s^2+6.4890s+144.1208)},\ \hat{C}^{\mathrm{LZG}}_{3/3}(s)=\frac{54.8760(s^2+0.4762s+0.2695)(s+12.8684)}{s(s^2+2.4488s+172.6147)}$$

$$\hat{C}^{\mathrm{LG}}_{7/7}(s)=\frac{49.2861(s^2+0.4604s+0.2583)(s+14.6555)(s^2+4.4854s+5.170)(s^2+3.6172s+3.2987)}{s(s+2.0408)^4(s^2+1.9056s+175.3277)}$$

On the basis of Table 1 and transfer functions of the obtained controllers, it can be concluded that controllers $\hat{C}^{\mathrm{LG}}(s)$ are of higher orders compared to those obtained by other methods, all having the same M_s. In addition, in the third example by applying LG method it was not possible to determine the controller for the given M_s. Controllers of type $\hat{C}^{\mathrm{LG}}(s)$ and $\hat{C}^{\mathrm{LZG}}(s)$ obtained for the second process contain double unstable pole, but in control systems it is not recommendable that the controller itself is unstable since this impairs robustness of the control loop. Controller of type

$\hat{C}^{\check{S}TG}(s)$ gives better results for processes having dominant transport delays compared to other two controller types. In design of the controller it is possible to include parameter ζ, as in [30], in order to achieve better indices of peformance and robustness.

TABLE I. THE RESULTS OBTAINED BY THE PROPOSED METHODS OF RATIONAL APPROXIMATION OF CONTROLLER C(S) FOR THE SAME VALUE OF M_s AND ζ=1.

Process	Method	λ	M_n	IAE	M_s	M_p
$G_{p1}(s)$	$\hat{C}^{\check{S}TG}_{4/4}(s)$	2.62	2.05	27.3	2.9	2.7
	$\hat{C}^{LZG}_{4/4}(s)$	2.62	29.0	27.3	2.9	2.7
	$\hat{C}^{LG}_{6/6}(s)$	2.62	9.62	27.3	2.9	2.7
$G_{p2}(s)$	$\hat{C}^{\check{S}TG}_{4/4}(s)$	1.71	6.48	58.1	14	14.6
	$\hat{C}^{LZG}_{4/4}(s)$	1.71	92.3	58.1	14	14.6
	$\hat{C}^{LG}_{7/7}(s)$	1.71	243	58.1	14	14.6
$G_{p3}(s)$	$\hat{C}^{\check{S}TG}_{3/3}(s)$	0.63	31.1	2.00	2.8	3
	$\hat{C}^{LZG}_{2/2}(s)$	0.47	555	0.90	2.8	3
	$\hat{C}^{LG}(s)$	-	-	-	-	-
$G_{p4}(s)$	$\hat{C}^{\check{S}TG}_{4/4}(s)$	0.49	100	0.95	3.8	3.9
	$\hat{C}^{LZG}_{3/3}(s)$	0.48	54.8	0.91	3.8	3.9
	$\hat{C}^{LG}_{7/7}(s)$	0.49	49.3	0.96	3.8	3.9

III. CONCLUSIONS

Design of controllers of unstable industrial processes including transport delays with restrictions on performance and robustness is of exceptional significance from the point of view of industry. The problem of control of complex processes (multiple instabilties, multiple astatisms, dominant time delay) can not be adequatly solved by using PID controllers, which is the basic reason for development of the methods for design of complex controllers. The conditions for a complex controller are that it is stable, of relativly lower order, and of adequate structure for practical realization. In this work three methods of rational approximation in the design of complex controllers have ben analyzed. The presented comparative analysis and simulation gave the expected results.

ACKNOWLEDGMENT

The authors would like to acknowledge financial support of the Serbian Ministry of Science and Education, under grants TR33020 (T.B.Š), TR32018 and TR33013 (M.R.R).

REFERENCES

[1] S. Yamamoto and I. Hashimoto, "Present status and future needs: the view from Japanese industry", In Arkun and Ray, Eds., *Chemical Process Control-CPCIV. Proc. 4th Inter. Conf. on Chemical Process Control,* TX, 1991.

[2] L. Desbourough, R. Miller, "Increasing customer value of industrial control performance monitoring—Honeywell's experience", in: *Sixth International Conference on Chemical Process Control, AIChE Symposium* Series Number 326, Vol.. 98, 2002.

[3] K.J. Åström, T. Hägglund, "Revisiting the Ziegler-Nichols step response method for PID control", *Journal of Process Control*, vol. 14, pp. 635-650, September 2004.

[4] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning", *Journal of Process Control*, Vol. 13, pp. 291–309, 2003.

[5] M. Shamsuzzoha, M. Lee, "IMC-PID controller design for improved disturbance rejection of time-delayed processes", *Industrial & Engineering Chemistry Research*, vol. 46, no.7, pp. 2077-2091, 2007.

[6] M.R. Mataušek, T.B. Šekara, "PID controller frequency-domain tuning for stable, integrating and unstable processes, including dead-time", *J. Process Control* vol. 21, pp. 17-27, 2011.

[7] T.B. Šekara, M.R. Mataušek, "Classification of dynamic processes and PID controller tuning in a parameter plane", *J. Process Control* vol. 21, Issue 4, pp. 620-626, 2011.

[8] H. Panagopoulos, K.J. Åström and T. Hägglund, "Design of PID controllers based on constrained optimization", *IEE Proceedings-Control Theory and Applications*, vol. 149, pp. 32-40 January 2002.

[9] A. Wallén, K.J. Åström, and T. Hägglund, "Loop-shaping design of PID controllers with constant T_i/T_d ratio", *Asian Journal of Control,* vol. 4, pp. 403-409, December 2002.

[10] C. Hwang and C-Y. Hsiao, "Solution of non-convex optimization arising in PI/PID control design", *Automatica* vol. 38, pp. 1895-1904, November 2002.

[11] B. Kristiansson and B. Lennartson, "Evaluation and simple tuning of PID controllers with high-frequency robustness", *Journal of Process Control*, vol..16, pp. 91-102, February 2006.

[12] B. Kristiansson and B. Lennartson, "Robust tuning of PI and PID controllers: using derivative action despite sensor noise" *IEEE Control Systems Magazine*, pp. 55-69, February 2006.

[13] A.J. Isaksson and S.F. Graebe, "Derivative filter is an integral part of PID design", *IEE Proceedings-Control Theory and Applications*, vol. 149 pp. 41-45, January 2002.

[14] A. Karimi, M. Kunze and R. Longchamp, "Robust controller design by linear programming with application to a double-axis positioning system", *Control Engineering Practice*, vol. 15, pp. 197-208, February 2007.

[15] T. B. Šekara, M. R. Mataušek, "A simple effective method to obtain a well-tuned PID controller", Proceedings of the 51st, Conference on. ETRAN, AU3.1, 2007 (in Serbian).

[16] T.B. Šekara, M.R. Mataušek, "Optimal tuning of a PI/PID controller for processes defined by a rational transfer function", INFOTEH Vol. 6, Paper A-2, p. 6-9, Jahorina, March 2007 (in Serbian).

[17] T.B. Šekara and M.R. Mataušek, "Optimization of PID controller based on maximization of the proportional gain under constraints on robustness and sensitivity to measurement noise ", *IEEE Trans. Automatic Control*, vol. 54, no.1, pp.184-189, Jan. 2009.

[18] T.B. Šekara and M.R. Mataušek, "Revisiting the Ziegler-Nichols process dynamics characterization", *J. Process Control* vol. 20, pp. 360-363, 2010.

[19] T. B. Šekara, M. R. Mataušek, "A four-parameter optimization of a PID controller", Proceedings of 52. Conf. ETRAN, Vol. 1, Palic, Junne 2008 (in Serbian).

[20] T.B. Šekara, M.R. Mataušek, "Optimal tuning of a PID controller in frequency domain,", INFOTEH , Paper A-6, p. 24-27, Bosnia and Herzegovina, Jahorina, March 2009 (in Serbian).

[21] T.B. Šekara, M.B. Trifunovic, Optimal tuning of a PID controller having a differential compensator connected in series in frequency domaine, Proceedings of INDEL, pp. 258-261, Banja Luka, 4-6 November 2010 (in Serbian).

[22] F.G. Shinskey, "How good are our controllers in absolute performance and robustness?", *Measurement and Control*, vol. 23, pp. 114-121, May 1990.

[23] Chien, I. L."IMC-PID controller design-an extension" *IFAC Proceedings Series*, vol. 6 (pp. 147-152), 1988.

[24] Fruehauf, P. S., Chien, I. L., & Lauritsen, M. D. "Simplified IMC-PID tuning rules." *ISA Transactions*, 33(1), 43-59, 1994.

[25] Lee, Y., Lee, M., Park, S., & Brosilow, C. “PID controller tuning for desired closed-loop responses for SI/SO system. ” A.I.Ch.E. Journal, 44(1), 106-115, 1998.

[26] Liu T, Zhang W, Gu D. Analytical design of two-degree-of-freedom control scheme for open-loop unstable process with time delay. *J Process Control* 2005; 15:559-72.

[27] M. Shamsuzzoha, M. Lee, “Analytical design of PID controller cascaded with a lead-lag filter for time-delay processes”, Korean J. Chem. Eng., vol. 26, pp. 622-630, 2009.

[28] M. Shamsuzzoha, M. Lee, “Enhanced disturbance rejection for open-loop unstable process with time delay” *ISA Transactions*, vol. 48, pp. 237-244, 2009.

[29] M. Liu, F. Gao, “Enhanced IMC design of load disturbance rejection for integrating and unstable processes with slow dynamics” *ISA Transactions*, vol. 50, pp. 239-248, 2011.

[30] T.B. Šekara, M. B. Trifunović, V. Govedarica, “ Frequency Domain Design of a Complex Controller under Constraints on Robustness and Sensitivity to Measurement Noise“, *Electronics*, vol. 15, pp. 40-44, 2011.

[31] A. George, J. Baker, G.M. Peter, Padé Approximants: Part I.Basic Theory, Addison Wesley, New York, 1981.

[32] K.J. Åström, T. Hägglund, Advanced PID control, ISA-The instrumentation, Systems, and Automation Society, 2006.

Adaptive Torque Control for Sensorless Induction Motor Drives in Wide-Speed Range

Aleksandar Ž. Rakić
University of Belgrade, School of Electrical Engineering
Belgrade, Serbia
rakic@etf.rs

Petar R. Matić
University of Banja Luka, School of Electrical Engineering
Banja Luka, Republic of Srpska, Bosnia and Herzegovina
pero@etfbl.net

Abstract—**In this paper, control structure is proposed to ensure desired performance of sensorless induction motor (IM) drives in both base speed range and field-weakening. Appropriate nonlinear IM model is utilized for derivation of adaptive slip manipulation based torque control law. In the base speed range, proposed solution reduces to indirect field-oriented control (IFOC), while in the field weakening it becomes voltage angle control with full dc bus utilization. Proposed solution is verified by means of simulation.**

Keywords - induction motor; sensorless drive; torque control.

I. INTRODUCTION

The high performance IM drives are mainly controlled by field-oriented control schemes [1], where direct torque control (DTC) make immediate use of stator voltage vector to control the torque, while IFOC-type torque control rely on embedded current control loops. During the last decade high-speed IM drives are being in the focus of research [2]-[13] for their efficiency, small cost and the ability to operate in wide speed range without the mechanical transmission gear. The operation of IM at speeds higher than the nominal one is enabled by the field-weakening (reduction of the rotor flux), where maximum torque capability can be obtained only by the full utilization of available inverter voltage, i.e. stator voltage vector is required to be in saturation with its amplitude set to maximal.

Though IFOC obtains superior dynamic performance in base-speed region, proper operation of its current loops in the field-weakening demands absence of the stator voltage saturation. On the other hand, authors of this paper proposed DTC-type voltage angle torque control (VATC) [14]-[17], which is intended for the full DC bus utilization and high performance in the field weakening, while its base-speed region performance is recognized to be inferior to IFOC. In this paper, adaptive torque control solution is proposed as a proper utilization of both IFOC and VATC in the regions of their superiority; IFOC in the base-speed region, and VATC in the field-weakening.

The rest of the paper is organized as follows. In Section II appropriate nonlinear state-space model of IM is presented. Outline of torque control derivation is presented in Section III, along with the structural block diagram of the overall control solution. Verification by simulation in few representative torque demand scenarios is given in Sect. IV.

II. NONLINEAR IM MODEL

Assuming mechanical transients are much slower than electrical ones, the state-space model of the IM, in normalized (per-unit [p.u]) values, is given as follows:

$$\frac{di_d}{dt} = \omega_b \left[-\frac{i_d}{T_\sigma} + \omega_{dq} i_q + \frac{k_r \Psi_D}{T_r L_\sigma} + \omega_r \frac{k_r \Psi_Q}{T_r L_\sigma} + \frac{u_d}{L_\sigma} \right], \quad (1)$$

$$\frac{di_q}{dt} = \omega_b \left[-\omega_{dq} i_d - \frac{i_q}{T_\sigma} - \omega_r \frac{k_r \Psi_D}{L_\sigma} + \frac{k_r \Psi_Q}{T_r L_\sigma} + \frac{u_q}{L_\sigma} \right], \quad (2)$$

$$\frac{d\Psi_D}{dt} = \omega_b \left[\frac{L_m}{T_r} i_d - \frac{1}{T_r} \Psi_D + \omega_{sl} \Psi_Q \right], \quad (3)$$

$$\frac{d\Psi_Q}{dt} = \omega_b \left[\frac{L_m}{T_r} i_q - \omega_{sl} \Psi_D - \frac{1}{T_r} \Psi_Q \right], \quad (4)$$

$$t_e = \frac{L_m}{L_r} \left(\Psi_D i_q - \Psi_Q i_d \right), \quad (5)$$

where the state vector consists of stator currents and rotor fluxes, motor torque t_e is primary output variable, and immediate controlling variables are stator voltage components u_d and u_q.

In the model (1) – (5), ω_b is base speed and all other variables and parameters are normalized (in [p.u.]): ω_{dq} is synchronous frequency, ω_{sl} is motor slip, ω_r is rotor angular velocity, R_s and R_r are stator and rotor resistance, L_s and L_r are stator and rotor self–inductances, L_m is mutual inductance, $k_r = L_m/L_r$ is rotor coupling coefficient, $T_\sigma = L_\sigma / R_\sigma$ and $T_r = L_r / R_r$ are stator and rotor transient time constants in [p.u], $\sigma = 1 - M^2 / L_s L_r$ is leakage coefficient and l_s is stator inductance in [p.u].

III. PROPOSED TORQUE CONTROL

Performing Laplace transform on (3) and (4), solving the system of equations for $\Psi_D(i_d, i_q)$ and $\Psi_Q(i_d, i_q)$, and introducing the obtained solutions in (5), result in:

$$t_e = \frac{T_r L_m^2}{L_r} \frac{i_d^2 + i_q^2}{T_r^2 p^2 + 2T_r p + 1 + \omega_{sl}^2 T_r^2} \omega_{sl}, \quad (6)$$

where p stands for the complex variable of the Laplace transform.

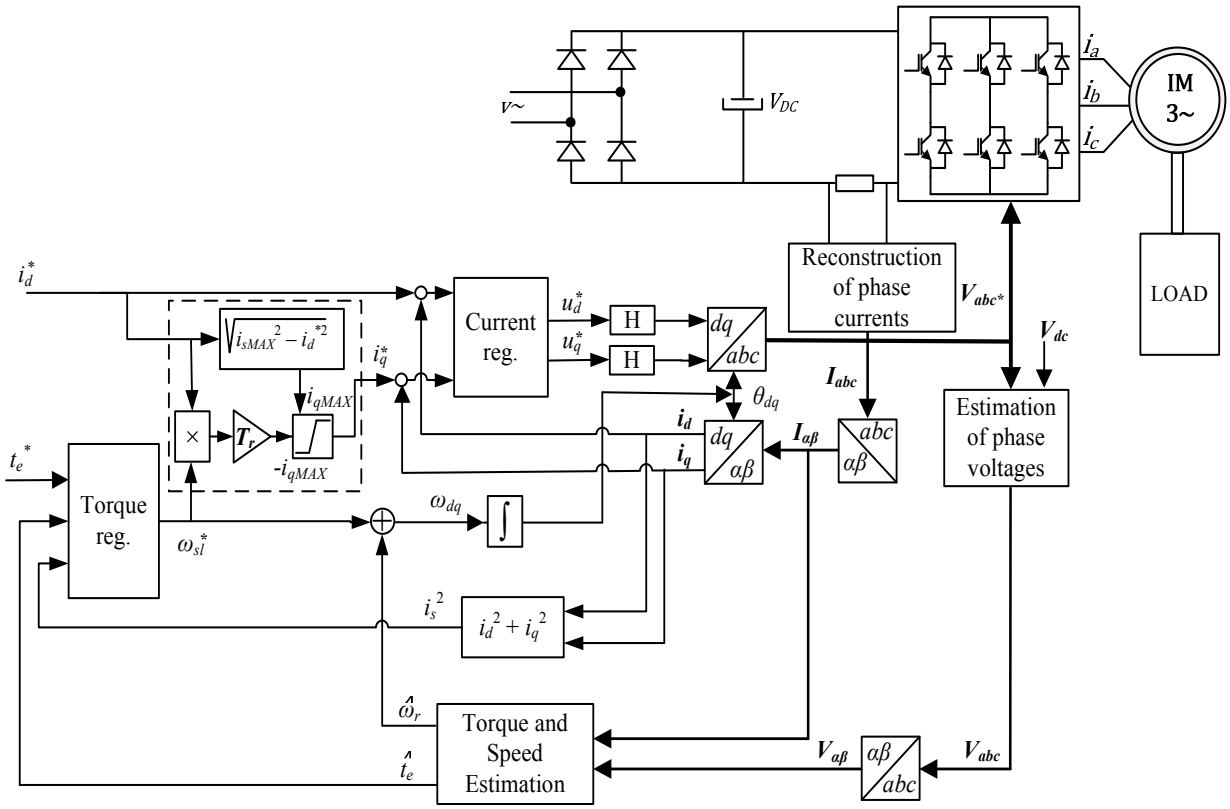


Figure 1. Structural block diagram of the proposed solution

In the vicinity of the operating regime $(\omega_{sl}, i_s) = (\omega_{sl}^0, i_s^0)$, where $i_s = \sqrt{i_d^2 + i_q^2}$ is the stator current modulus, transfer from the control variable ω_{sl} to the output t_e is adopted as a low-frequency approximation of (6):

$$G_0 = \frac{L_m^2}{R_r} \frac{\left(i_s^0\right)^2}{1 + \left(\omega_{sl}^0 T_r\right)^2}. \tag{7}$$

For the plant (7) and adopted bandwidth

$$\omega_0 = \frac{\omega_b}{T_r}, \tag{8}$$

gain-scheduling integral torque controller is proposed:

$$K(p) = \frac{\omega_0}{G_0} \frac{1}{p}. \tag{9}$$

In order to prevent excessive control actions which could lead to torque breakdown, limit of $\pm 1/(\sigma T_r)$ is set to integration in (9).

Block diagram of proposed torque control method is shown on Fig. 1. The basic idea is to utilize unique torque regulator to perform motor-slip command ω_{sl}^* calculation for both base-speed region (BSR) and field-weakening (FW), since the proposed control law (9) uses model (7), which is valid in both drive operation regimes. When the drive operates in BSR, reference currents i_d^* and i_q^* are fed to current regulators and proposed solution reduces to effective IFOC algorithm. On the other hand, when output voltage commands u_d^* and u_q^* reach the limit $U_{s\max}$ of maximal available inverter voltage amplitude ($\sqrt{u_d^{*2} + u_q^{*2}} = U_{s\max}$), the output voltage commands u_d^* and u_q^* of the current regulators are held by holding circuits H and drive enters the FW regime. Inverter voltage is fully utilized and the torque control is effectively performed by the stator voltage angle control, i.e. the torque regulator adjusts the angle of the (maximal amplitude) stator voltage vector only by motor-slip manipulation.

Though voltage limit is explicitly addressed by the proposed control solution, the inverter current limit $I_{s\max}$ violations are possible so far. One way to impose the current limitation in torque control is to calculate maximal reachable motor torque in the existing operating regime for the current engagement i_s limited to $I_{s\max}$. Utilizing model (7), maximal reachable motor torque

$$t_{eLIM}^* = \frac{L_m^2 I_{s\max}^2}{R_r} \frac{\omega_{sl}^0}{1 + \left(\omega_{sl}^0 T_r\right)^2} \tag{10}$$

can be used as a dynamic limitation of the torque command t_e^* in order to effectively impose the limitation of the motor current.

IV. Simulation Results

Simulations were conducted in Matlab/Simulink in order to verify expected performance. Proposed torque control

solution is tested through scenarios of torque command t_e^* application in different IM operating regimes, i.e. BSR and FW, as well as in transition between the aforementioned regimes. Control system is tested without any outer speed or position control loop and torque commands are chosen to be in the form of sequenced step changes. In that way torque control solution is exposed to worst case demands and obtained performance is expected only to be more favorable in the real-time exploitation than in the tested cases (outer control loops would set less demanding t_e^* signal shapes than the adopted step changes). Parameters of tested motor are given in Appendix.

The first test scenario is the sequential bidirectional application of maximal amplitude torque commands. The timing of step changes is chosen to lead the IM from BSR to FW and vice versa several times and in both rotation directions. The responses of all relevant variables are given in Fig. 2. Within first 2 seconds motor is brought from BSR into the FW and quickly returned to BSR. At t = 2sec, maximal positive torque command is applied and it drives motor speed again into the FW and up to 2 p.u. speed in forward direction. At t = 5 sec, maximal negative torque command is applied and it first drives motor back to BSR in forward direction and then, at $t \approx 6.8$ sec, reverses the direction of rotation and forces the motor to enter FW in reverse direction and build the speed up to approx. –2 p.u.

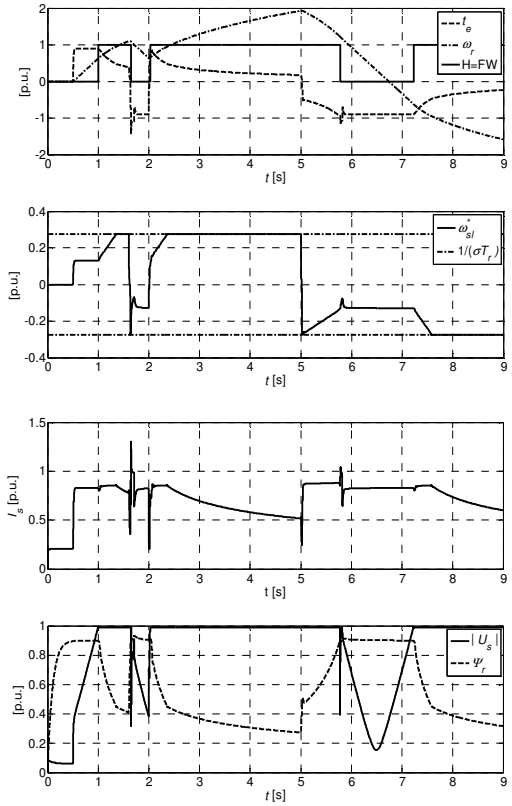


Figure 2. Test scenario with multiple bidirectional transitions to/from field weakening

Voltage saturation in FW results in automatically adjusted (decreased) levels of the rotor flux and obtainable torque, but it influences neither stability nor the performance of the torque control loop. The stator current amplitude is efficiently kept below the $I_{s\,\max}$ = 1 p.u. at all times, except at the FW→BSR transients. The reason of this current limit breach is the initial response of the reactivated IFOC current regulation. The effect is negligible, since $I_{s\,\max}$ stands for the steady current load limitation and the short-term current overloads are permitted (within the voltage inverter current margin).

The second test scenario is the sequential bidirectional application of 50% amplitude torque commands which keeps the IM in BSR, but drives it in both rotation directions. Wave-forms of the relevant variables are given in Fig. 3. Since both current and voltage are within limits, rotor flux is at its nominal value and the torque command is always fully attained.

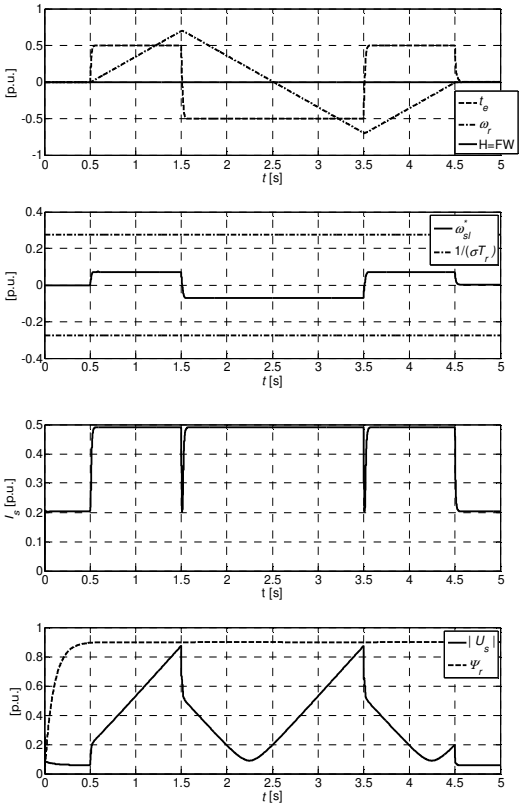


Figure 3. Test scenario for base speed performance evaluation.

Reversing of the drive rotation direction happens at t = 2.5 sec. No issues around zero speed were detected in torque or flux, mainly because inverter nonlinearities are neglected in the simulation model. Proposed torque control is primarily intended for high speed mode of operation in order to fully utilize available inverter voltage. However, operation at low speeds and reversing of the real drive is also possible, but with the speed estimation scheme suitable for low speeds.

The third test scenario is the sequential unidirectional application of maximal amplitude torque commands, which leads the drive deep into the FW in forward rotation direction. The responses of all relevant variables are given in Fig. 4. As in the case of the first scenario, voltage limit is reached in the FW region and rotor flux and torque producing capability is decreased. However, the current limit is not being breached, since the current demand in FW only decreases with the increase of drive speed. The quality of all the transients is desired.

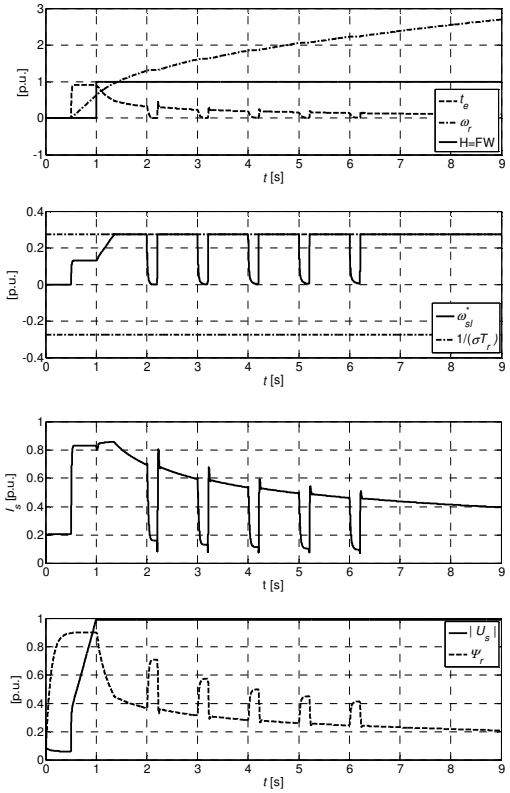


Figure 4. Test scenario for field weakening performance evaluation.

It can be observed in the FW drive operation that the rotor flux automatically regenerates (builds up) and the stator current drops to level lower than in BSR, when the torque demand is absent. This property may lead to conclusion that the proposed torque control represents optimal (or near optimal) solution in the sense of minimal power consumption with maximal torque producing capability, which is worth of the further investigation.

APPENDIX

Motor data: 750W, 195V, 70Hz, R_s = 10.8Ω, R_r = 5.673Ω, L_s = L_r = 0.552 H, L_m = 0.518 H.

REFERENCES

[1] D. Casadei, F. Profumo, G. Serra, A. Tani, "FOC and DTC: Two Viable Schemes for Induction Motor Torque Control", *IEEE Transactions on Power Electronics,* Vol. 17, No.5, September 2002, pp. 779-787.

[2] J. Pyrhonen, J. Nerg, P. Kurronen, U. Lauber, " High-Speed High-Output Solid-Rotor Induction-Motor Technology for Gas Compression", *IEEE Transactions on Industrial Electronics*, Vol. 57, No.1, January 2010, pp. 272-280.

[3] M. Centner, U. Schäfer, "Optimized Design of High-Speed Induction Motors in Respect of the Electrical Steel Grade", *IEEE Transactions on Industrial Electronics*, Vol. 57, No.1, January 2010, pp. 288-295.

[4] J. F. Gieras, J. Saari, "Performance Calculation for a High-Speed Solid-Rotor Induction Motor", *IEEE Transactions on Industrial Electronics*, Vol. 59, No.6, June 2012, pp. 2689 - 2700

[5] M. Mengoni, L. Yarri, A. Tani, G. Serra, D. Casadei, "Stator Flux Vector Control of Induction Motor Drive in the Field Weakening Region", *IEEE Transactions on Power Electronics*, Vol. 23, No. 2, March 2008, pp. 941-949.

[6] P. Y. Lin, Y. S. Lai, "Novel Voltage Trajectory Control for Field-Weakening Operation of Induction Motor Drives", *IEEE Transactions on Industry Applications,* Vol. 47, No. 4, January/February 2011, pp. 122-127.

[7] R. Sepulchre, T. Devos, F. Jadot, F. Malrait, "Antiwindup Design for Induction Motor Control in the Field Weakening Domain", *IEEE Transactions on Control System Technology*, Vol. 21, No.1, January 2013, pp. 52-66.

[8] F. Briz, A. Diez, M. W. Degner, R. D. Lorenz, "Current and Flux Regulation in Field –Weakening Operation", *IEEE Transactions on Industry Applications,* Vol. 37, No.1, January/February 2001, pp. 42-50.

[9] G. Gallegos-López, F. S. Gunawan, J. E. Walters, "Current Control of Induction Machines in the Field-Weakened Region", *IEEE Transactions on Industry Applications,* Vol. 43, No.4, July/August 2007, pp. 981-898.

[10] A. B. Jidin, N. R. B. Idris, A. H. B. M. Yatim, M. E. Elbuluk, T. Sutkino „A Wide-Speed High Torque Capability Utilizing Overmodulation Strategy in DTC of Induction Machines With Constant Switching Frequency Controller", *IEEE Transactions on Power Electronics*, Vol. 27, No. 5, May 2012, pp. 2566-2575.

[11] D. Casadei, G. Serra, A. Tani, L. Zarri, F. Profumo, "Performance Analysis of a Speed Sensorless Induction Motor Drive Based on a Constant Switching Frequency DTC Scheme", *IEEE Transactions on Industry Applications,* Vol. 39, No. 2, March/April 2003.

[12] M. Hajian, J. Soltani, G. A. Markadeh, S. Hosseinnia, "Adaptive Nonlinear Direct Torque Control of Sensorless IM Drives With Efficiency Optimization" *IEEE Transactions on Industrial Electronics,* Vol. 57, No.3, March 2010, pp. 975-985.

[13] J. Rodríguez, R. M. Kennel, J. R. Espinoza, M. Trincado, C. A. Silva, C. A. Rojas, "High-Performance Control Strategies for Electrical Drives: An Experimental Assessment", *IEEE Transactions on Industrial Electronics*, Vol. 59, No.2, February 2012, pp. 812-820.

[14] P. Matić, A. Rakić, and S. Vukosavić, "Induction Motor Torque Control in Field Weakening by Voltage Angle Control," in Proc. 14th Int. Power Electronics and Motion Control Conf. (EPE/PEMC), Ohrid, Macedonia, Sep. 6–8, 2010, pp. T4-108–T4-115.

[15] P. Matić, A. Rakić, and S. Vukosavić, "Direct Torque Control of Induction Motor in Field Weakening Based on Gain-Scheduling Approach," in Proc. 16th Int. Symp. Power Electronics (Ee 2011), Novi Sad, Serbia, Oct. 26–28, 2011.

[16] A. Rakić, P. Matić, "Robust Modeling and Reference Tracking Control of Voltage Angle Controlled Induction Motor in Field Weakening Regime," in Proc. INDEL Conf., Banja Luka, Republic of Srpska - Bosnia and Herzegovina, Nov. 4–6, 2010, pp. 262-267.

[17] A. Rakić, P. Matić, T. Petrović, "Robust Modeling and Gain-Scheduling Control of the Induction Motor in the Field Weakening Regime," in Proc. LV ETRAN Conf., Banja Vrućica-Teslić, Republic of Srpska - Bosnia and Herzegovina, Jun. 6-9, 2011, pp. AU4.5-1-4.

Session TO-8

RENEWABLE ENERGY AND ENERGY EFFICIENCY

Sophisticated research and development station for control of grid connected distributed energy sources

Bane Popadic, Boris Dumnic, Dragan Milicevic, Vladimir Katic, Corba Zoltan
Department of Power, Electronic, and Telecommunication Engineering
Chair of Power Electronics and Converters
Faculty of Technical Sciences
Novi Sad, Serbia
banep@uns.ac.rs

***Abstract*—A sophisticated research and development station for control of the grid connected distributed energy sources have been developed at the Faculty of Technical Sciences. This research station should allow researchers to investigate the influence of grid connected converter control on the grid conditions and vice versa. Additionally, the control of the grid connected converter during grid disturbances can be investigated. In addition to its R&D features, this station will present students with the possibilities to get high quality knowledge in the respected area. This paper aims to present the research potential of the station, while displaying some of the many versatile features gained by the introduction of the grid emulator.**

Keywords - distributed energy resource; grid connected converter; grid emulation; research and development station;

I. INTRODUCTION

In order to achieve sustainable development of the society in the future, it is necessary to address one of the key features that follow the development – energy demand. Energy market in the future, concerning the rather rapid increase in demand, will have to develop and adapt to more decentralized power system. The distinction between consumer and supplier is going to become less apparent, as renewable energy sources are integrated in the distribution grid. This allows for utility grid "clients" to alter its nature from consumer to supplier, and vice versa, depending on its current energy demand and generation capacity. In addition, with introduction of energy storage this alteration in nature can be postponed and made when energy prices are higher and it is more beneficiary for the grid connected subject to act as a supplier rather than consumer.

By the same token the development in the automotive industry is bringing forth ever so progressive inclusion of electrical vehicles in the transportation system, consequently leading to increase number of vehicles connected to the grid. The nature of the electrical vehicle and its possibility to run in V2G mode or G2V mode also provides the opportunity for it to behave as a source or as a load while connected to utility grid.

The basis of the sustainable development of the society therefore lies within the ideas and features proposed by the smart grid technology. However, with improvement of efficiency, reliability, economics, and sustainability of energy supply in mind, smart grid introduces imminent increase in number of grid connected converters. In such a number the influence of the converters on the grid power quality becomes a serious issue. Additionally, reduced power quality of the grid influences other converters, which in return influence the grid power quality even more. In that regard, developing adequate control for the grid connected converter is very important. These control algorithms, in addition to being able to actively participate in sustaining the grid power quality, have to be able to operate under different grid faults. Testing such an algorithm can prove to be a difficult task, especially due to the fact that grid fault appearances are relatively rare, and it is not rational to induce them potentially causing instability in the distribution network.

To accommodate the need for such a particular testing, Chair of Power Electronic and Converters, at Faculty of Technical Sciences, has put together a sophisticated research and development station for control of grid connected distributed energy sources. The station consists of cutting edge high-tech tools for development and testing of control algorithm of grid connected systems. The components allow for testing of these systems under various grid conditions in both operating regimes.

Another benefit of such a system is that it will serve as an educational tool for MSc and PhD students where they will get high quality knowledge in the area of grid connected distributed energy sources. Consequently, this will allow students to compete with other skilled individuals in the field [1]. The setup also includes a high level of versatility, allowing for a full control of the grid side converter, machine side converter, DC bus parameters, multi-phase machines etc.

This paper looks to present this sophisticated research and development station for control of grid connected distributed energy sources developed at the Faculty of Technical Sciences. Together with the description of basic elements of the research station, the paper will present possibilities for grid condition variations. There exist, but a few, experimental stations such as this one, that accompanies such level of versatility and modularity, and with its functionality it should position itself as a top tool for research and development of control algorithms.

II. RESEARCH AND DEVELOPMENT STATION DESCRIPTION

Basis of this research station is an advanced laboratory setup for control of electrical drives [1]. The laboratory setup presented in Ref. [1] is expanded and adapted to include a grid connected converter topology. In addition to the grid connected converter, grid emulator is integrated in the system to allow variable grid conditions.

As it is presented, the laboratory setup consists of the state of the art hardware in the field of electrical drives and control. Its modularity allowed for simple integration of a grid emulation system, with no noticeable change to the system. The basis of the system remains highly modular dSPACE hardware with modified industrial converters [2, 3, 4]. Versatility of the setup is increased by adding grid emulator to the grid side converter [5].

Fig. 1 shows the block diagram of the research station with included basic control elements. The actual outlook of the system is presented in Fig. 2, showing the grid connected topology and grid emulator separately. The focus of this system is to observe the behavior of grid side converter under different operating conditions.

A. Control Hardware

dSPACE is a highly modular, highly sophisticated control system hardware, which allows for a fast developing and testing of control algorithms. Major advantage of this control hardware is that it minimizes the development time and cost, while opening up a large number of possibilities of control optimization [2]. It consists of different control boards (modules) described in details in [1] and [6]. Control boards DS1006 (processor board), DS5101 (DWO board), DS2004 (A/D board), DS3001 (encoder board) and DS2201 (Multi I/O board) are used for the control of the complete research station. Fig. 3 shows all basic modules of a dSPACE hardware.

B. Power Converters

Power converters represent the basis for any modern renewable energy system, and thus is an integral part of the research station such as this one. Standard industrial converters are an obvious choice, since they are generally very robust, but they lack in the area of controllability especially for the laboratory testing. To overcome this, these converters are equipped with Interface and protection card (IPC), developed at Aalborg University [7].

The converter with the integrated IPC card is shown in Fig. 4. It is clearly intended to allow the user full operational control of the inverter switches, while maintaining its predominant quality of a highly reliable industrial converter. Another important quality of standard industrial drive is that it has relatively small packaging. This, together with the full controllability of the system allowed by the IPC, clearly shifts the edge towards using this solution for laboratory testing.

Some more details on the power converters can be found in the references [1], [3] and [4].

In contrast to the experimentation done in [1], with proper selectors and switch manipulation, one of the converters

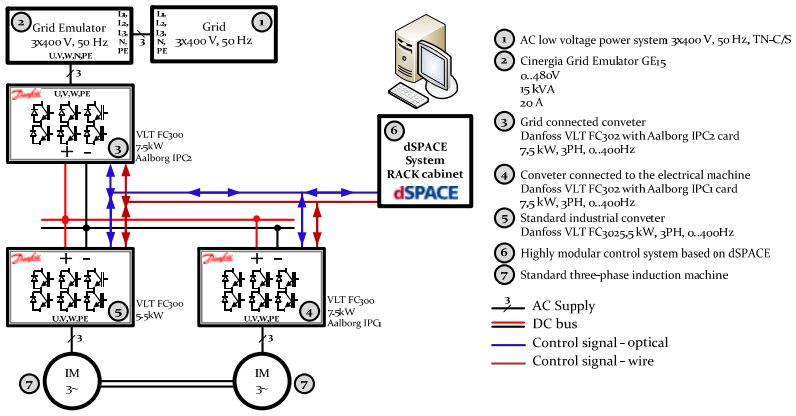


Figure 1. Block diagram of a research station

Figure 2. Physical overview of the system and the grid emulator

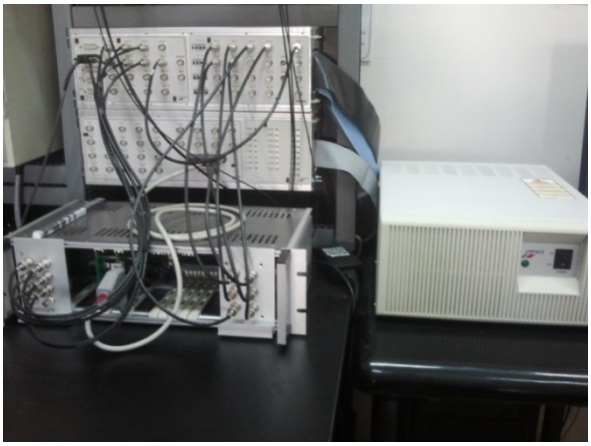


Figure 3. dSPACE based control hardware

Figure 4. Standard industrial converter with integrated IPC card

with an IPC card is connected to the grid at its active end, thus playing the role of a grid connected converter in renewable energy sources application.

C. Grid Emulator

To fully understand operation of grid connected converter it is necessary to observe its behavior not only under normal grid conditions, but also during faults, and different power quality conditions. To allow this kind of experimentation setup described in [1] needs to be complemented with a device capable of altering grid operating conditions. In addition, since the grid connected converter can act as a load, as well as the supplier to the electrical grid, this device should have possibility of bidirectional power flow. This is a highly valued feature, since it will allow for testing the grid connected converter both as an active rectifier, and as an inverter depending on the power flow. With that in mind, such a device would fit in well, with the modular versatile concept, this sophisticated and modern research station is trying to implement.

Grid emulator is a product of CINERGIA, a company which specializes in state of the art digital control of customized power electronics solutions. The device based on controlled power electronic converters capable of emulating the grid conditions. Modern design and powerful control features allow for a wide range of different operating regimes. Presented grid emulator is fully capable of a bidirectional power flow, which can be concluding by observing the block diagram in Fig. 5. It can be concluded that the basis for grid emulator is a Back-to-Back grid connected converter controlled by the separate DSP at the input and the output side. Control algorithms based on resonant control are employed for both the input and the output AC side. These control algorithms allow each harmonic to be controlled independently, consequently generating or suppressing it at a given set-point value. A large number of sensors are integrated for measurement of necessary parameters to achieve desired output parameters of the emulated grid. A 12 bit analogue to digital conversion with digital processing allows high resolution output up to 0.1 % paired with a high stability of the system [5]. Additional elements including filters, protection and switching equipment are present.

User has the possibility to operate the grid emulator using a PC software, digital I/O ports, analogue I/O ports, RS485 connector and associated protocol, CAN bus and standard RJ45 connector and MODBUS protocol. It is also possible to control the grid emulator using local touchscreen 3.2″ panel. The control of the grid emulator using local panel is very limited, without possibility to generate faults.

In additions to generating different type of grids, grid emulator is specially designed to also emulate their common faults and disturbances [5]. At the output side, grid emulator can produce three-phase grids with voltages ranging from 0 to 480 V, and frequencies from 40 to 400 Hz. Total distortion of these voltages (without including harmonic generation) is less than 0.5 %. Furthermore, voltage amplitude and phase referent value can be set independently for each phase, thus making it possible to create a non symmetrical three phase system (either by amplitude or by phase).

When non-ideal grid conditions are discussed, usually voltage harmonic distortion is considered. Grid emulator is fully capable of generating full range harmonics for orders up to 15^{th} for grids frequencies of 50 Hz (60 Hz). Any harmonic in that range can be set to any amplitude (with regard to maximum output voltage) independently for any phase. In addition to the harmonic spectrum that can be set for constant operating regime, harmonic sequence can be set as a grid fault. Harmonic sequence represents a run through, or cyclic repetition, of several different reference values of harmonic spectra. To test the fault ride through of the grid connected converter, grid emulator is capable of generating various types of voltage dips, over and under voltage disturbances, frequency variation and variable grid impedance.

While maintaining the desired output grid emulator always consumes sinusoidal current from the grid (with distortion less than 2%). Input side also has an optional control for the reactive power supplied by the grid emulator on the grid side. Table I shows the basic characteristics of a grid emulator integrated in the research and development station. The grid emulator acquired by the Faculty of Technical Sciences is a three-phase 15 kVA model, with a possible 13.5 kW of active output power.

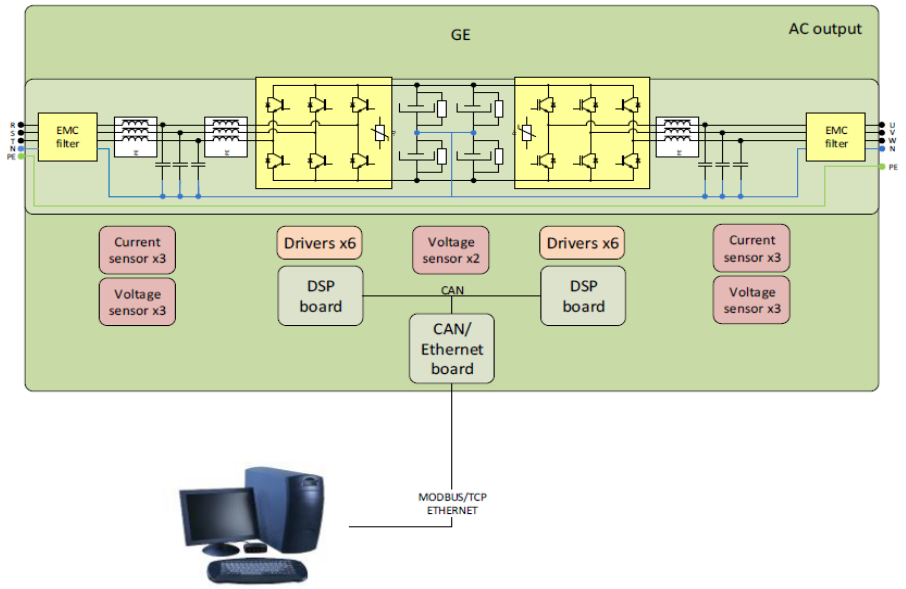


Figure 5. Block diagram of a grid emulator taken from grid emulator operating instruction

TABLE I. GRID EMULATOR BASIC CHARACTERISTICS

GE 15		
Feature	**Description**	**Value**
Input voltage	Rated	3 x 400 V (with neutral and earth)
Input Voltage range		+15 % / -20%
Input AC current	Rated	20 A
Input Frequency		50/60 Hz
Power Factor		1
Efficiency	At full load	> 92 %
Overload		125 % for 10 min 150 % for 60 sec
Output voltage	Phase-Phase	0-480 V
Output AC current	Single-Phase	0-20 A
Output Frequency		40 - 400 Hz
Harmonic content	Per Phase	1^{st} - 15^{th}
Faults	Voltage dip Over and Under Voltage Frequency variation Flicker Harmonic sequence	
Measurements	Input Power Output Voltages Output Currents Output Power Temperatures	
Protection	Over Current and Over Voltage Short circuit Over temperature	

III. IMPLEMENTATION AND EXPERIMENTATION SOFTWARE TOOLS

For the fast prototyping systems, the implementation of the control algorithm must be equally as fast. The new approach to development of control algorithms is recognized as Total Development Environment (TDE). This concept allows for a full visual block-oriented programming of dynamic real time systems. TDE has been reported in a large number of papers and it is hard to distinguish who was the pioneer and where TDE was mentioned firstly [7, 8, 9, 10]. The principles of TDE introduction to the particular laboratory setup is described in [1]. The main advantage of TDE is that the programming method is highly modular and fairly easy to comprehend. This makes the research station ideal tool for beginner, as well as experienced users. Implementation and experimentation tools for the grid connected converter control are Matlab/Simulink and Control-Desk respectively.

The addition in the control is the introduction of CINERGIA software for control of grid emulator. Even though the grid emulator can be controlled through several different protocols, even directly from the dSPACE system (thus implementing it in TDE), software tool for control of grid emulator provides very good functionality. Main advantage of the software is its easy to use nature. The overview of the control software is given in the Fig. 6. The software allows for a control of grid parameters on the output side, with a control option of reactive power injection at the input side. All of the measured values are available to be observed by the control software as well. In addition to control of voltage parameters in the stationary regime, software allows for voltage parameters during all of the mentioned faults to be set up and generated.

Figure 6. Grid emulator control software

IV. EXPERIMENTAL EXAMPLES

In addition to the research oriented at electrical drives and control shown in [1], the presented station without grid emulator has been used to investigate the grid connected converter behavior to some extent. Results achieved are shown in [11], where current controller parameter influence on power quality was investigated. With addition of the grid emulator, the research can be expanded to investigate the influence of the grid conditions on control quality and power quality issues of the grid connected converter.

These experimental results aim to demonstrate the possibilities and the versatility of the grid emulator. The measurement was carried out using Chauvin Arnoux power quality analyzer. The tests were done with the open circuit output side of the grid emulator. To establish a base point, the output reference was set at 230 V, with no phase shift. In addition, no disturbance has been activated for the first test. It can be observed from the Fig. 7 that the output signal is almost ideally sinus waveform. Total harmonic distortion for this particular case is less than 0.5 %, thus we can consider this mode of operation an idealized stat of the grid voltages.

In order to fully demonstrate the possibilities of generation of the voltage with a certain harmonic content, set-point values for the voltage harmonic content for every phase of the three-phase system is set independently. Waveform of the voltages with desired harmonic content is shown in Fig. 8. By comparison of the spectrum with the reference values entered into the control software, it can be concluded that grid emulator has almost ideal matching between the reference and the output harmonic content. It is also apparent that this grid emulator can emulate any harmonic content of the voltage, thus the power quality issues of the grid connected converter can be further investigated. Furthermore, it can act as an ideal network, any classical network (by adopting the standard harmonic content of the grid voltages), or any other type of network with different harmonic content of the voltages. This last part is particularly interesting in case of grid connected converters running in an isolated network, i.e. islanding mode.

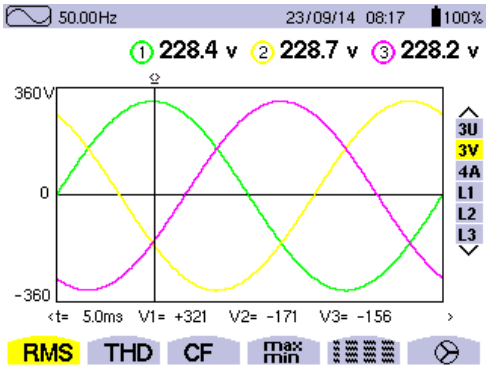


Figure 7. Emulated grid voltages

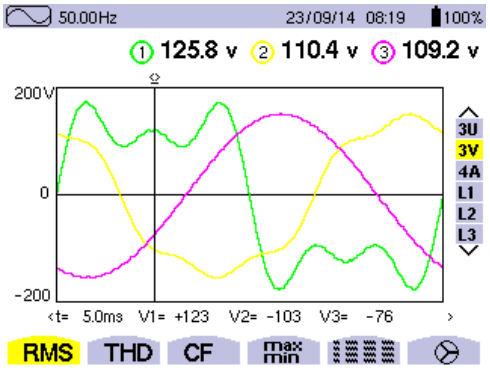


Figure 8. Emulated grid voltages with harmonic content

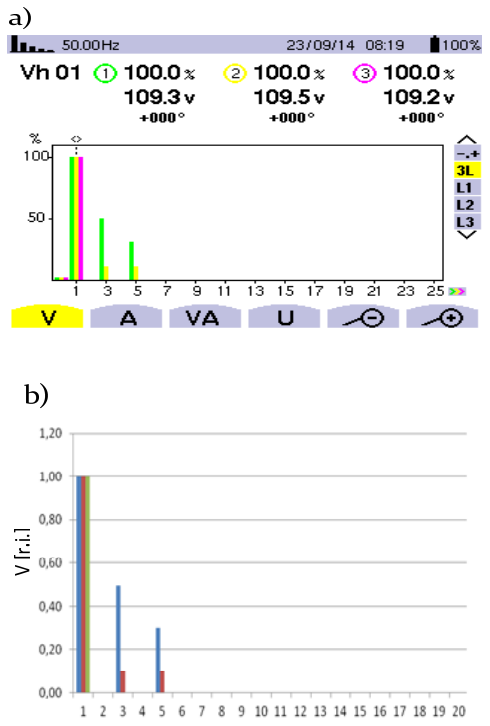


Figure 9. Harmonic content of emulated grid voltages (a) and reference values (b)

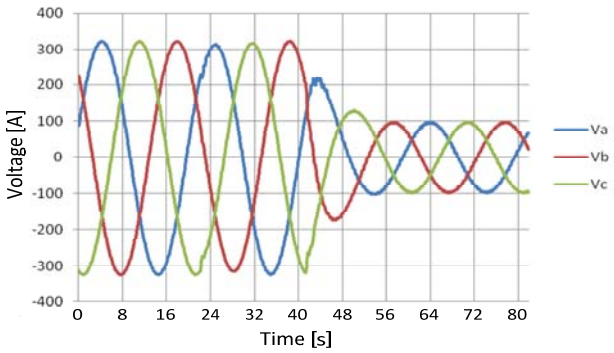


Figure 10. Symmetrical voltage dip

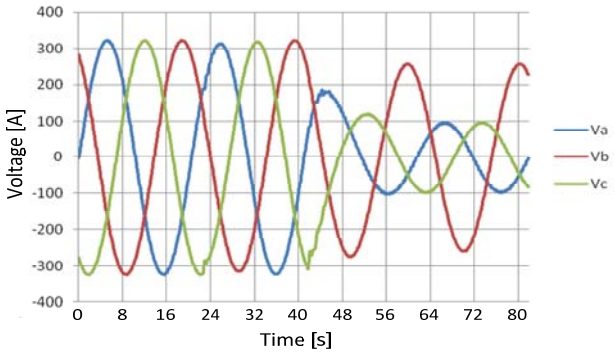


Figure 11. Asymmetrical voltage dip

In order to investigate the behavior of the grid connected converter and its associated control algorithm during faults, grid emulator needs to be able to generate faults like voltage dips. Figs. 10 and 11 show the voltage dip emulated by the grid emulator. The set-point value for the voltage, during the voltage dip shown in Fig. 10, has been set at 30 % for all three phases and the length of the disturbance is set at 3 periods of the waveform. As it can be seen in Fig. 10 the voltage dip has been emulated successfully. To demonstrate the versatility of the emulator, an asymmetrical voltage dip has been emulated and the result is shown in the Fig. 11. It is shown that every voltage can be controlled independently, allowing for a large number of possibilities when testing the grid connected converter during disturbances. For the test shown in Fig. 11 two phases have the voltage set-point at 30 %, where as one has the set-point at 80 %. In addition to these, successful test have been carried out in generating flicker, as well as generating harmonic sequence disturbance.

V. CONCLUSION

As the number of the grid connected converters increase in the future, it will become necessary to study the influence of the grid connected converter on the power system. In addition to this research area, the influence of grid conditions on the operation of the grid connected converters will be investigated as well. Together with those, future converters will have to operate under different grid disturbances. In order to prepare for the research in those areas at the Faculty of Technical Sciences, Chair of Power Electronic and Converters has developed a sophisticated research and development station for control of grid connected distributed energy sources. This research station is an expansion of the highly modular advanced laboratory setup for control of electrical drives. With the introduction of the grid emulator, this research station becomes more than capable of reproducing any desired grid conditions, without influencing the actual power system. As it has been shown, the sophisticated research station is a versatile tool for research, development and testing the grid connected converter control algorithms, under varying grid conditions. In addition to its R&D features, this station becomes a great tool for educating young scientists in the area of grid connected distributed energy sources.

ACKNOWLEDGMENT

This paper is a result of the scientific project No. III 042004 of Integrated and Interdisciplinary Research entitled “Smart Electricity Distribution Grids Based on Distribution Management System and Distributed Generation“, funded by Republic of Serbia, Ministry of Education and Science.

REFERENCES

[1] Boris Dumnic, Dragan Milicevic, Bane Popadic, Vladimir Katic, Zoltan Corba, “Advanced laboratory setup for control of electrical drives as an educational and developmental tool”, The IEEE Eurocon 2013, 1- 4 July, Zagreb, Croatia, pp. 903-909.

[2] The dSpace webpage. [Online]. Available: http://www.dspace.com/en/ pub/home/products/hw/modular_hardware_introduction.cfm

[3] The Danfoss website. [Online]. Available: http://www.danfoss.com/

[4] F. Blaabjerg, J.K. Pedersen, M.P. Kazmierkowski, P. Thøgersen and M. Tønnes, The Danfoss Professor Programme in Power Electronics and Drives at Aalborg University, Proc. of EAEEIE '98, Lisbon, Portugal, pp. 197-203, October 1998.

[5] The Cinergia webpage [Online]. Available: http://www.cinergia.coop/ cinergia-product/ge-grid-emulator-series

[6] Dumnic, Boris P. "Control of Induction Generator with Back-to-Back Converter Connected to Power Grids" PhD Thesis, Novi Sad, Serbia, 2013.

[7] R. Teodorescu, M. M. Bech, F. Blaabjerg, J. K. Pedersen, A Modern Laboratory for Teaching Electrical Drives at Aalborg University, Tomorrow's Education in Electrical Technologies, March 2001, Liège, Belgium.

[8] J. B. Dabney, T. L. Harman, *Mastering Simulink*, Pearson Prentice Hall, NJ, 2004.

[9] R. Teodorescu, M. Bech, F. Blaabjerg, J. Pedersen, A new approach in teaching power electronics control of electrical drives using real-time systems, The 7th Workshop on Computers in Power Electronics, COMPEL 2000.

[10] H. Sschutte, TDE an integrated toolset for real-time control applications, proceedings of MOVIC 1998.

[11] Bane Popadić, Vladimir Katić, Boris Dumnić, Dragan Milićević, Zoltan Čorba, "Current controller parameter influences on THD for grid connected converter", 17th International Symposium on Power Electronics Ee2013, October 2013, Novi Sad, Serbia

Low voltage ride-through capability of distributed generator connected to the grid through the back-to-back converter

Zoran Ivanović, Evgenije Adžić, Stevan Grabić, Vlado Porobić, Vladimir Katić
University of Novi Sad
Faculty of Technical Sciences
Novi Sad, Serbia
zorani@uns.ac.rs, evgenije@uns.ac.rs, grabics@uns.ac.rs, poroba@uns.ac.rs, katav@uns.ac.rs

***Abstract*—This paper deals with low voltage ride-through capability of distributed generator unit. Distributed generator, considered in this paper, is based on squirrel cage induction machine and it is connected to the grid through the back-to-back converter and LCL filter. With an increasing distributed generation grid connection requirements in almost every country require generation unit to stay connected to the grid and ensure active and reactive power injection. In this paper we proposed improved dual vector current controller to deal with the unbalance imposed by the electrical grid. Controller provides injection of active and reactive power to the grid, even if the voltages are lower than the nominal one. Results are validated using contemporary hardware-in-the-loop emulation platform, while controller is based on TMS320F2812 DSP.**

Keywords-distributed generation; voltage ride-through capability, back-to-back converter; hardware-in-the-loop.

I. INTRODUCTION

Distributed generation (DG) play very important role as a means of achieving increased power system reliability [1]. Wind energy systems, photovoltaic and small hydro plants are commonly employed type of DG units. They have a lot of positive impacts on the grid such as, lower capital costs due to smaller size and possibility to contribute to the overall system stability. Most of distributed generation use renewable sources and they require a lot of power electronics interface to match DG characteristic with the grid requirements [2], [3]. Controllable power electronics interface is most often realized in the form of voltage source convertor (VSC). One complete DG system is shown in Fig. 1. Squirrel cage induction generator is connected through the back-to-back converter, LCL filter and transformer to the grid.

In the past there was no requirement by the grid code for a DG to stay connected to the grid during the fault or grid disturbances [4]. The main focus was on protection of DG unit itself. In the last decade, increasing DG penetration, especially wind energy systems lead to fast establishment of variety of grid codes, which define behavior of DG unit in unconventional conditions. Grid code requirements vary from country to country. One of the most commonly used is adopted by the company E.ON Netz. According to this grid code requirements it is defined that in the event of faults in the grid, the generating plants must stay connected and inject active and reactive power. These grid code requirements are illustrated in the Fig. 2. It shows the limit curve for the voltage pattern at the grid connection of induction generator based generating plant.

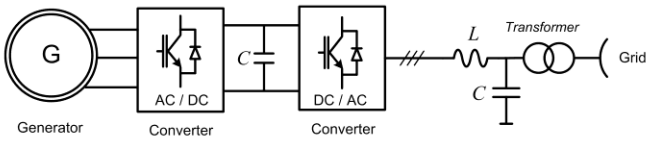


Figure 1. Figure 2 DG unit connected to the grid through the power electronic interface

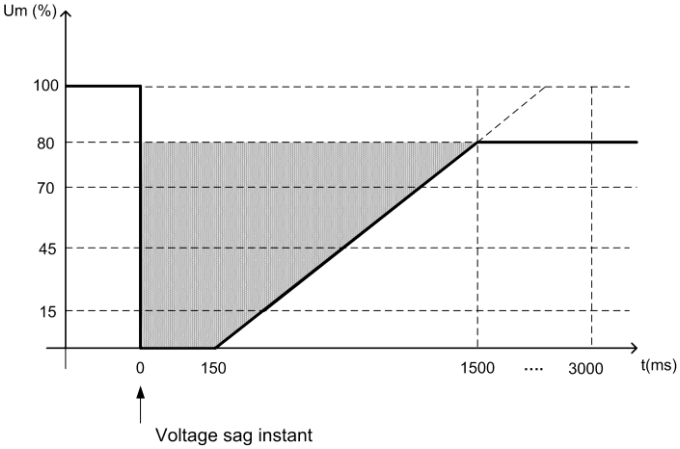


Figure 2. Figure 1 E.ON grid requirements for DG connection

There are many grid codes which require grid support by active and reactive power injection during the fault in the event of voltage sags. However, specific requirements depend on the specific characteristics of each power system and the protection method employed.

In this paper one control algorithm which solves some power quality problems in case of grid voltage sags is presented. Balanced and unbalanced voltage sags are considered. Controller is based on standard dual vector current

control technique and introduces regulation of negative sequence component.

Tests on DG systems is difficult to carry out in real laboratory due to high power rating of the hardware, system complexity, difficulty associated with the generation of desired grid voltage profile, and impartibility to disconnect DG unit form the system. Instead of working on a real system we can use real-time emulator where power stage is emulated on appropriate high speed platform, while controller is real. Here, we used HIL emulation platform to evaluate the ability of proposed control algorithm to meet grid code requirements in severely balanced and unbalanced grid conditions[5], [6]. The entire hardware is emulated in real–time on the FPGA platform with a fixed simulation step time of 1 μs. The FPGA based platform interacts with the controller through the custom made I/O board. The control algorithm is realized using a control platform based on the TMS320F2812 DSP.

II. VOLTAGE SAGS

A three phase fault in the system leads to an equal voltage drop in each phase. Nonsymmetrical faults leads to drops in one, two or three phases, with not all phases having the same drops [7]. Unbalanced voltage sags caused by network faults introduce negative sequence grid voltage and current components. The control and operation of a grid-connected VSC under these circumstances have been widely investigated in the literature [8]–[18].

One example of unbalanced voltage sag is shown in Fig. 3. Voltages in two phases drop down to 50 % of the nominal one, while third phase remain the same. Disturbance lasts for 3 periods.

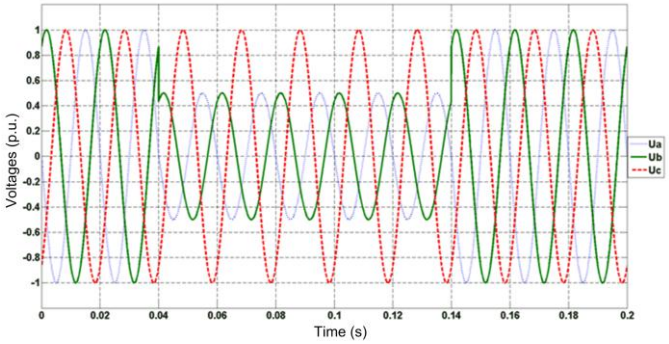


Figure 3. Unbalanced voltage sag

In order to calculate balanced voltage sag amplitude in the radial distribution networks, we can use the schematic shown in Fig. 4. Z_s represents impedance of the source connected to the point of common coupling (*pcc*), while Z_F represents impedance between *pcc* and place of faults.

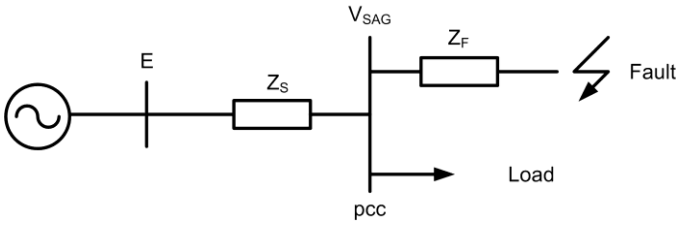


Figure 4. Unbalanced voltage sag

Remained voltage (voltage at pcc) could be calculated as:

$$V_{sag} = \frac{Z_F}{Z_S + Z_F} \cdot E \tag{1}$$

If the fault location is closer to the *pcc*, than impedance Z_F will be smaller, so remained voltage will be lower (Z_F has lower value). In case of unbalanced voltages, schematic from Fig. 4 cannot be used. It is necessary to use symmetrical components domain. Equivalent circuit from Fig. 4 should be separated to positive, negative and zero sequence components in order to calculate voltages and currents properly. Detailed analyses of all possibilities are thoroughly explained in [19] and [20]. There are seven basic voltage sag types according to the ABC classification. They are shown in the Figs. 5 and 6.

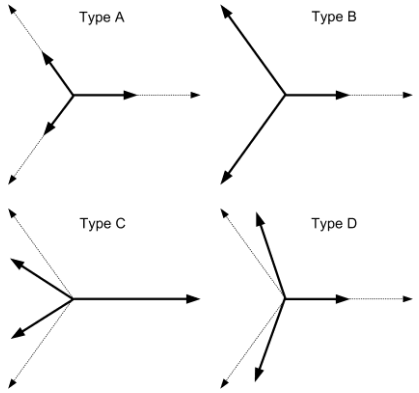


Figure 5. Voltage sags types – ABC classification (a)

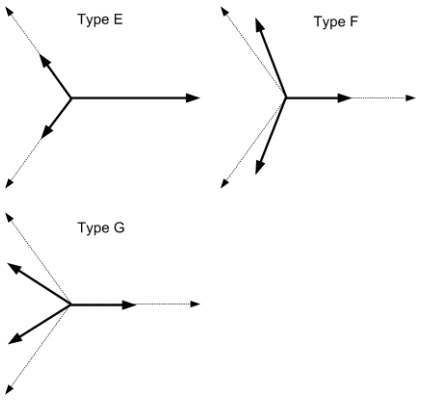


Figure 6. Voltage sags types – ABC classification (b)

Voltage sags A – D are caused by one-phase, two-phase or three-phase short circuit, while voltage sags E – G are consequence of two-phase to ground short circuit. If there is any transformer in the system, like in Fig. 7 it can lead to changing of voltage types at the converter terminals from one form to another one [7].

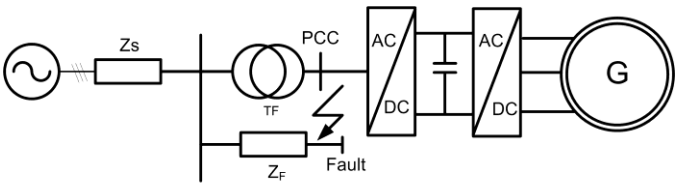


Figure 7. DG unit connected to the grid

Regardless winding connection of the transformer, voltage sag classification mentioned in Figs. 5 and 6 include all possible cases and explain the propagation of three-phase unbalanced sags from one voltage level to another [7].

III. SYSTEM DESCRIPTION UNDER UBALANCED GRID VOLTAGES

In unbalanced systems voltages and currents can be represented by its positive and negative sequence equivalents. A fast and precise detection of both, the positive and the negative sequence angle and magnitude of the voltage component is a vital issue during transient faults in the grid.

Therefore, an unbalanced system of the three phase-voltages (u_a, u_b, u_c) could be represented with its positive ($u_{dq}^p = u_d^p + ju_q^p$) and negative sequence ($u_{dq}^n = u_d^n + ju_q^n$) components, as given by:

$$u_{\alpha\beta} = e^{j\omega t} u_{dq}^p + e^{-j\omega t} u_{dq}^n \tag{2}$$

where $u_{\alpha\beta} = \sqrt{2/3}\left(u_a + u_b e^{j2\pi/3} + u_c e^{-j2\pi/3}\right)$ is the grid voltage vector expressed in the stationary reference frame (using a power-invariant transformation) and ω is the angular grid-frequency. In the same manner, unbalanced grid-currents also appear and they could be represented in terms of positive and negative sequence current components, similarly to (2):

$$i_{\alpha\beta} = e^{j\omega t} i_{dq}^p + e^{-j\omega t} i_{dq}^n \tag{3}$$

where ($i_{dq}^p = i_d^p + ji_q^p$) and ($i_{dq}^n = i_d^n + ji_q^n$).

One case of unbalanced grid-voltages in the original and synchronously rotating reference frame is shown in Fig. 8. It should be noted that in the positive sequence reference frame, a positive component appears as DC, whereas a negative component oscillates at twice the grid frequency. In the negative reference frame, it is the opposite, which is explained thoroughly in [12].

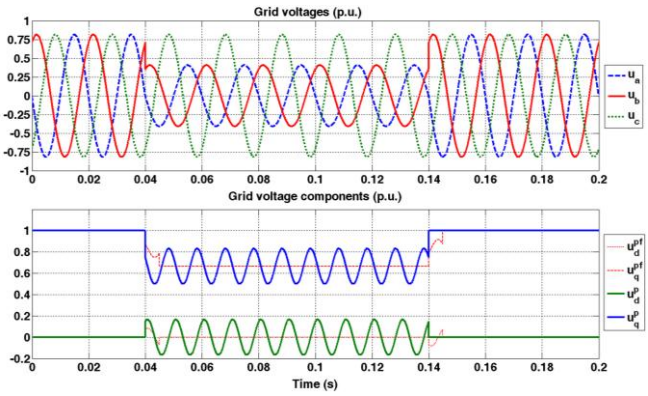


Figure 8. DG unit connected to the grid

The representation of a two-level VSC, used as an actuator in DG application, could be described by differential equation (4) in the stationary reference frame:

$$v_{\alpha\beta} = u_{\alpha\beta} + L\frac{di_{\alpha\beta}}{dt} + Ri_{\alpha\beta} \tag{4}$$

where R is grid resistance, L is grid inductance and

$$v_{\alpha\beta} = \sqrt{\frac{2}{3}}\left(v_a + v_b e^{j2\pi/3} + v_c e^{-j2\pi/3}\right) \tag{5}$$

$$i_{\alpha\beta} = \sqrt{\frac{2}{3}}\left(i_a + i_b e^{j2\pi/3} + i_c e^{-j2\pi/3}\right) \tag{6}$$

where $v_{\alpha\beta}$ and $i_{\alpha\beta}$ denote converter pole voltages and the line currents respectively.

Equation (4) can now be transformed and decomposed into two parts in the positive and negative synchronous rotating reference frames, respectively, as shown in (7) and (8) [12]:

$$v_{dq}^p = L\frac{di_{dq}^p}{dt} + Ri_{dq}^p + j\omega Li_{dq}^p + u_{dq}^p \tag{7}$$

$$v_{dq}^n = L\frac{di_{dq}^n}{dt} + Ri_{dq}^n - j\omega Li_{dq}^n + u_{dq}^n \tag{8}$$

With regards to this, instantaneous apparent power could be expressed as:

$$s = u_{\alpha\beta} i_{\alpha\beta}^* = p(t) + jq(t) \tag{9}$$

where active power $p(t)$ and reactive power $q(t)$ are:

$$p(t) = P_0 + P_{c2}\cos(2\omega t) + P_{s2}\sin(2\omega t) \tag{10}$$

$$q(t) = Q_0 + Q_{c2}\cos(2\omega t) + Q_{s2}\sin(2\omega t) \tag{11}$$

Terms P_0 and Q_0 designate the value of the average power, while P_{c2}, P_{s2}, Q_{c2} and Q_{s2} are the magnitudes of the power oscillations caused by the unbalance. Detailed expressions for all six terms are given in [12].

IV. CONTROL ALGORITHM DESCRIPTION

The block scheme of the DG system is represented in Fig. 9. The back to back converter which is fed from squirrel cage induction generator unit is connected to the grid through the LCL filter and grid impedance. Transformer is also used between grid and DG unit. Three phase grid currents and voltages are measured and transformed to dq domain using the transformation angle which is obtained by employing phase look loop (PLL) estimator. Voltages and currents in dq domain are used to calculate active and reactive power according to Equation (10) and (11). These expressions are used in order to design dual vector current controller, which is employed in case of unbalanced voltages.

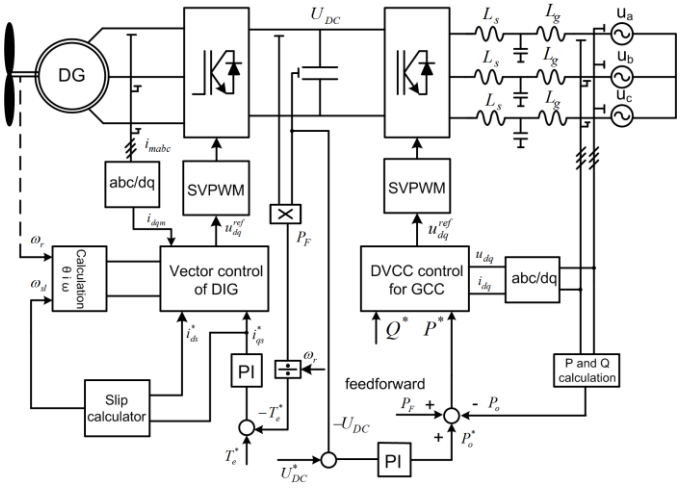


Figure 9. DG unit connected to the grid

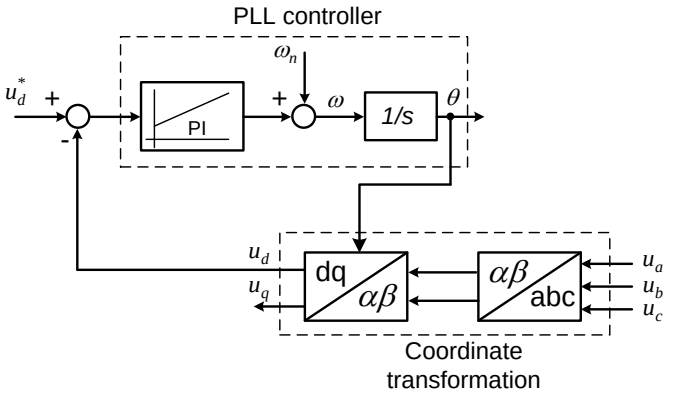


Figure 10. Determination of grid voltage angle (PLL)

The current controller (DVCC) used here consists of a pair of PI controllers that control the positive and negative sequence components separately and are implemented in two different, rotating reference frames. Details about the controllers and the extraction of sequence components can be found in [8] and [12]. DVCC controller is shown in Fig. 11.

In order to generate the proper current references, we should consider the following equation:

$$\begin{bmatrix} P_0 \\ Q_0 \\ P_{C2} \\ P_{S2} \end{bmatrix} = \begin{bmatrix} u_d^p & u_q^p & u_d^n & u_q^n \\ u_q^p & -u_d^p & u_q^n & -u_d^n \\ u_q^n & -u_d^n & -u_q^p & u_d^p \\ u_d^n & u_q^n & u_d^p & u_q^p \end{bmatrix} \begin{bmatrix} i_d^p \\ i_q^p \\ i_d^n \\ i_d^n \end{bmatrix} \tag{12}$$

Solving this equation in terms of grid currents components we could get current references for DVCC controller.

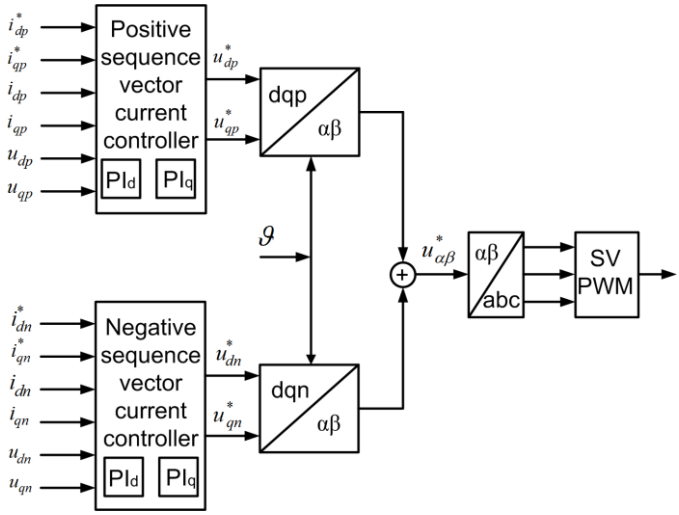


Figure 11. Dual vector current controller

Using this control structure, it is possible to eliminate DC link voltage oscillations, caused by unbalance, as well as oscillations in active power.

In order to protect power electronics converter currents in case of voltage sags are limited. As a consequence we will generate less power in case of disturbances, which will reflect on electrical torque of DG unit. Control of induction generator is classical indirect field vector control.

V. HIL EMULATION AND REAL-TIME IMPLEMENTATION

In this paper, a universal ultra-low latency HIL FPGA-based platform, dedicated to power electronics applications, is used [7]. The proposed platform comprises the emulator hardware and application software that supports a variety of circuit configurations. Electrical scheme which is emulated here is shown in Fig. 12. Power electronics hardware, which include generator, transformer, power part of power electronics converter and passive elements are emulated with the HIL600 platform with a time step of 1 μs. More details about HIL device can be found in [6]. The system parameters can be easily modified through the appropriate graphical user interface. Used concept is shown in Fig. 13 Control algorithms

are realized using a control platform based on TMS320F2812 DSP.

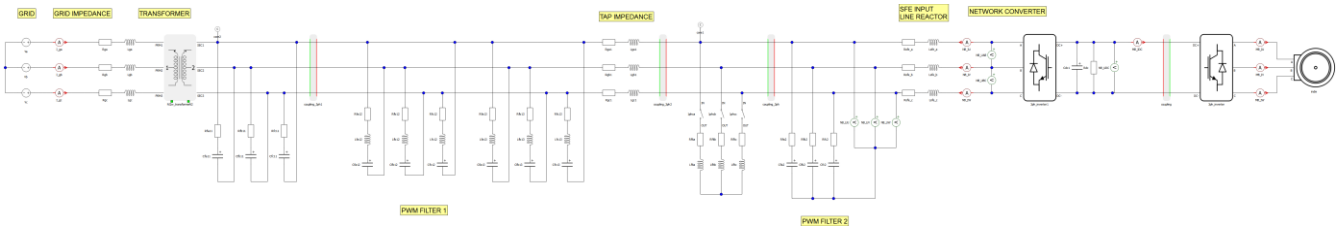


Figure 12. Schematic diagram of DG system

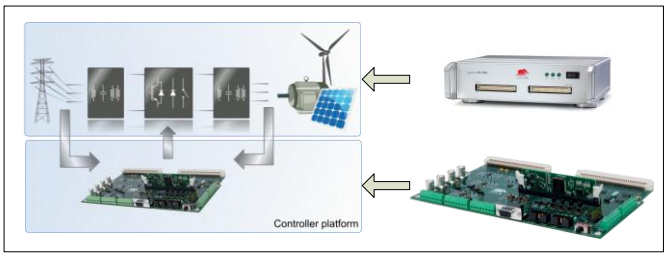


Figure 13. HIL principle

VI. HIL EXPERIMENTAL RESULTS

In this section the performance of the proposed control principle is verified. Control algorithm is tested for three cases: performance in normal condition, behavior in case of balanced voltage sag and behavior in case of unbalanced voltage sag. Figs. 14 – 18 illustrate behavior of DG based system in normal voltage condition.

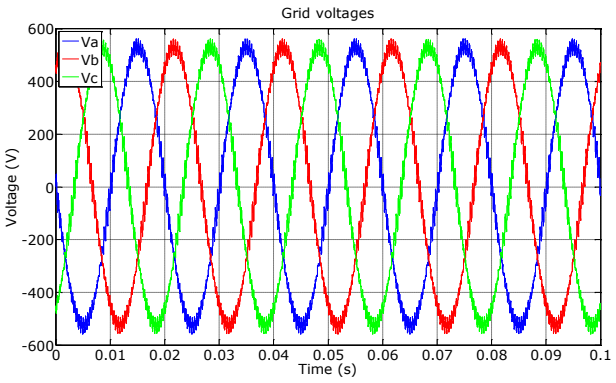


Figure 14. Grid voltages – no fault

Grid voltages are nominal with the line voltage of 400 V. Parameters of electrical transformer are: 1.5 MVA, 11.4/0.63 kV/kV. When we say voltage, we think of voltage after the transformer. From Fig. 15 we can see that grid currents are balanced. There is some ripple due to chosen switching frequency of 2 kHz, but PWM filter cut off most of it. It is the similar with the generator line currents, shown in Fig. 16. Machine electrical torque is quite constant and equals 2500 Nm. DC bus voltage is also stable, which indicate that the power transfer is correctly managed. In this case generator produce 560 kW of active power, while the reactive power is keep to zero.

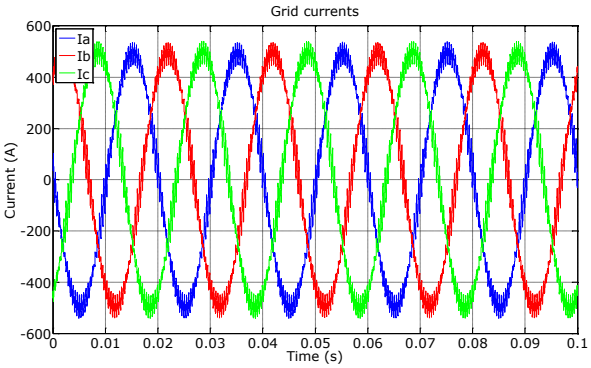


Figure 15. Grid currents – no fault

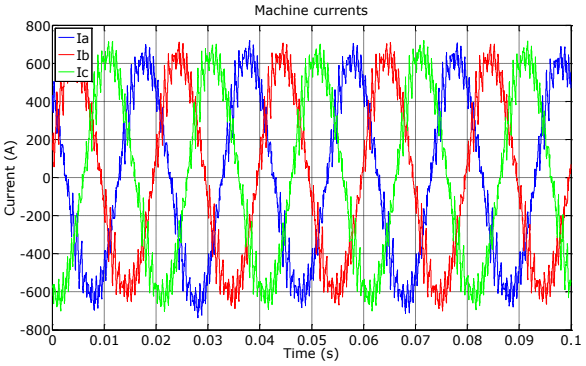


Figure 16. Machine currents – no fault

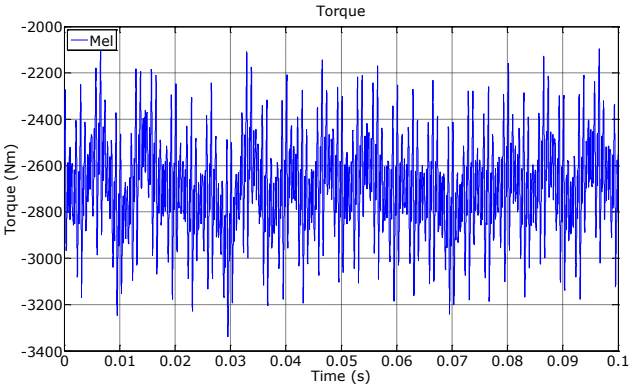


Figure 17. Machine electrical torque – no fault

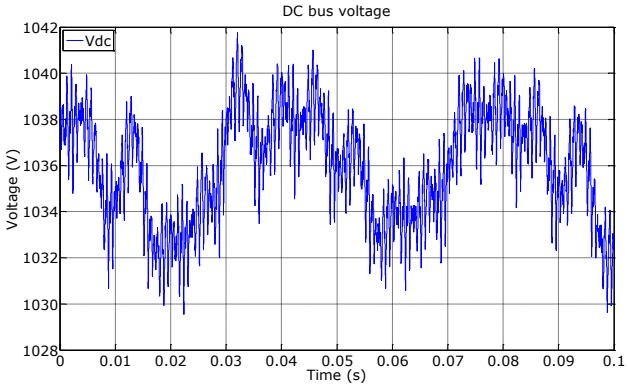


Figure 18. DC bus voltage – no fault

Figs. 19 – 23 show test results in case of balanced voltage sag. Balanced voltage sag is introduced using RL impedance and contactor. With the proper selection of R and L we can chose the depth of voltage sag.

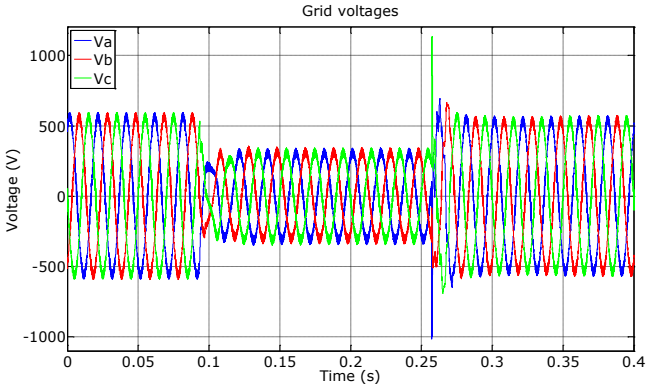


Figure 19. Grid voltages – Voltage sag type A

Fig. 19 shows grid voltage in case of balanced voltage sag to the 49% of the nominal voltage. We can see transient in the instant of voltage change which is due to RL circuit dynamic. During the voltage sag grid currents are higher than nominal, but they are limited in accordance to the grid side converter protection. A lower grid voltage and limited grid current imply a decrease in active power flow from the generator to the grid.

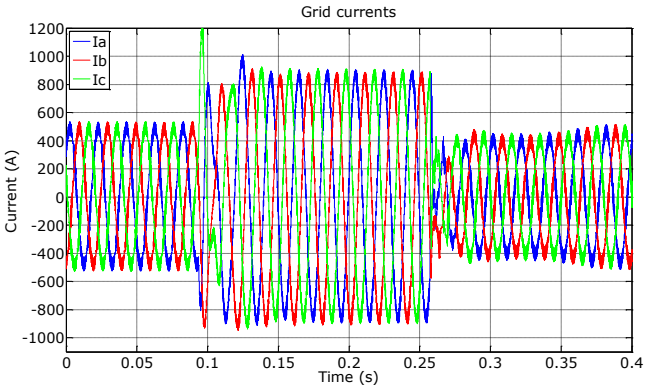


Figure 20. Grid currents – Voltage sag type A

In order to keep power flow balance it is necessary to reduce energy production of the distributed generator. This is achieved by proper reduction of torque reference. We can see in Fig. 22 that electrical torque during the voltage sag is reduced. In accordance with that generator currents are also reduced. DC bus voltage is still stable, except during transient, where we have a little deviation. DC bus stability is indicator that the power flow is correctly managed.

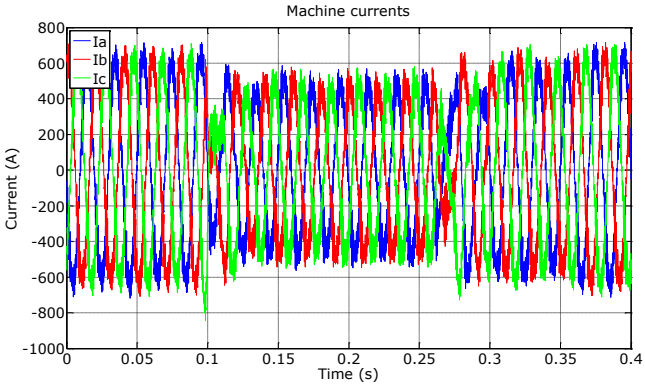


Figure 21. Machhine currents – Voltage sag type A

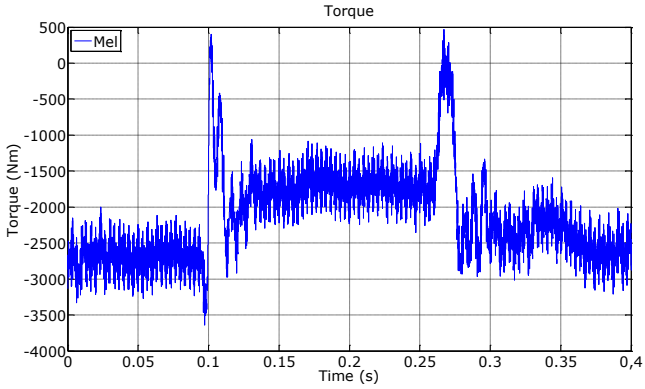


Figure 22. Electrical torque – Voltage sag type A

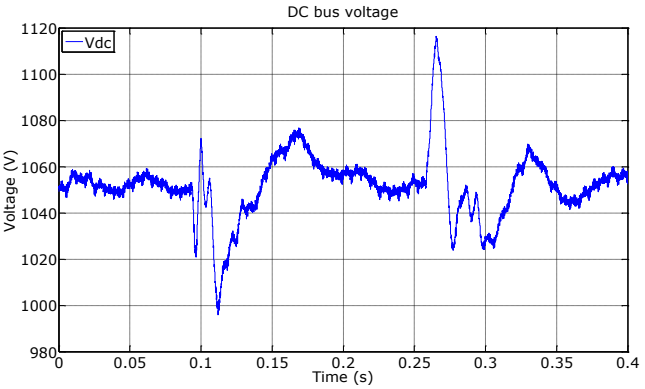


Figure 23. DC bus voltage – Voltage sag type A

The results shown in Fig. 24 – 28 show the response on unbalanced voltage sag type E. Two phases drop to 50 % of the nominal voltage, while third one stay the same as before the voltage sag. Before the voltage sag active power injected to the

grid was 300 kW. We can see that during the voltage sags currents are sinusoidal, but unbalanced. Currents are a little bit higher than the values before sags. Due to limited currents and lower grid voltages, power injected to grid is lower. This means that the production of the generator is a little bit lower, due to the fact that generator speed is kept constant. DC bus voltage is stable in this case, which means that our DVCC controller work properly.

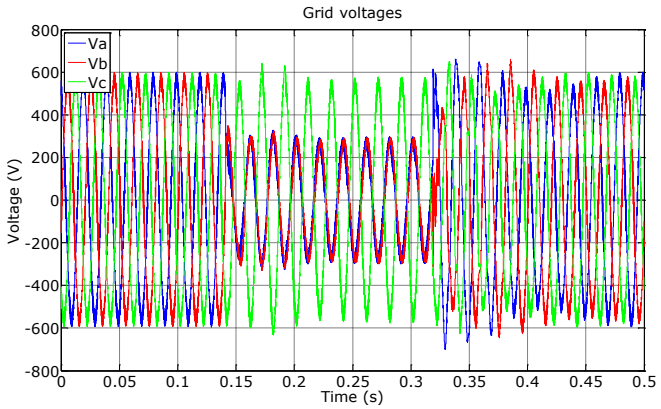


Figure 24. Grid voltages – Voltage sag type E

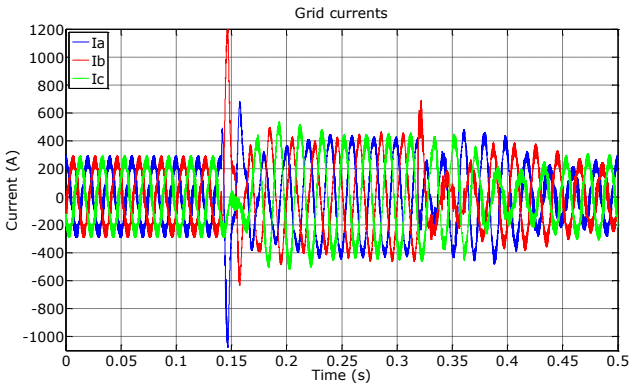


Figure 25. Grid currents – Voltage sag type E

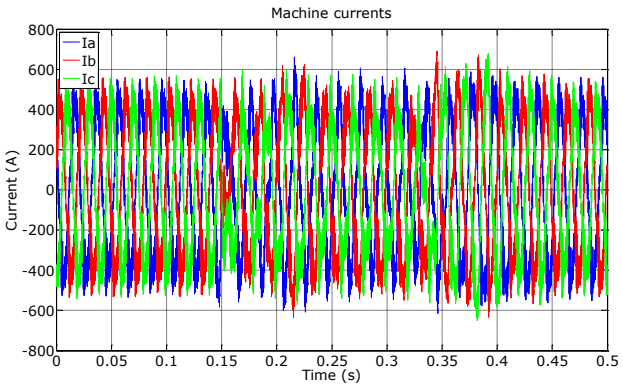


Figure 26. Machine currents – Voltage sag type E

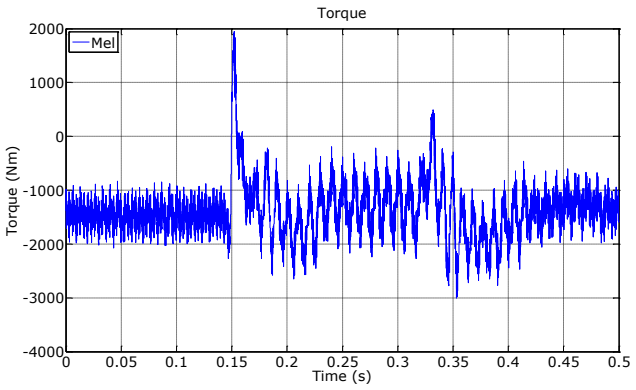


Figure 27. Electrical torque – Voltage sag type E

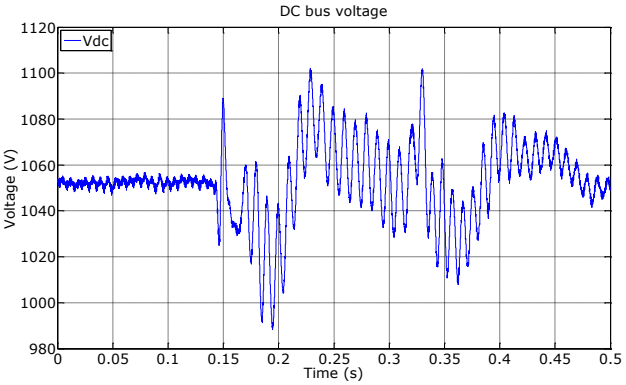


Figure 28. DC bus voltage – Voltage sag type E

VII. CONCLUSION

In this paper voltage ride-through behavior of distributed generator has been investigated. The performance of DVCC has been presented for balanced and unbalanced voltage sags. Proposed control algorithm enable distributed generator to stay connected to the grid during the disturbances and to inject active and reactive power in accordance with its capability. Algorithm principles are verified using Hardware-in-the-Loop emulation platform and TMS320F2812 DSP platform.

ACKNOWLEDGMENT (HEADING 5)

This research was partially co-funded by the Ministry of Education, Science and Technological Development of Republic of Serbia under contract No. III 042004 and by the Provincial Secretariat for Science and Technological Development of AP Vojvodina under contract No. 114-451-3508/2013-04.

REFERENCES

[1] M. Ezzat, M. Benbouzid, S.M. Muyeen and L. Harnefors, "Low-voltage ride-through techniques for DFIG-based wind turbines: state-of-the-art review and future trends," *In Proc. IEEE IECON 2013*, Viene Austria, November 2013, pp.236-242.

[2] Y. Yang and F. Blaabjerg, "Low-voltage ride-through capability of a single-stage single-phase photovoltaic system connected to the low-

voltage grid," *International Journal of Photoenergy*, Article ID 257487, 2013.

[3] F. Hassan, "On power electronics interface for distributed generation applications and its impact on system reliability to customers," *Technical report*, Chalmers University of Technology, Sweeden 2005.

[4] F. Magueed, A. Sannino, "Design of robust converter interface for wind power applications," *Wind energy, published by Wiley Interscience*, vol. 8, pp. 319–332, Jul. 2005.

[5] Z. Ivanović, E. Adzić, M. Vekić, S. Grabić, N. Čelanović, V. Katić, "HIL evaluation of power flow control strategies for energy storage connected to smart grid under unbalanced conditions," *IEEE Trans. Power electron.*, USA, vol. 27, no. 11, pp. 4699–4710, Nov. 2012.

[6] N. Čelanović, I. Čelanović, Z. Ivanović, "Cyber Physical Systems: A new approach to power electronics simulation, control and testing, " *Advances in electrical and ųomputer engineering*, Faculty of electrical engineering and computer sciences, University of Suceava, Romania, issue 1, Feb. 2012, pp. 33–38.

[7] L. Zhang, M. Boolen, "Characteristic of voltage dips (sags) in power systems," *IEEE Trans. Power Del*., Vol. 15, No. 2, pp. 827 – 832, April 2000.

[8] F. Magueed, A. Sannino, J. Svensson, "Design of robust interface for wind power applications", *In Proc. Nordic Power Conference*, March 2004, pp. 1–6.

[9] S. Bifaretti, P. Zanchetta, A. Watson, L. Tarisciotti, J. C. Clare, "Advanced power electronic conversion and control system for universal and flexible power management," *IEEE Trans. Smart Grid*, vol. 2, issue 2, pp. 231–243, May 2011.

[10] T. Noguchi, H. Tomiki, S. Kondo, I. Takahashi, "Direct power control of PWM converter without power-source voltage sensors," *IEEE Trans. Ind. Appl*., Vol. 34, pp. 473 – 479, May/Jun. 1998.

[11] O. Bellmunt, A. Ferre, A. Sumper, J. Jane, "Ride-through control of a doubly fed induction generator under unbalanced voltage sags," *IEEE Trans. Energy Convers.*, vol. 23, no. 4, pp. 1036–1045, Dec. 2008.

[12] H. Song, K. Nam, "Dual current control scheme for PWM converter under unbalanced input voltage conditions," *IEEE Trans. Ind. Electron.*, vol. 46, no. 5, pp. 953–959, Oct. 1999.

[13] G. Saccomando, J. Svensson, A. Sannino, "Improving voltage disturbances rejection for variable-speed wind turbines," *IEEE Trans. Energy Convers.*, vol. 17, no. 3, pp. 422–428, Sept. 2002.

[14] Z. Ivanovic, M. Vekic, S. Grabic, V. Katic, "Control of multilevel converter driving variable speed wind turbine in case of grid disturbances," *in Proc. 12th IEEE EPE-PEMC*, Portoroz, Slovenia, Aug./Sept. 2006, pp. 1569–1573.

[15] C. Ng, L. Ran, "Unbalanced-grid-fault ride-through control for a wind turbine inverter," *IEEE Trans. Ind. Appl.*, vol. 44, no. 3, pp. 845–856, May/Jun. 2008.

[16] Y. Suh, Y. Go, D. Rho, "A comparative study on control algorithm for active front-end rectifier of large motor drives under unbalanced input," *IEEE Trans. Ind. Appl.*, vol. 47, no. 3, pp. 1419–1431, May/Jun. 2011.

[17] Y. Suh, T. Lipo, "Control scheme in hybrid synchronous stationary reference frame for PWM AC/DC converter under generalized unbalanced operating conditions," *IEEE Trans. Ind. Appl.*, vol. 42, no. 3, pp. 825–835, May/Jun. 2006.

[18] F. Wang, J. L. Duarte, M. Hendrix, "Pliant active and reactive power control for grid interactive converters under unbalanced voltage dips", *IEEE Trans. Power Electron.*, vol. 26, no. 5, pp. 1511–1521, May. 2011.

[19] R. K. Sinha, R. Kumar, M.Venmathi, L. Ramesh, " Analyses of voltage sags with different DG for various faulty conditions," *International Journal of computer communication and information system*, Vol. 12., No. 1, ISSN 0976 1349, Dec. 2010.

[20] M. Boolen, "Voltage recovery after unbalanced and balanced voltage dips in three-phase systems," *IEEE Trans. Power Del*., Vol. 18, No. 4, pp. 1376 – 1381, October 2003.

Effects of Individual Battery Charger Station on Power Quality

Vladimir A. Katić,Milan Pecelj,Ivan M. Todorović
University of Novi Sad, Faculty of Technical Sciences
Novi Sad, Serbia

Abstract — **A number of new devices are entering the public grid in a form of battery charging station, either as individual units or as a group connected at the same grid. This paper addresses the question of the effect of individual charging station operation as non-linear unit on the local distribution grid power quality by investigating current and voltage spectrum, i.e. harmonics. Two type of charging modes are considered: Mode 3 (moderate speed charging) and Mode 4 (fast-charging) in case of three-phase AC chargers. The operation is tested on one of traditional topologies using computer simulations. Results showed low voltage distortion and rather high current one. Both distortions are in accordance with IEEE and IEC/EN standards. However, current distortion is very close to limits and in some cases may have effect on transformer overheating and resulted in its de-rating.**

Keywords - battery charger station, power quality, harmonics.

I. INTRODUCTION

Significant efforts have been made to lower the air pollutants emission, both on industrial and domestic level. The long term projections in 2007 showed that there was immediate need for wider usage of renewable energy sources (RES) for powering industry plants and households [1]. Nowadays, greenhouse gasses (GHG) levels are still high and above expected, although contribution of the RES operation in EU, USA and other developed countries is evident.

Similar efforts have been made in transportation sector, especially for passenger vehicles, with the goal to decrease CO_2 emission below 100 g/100km or to achieve zero emission. In that sense, vehicles that combine gasoline and electrical energy (hybrid electric vehicle - HEV) or use just electrical energy (battery electric vehicle - BEV, or simply, electric vehicle - EV) for propulsion are becoming more attractive [2]. There are several reasons for that, like regulation in some countries (USA, especially California), special incentives, spreading of environmental awareness, available adequate infrastructure and probably the most importantly, the performance and prices of these types of vehicles are becoming competitive.

Consequently, there is a great interest in these means of transportation and thus more and more HEVs and EVs are sold. The U.S. Department of Energy expects that about one million of HEVs/EVs will be on the road and that about 400,000 of HEVs/EVs will be sold in 2015 [3]. By 2050 more than 60% of all vehicles in U.S. will be HEV/EV. These numbers show that a significant rise in the number of these vehicles will soon be evident.

Traditionally, HEVs and BEVs had the notorious problem with the driving range and recharging time [2]. Also, weight of the batteries (specifically energy and power density) was generally a limiting factor.

Nowadays, these problems are reduced to some extent whether by implementation of new technologies or via combining different sources with different dynamics and exploitation of their good characteristics in different driving modes. For example, average lithium-based batteries powered cars have up to 160 km (100 miles) driving range, while Tesla Model Seven longer - up to 430 km (270 miles) [4, 5]. Fig. 1 shows estimated driving range of some modern BEV models represented in miles (1 mile=1.61 km) [4].

With the increase of these vehicles on the road, it would be impractical and unreasonable to expect that battery charging will happen only at home, with house mounted chargers, during the night hours. The drivers will demand adequate charging stations all along their driving path, i.e. along the main roads and even more inside city limits. Furthermore, their expectations are that charging time should last no longer than average filling time of gasoline tank or average spending time in gasoline stations (for tank filling, small shopping or refreshment).

There are three different types of chargers depending on charging time or by charger construction: slow, moderate and fast charging stations or level 1, level 2 and level 3 types of chargers. Fig. 2 shows cumulative number of fast-charging stations operational Worldwide [6]. It can be seen that this number is relatively low (around 15,000 in 2014) and that it is lagging to the demand.

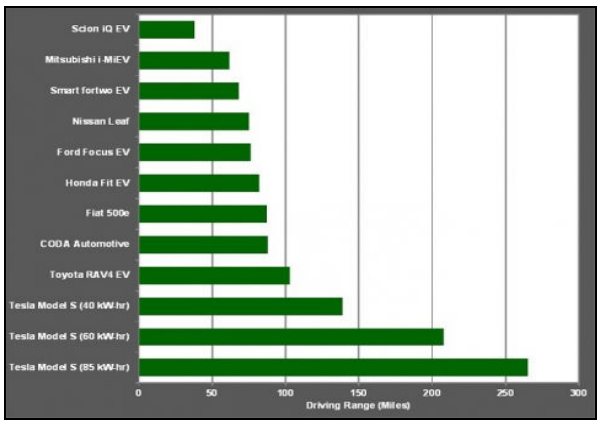


Figure1. Estimated driving range of some modern BEV models [4].

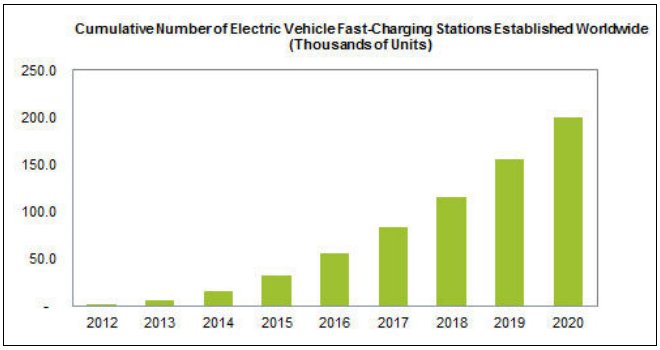


Figure 2. Number of EVs fast-charging stations available Worldwide [6].

Special problem is charger impact to electrical grid, i.e. their effect on power quality. Although, there have been significant improvements in charger construction and operation over the time, this item is still attracting high attention of researchers and wider public. Some surveys suggest that current THD has decreased from THDI=50.1% to THDI=6.12% since 1993 [7]. Even with these, probably deflated, numbers current THD is still high comparing with IEEE Standard 519-1992, IEC 61000 standards family or European EN50160 standard. Additional problem is that manufacturers produce different charging types and in different power levels using different technology. Therefore, it can be noted that chargers can have a strong influence on power quality.

Considering how switching converters work, and their impact on current and voltage THD, the problem of power quality should be addressed. To get into in-depth of this phenomenon it is necessary to develop adequate simulation models and apply computer simulations.

In this paper a model of a charger station is developed for three-phase public supply grid with 0.4 kV rated voltage using Matlab/Simulink software tools. The model represent the most common charger topology, i.e. a train of power electronics converters applied for charging purposes [8]. The model is applied for investigation of impact of mode 3 (moderate speed charging) and mode 4 (fast-charging) charging on power quality in the public grid. The simulation results are shown and some comments on observed harmonic spectrum and obtained THD values are stated.

II. CHOICE OF CHARGERS

Chargers are generally classified based on their charging power (level 1 to 3), or on charging modes (Mode 1 to 4).In the first case, they are divided according to power size, rated voltage levels, number of phases, and by rated current. Charging time depends strongly on the type of charger, charging current (AC or DC) and applied batteries, so different charging modes can be distinguished. There is a difference in classification in the Europe and U.S., as a consequence of using different standards. In Europe and rest of the World the IEC 61851-1 is applied, while in USA the national SAE is in use [9, 10]. Table I shows main characteristics of different types and charging modes [11]. Table I can be used as a guideline for which voltage and current levels are expected to be met in commercially available chargers.

TABLE I CHARGERS CLASSIFICATION.

Type of charger	Power [kW]	Voltage [V]	Current [A]	Charging time
Slow charging speed - Mode 1	3.3	230, AC 1~	16	6 – 8 h
Slow charging speed - Mode 2	10	400, AC 3~	16	2 – 3 h
Moderate charging speed - Mode3	7	230, AC 1~	32	3 – 4 h
Moderate charging speed - Mode3	22	400, AC 3~	32	1 – 2 h
High charging speed - Mode4	43	400, AC 3~	63	20 – 30 min
High charging speed - Mode4	50	400 – 500, DC	100 - 125	20 – 30 min
USA level 1 – AC	1.8	120, AC	15	12 – 14 h
USA level 1 – DC	≤19.2	200-450DC	≤80	~20 min
USA level 2 – AC	7.2	240, AC	30	3 – 4 h
USA level 2 – DC	90	200-450DC	200	~15 min
USA level 3	TBD	TBD	TBD	TBD

The charging is connected with power conversion using different power electronics converters. This type of devices is well known as non-linear one consuming distorted currents, i.e. producing current harmonics. Over the time, many new, sophisticated chargers with good performances (low THD, adjustable power factor) are proposed in literature [12, 13, 14, 15], but they tend to be rather complex both to control and to produce, and thus most of them aren't being commercialized. Nowadays, several common topologies depending on charging mode and power sizes can be distinguished [8]. As manufacturers produce them for specific operating conditions, with various topologies and for different charging time, similar chargers may exhibit different effects on the grid. Still, the goal here is to test worst case scenario and thus a more traditional topology for charger is chosen.

Fig. 3 shows a topology of battery charging station connected to the public grid. The case of a charger plug-in the three-phase public supply grid with 0.4 kV rated voltage, which is common in Europe is considered. Difference between moderate (Mode 3) and fast (Mode 4) charging speed chargers is only in referenced currents, i.e. they are the same but inject different currents into the battery. An example of such a topology, similar to the one shown in Fig. 3 and with nominal current and voltage for appropriate chargers shown in Table I can be found in chargers produced by ABB (Terra series) [16].

The model is configured in such a way to reflect a real condition in the field, as much as it is possible. The medium voltage (MV) public grid is three phase source with nominal voltage of 10 kV and frequency of 50 Hz. Instead of pure sinusoidal voltages, it is assumed presence of voltage harmonic distortion of THDU≈2% (5^{th} and 7^{th} harmonics with 0.02 and 0.015 per unit amplitudes, respectively). This is done in accordance with actual situation in MV grid, so that the low voltage (LV) local grid has dominant 5^{th} and 7^{th} harmonic, already present [17].

Standard transformer of rated power 630 kVA and the transfer ratio of 10kV/0.4kV is used. The charging station is connected to LV grid, 0.4 kV, at the transformer substation.

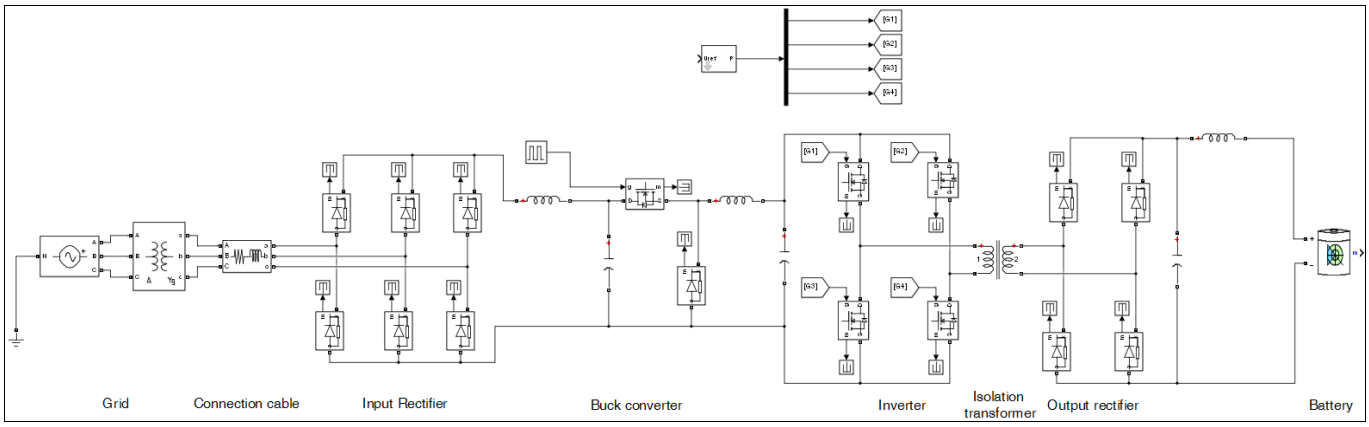


Figure 3. Matlab/Simulink scheme of selected battery charger topology.

The distance from substation to the possible location, where a charger station may be placed is assumed to 50 meters, so cables are modelled in simulation accordingly. The cable has equivalent resistance of 10 mΩ and inductance of 0.3 mH.

For input rectifier, simple uncontrolled three-phase bridge diode rectifier is used. Output of such a converter is rated to 450 V.

Next, considering the voltage level after the rectifier and nominal voltage of the batteries (in this case 300 V), it is evident that step-down converter should be used. Out of many converters, for this purpose simple buck converter was used and modelled. Buck's inductance is set to 2mH and capacitance is set to 660µF. Taking into account the type of the rectifier and DC/DC converter used, it is obvious that only one direction of energy is assumed in this paper, i.e. the possibility that the batteries are used as the temporary energy bank for distribution grid is excluded. After buck converter a filter is connected. This filter has inductance of 15 mH and capacitance set to 33 µF.

In order to have galvanic isolation, a single phase inverter is used. At the output of the isolation transformer a single-phase bridge rectifier is connected. Transformer has nominal power 250 kVA, nominal frequency 10 kHz and 400V/400V primary/secondary voltages. A LC filter is connected to the rectifier output. Filter's inductance is set to 1mH and capacitance is set to 440µF.At the end a lithium-ion battery was chosen. The battery model used in the simulation is taken from Simulink's/SymPower System toolbox library and is set to have nominal voltage of 300 V and maximum capacity of 25 Ah. Initial state-of-charge is set to be 50%. Also, parameters for all diodes and transistors are left on their default values. The buck and the inverter are controlled in open loop, since dynamic behaviour of the charger is not of interest for this paper.

III. SIMULATION RESULTS

In this paper two types of battery charger station operation are simulated: 3 phase AC, 0.4 kV Mode 3 (moderate charging speed) charger and 3 phase AC, 0.4 kV Mode 4 (high charging speed). Grid voltage, current and harmonic content of both voltage and current have been observed and calculated. Results are presented in Figs. 4-9.

A. *Mode 3 operation simulation results*

Figs. 4, 5 and 6 depict simulation results obtained for a single battery charger station operated in the mode 3 (moderate charging speed/time) and connected to the substation, together with MV grid. Fig. 4 shows line voltage, phase voltage and phase current at the charging station input terminals. Fig. 5 shows harmonic spectrum of the phase voltage at the charging station input terminals. Fig. 6 shows harmonic spectrum of the phase (line) current at the charging station input terminals.

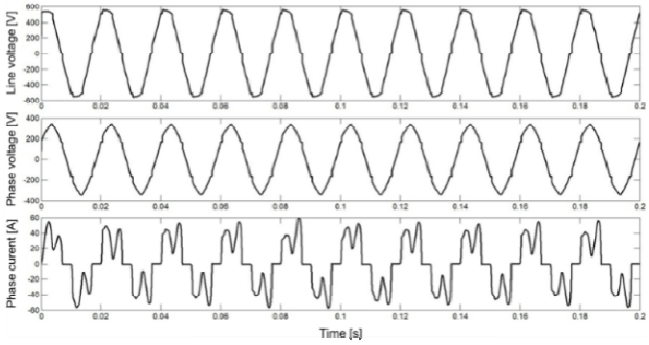


Figure 4. Line voltage, phase voltage and phase current at the charging station input terminals (Mode 3).

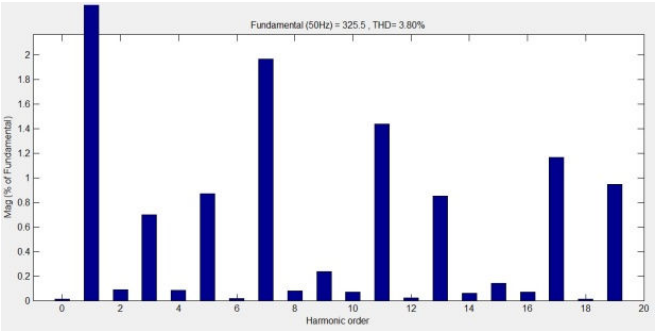


Figure 5. Phase voltage harmonic spectrum at the charging station input terminals (Mode 3).

B. *Mode 4 operation simulation results*

Figs. 7, 8 and 9 depict simulation results obtained for a single battery charger station operated in the mode 4 (fast charging speed / time) and connected to the substation, together

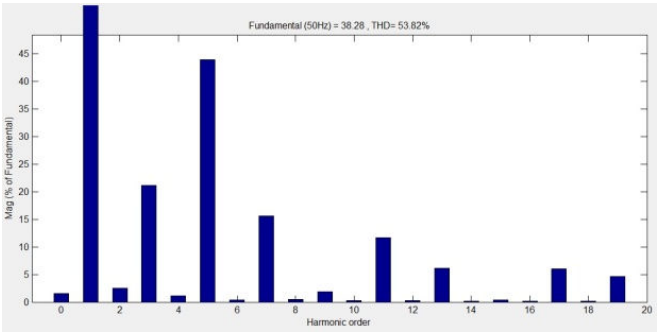


Figure 6. Phase current harmonic spectrum at the charging station input terminals (Mode 3).

with MV grid. Fig. 7 shows line voltage, phase voltage and phase current at the charging station input terminals. Fig. 8 shows harmonic spectrum of the phase voltage at the charging station input terminals. Fig. 9 shows harmonic spectrum of the phase (line) current at the charging station input terminals.

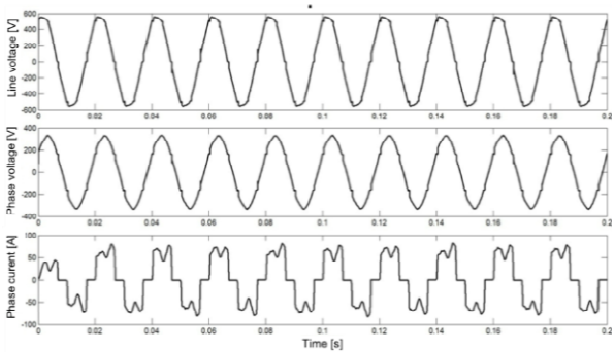


Figure 7. Line, phase voltage and phase current at the charging station input terminals (Mode 4).

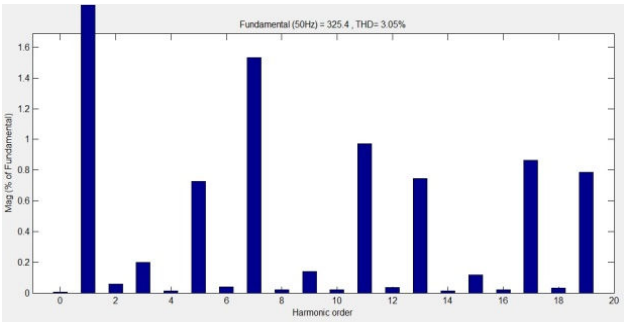


Figure 8. Phase voltage harmonic spectrum at the charging station input terminals (Mode 4).

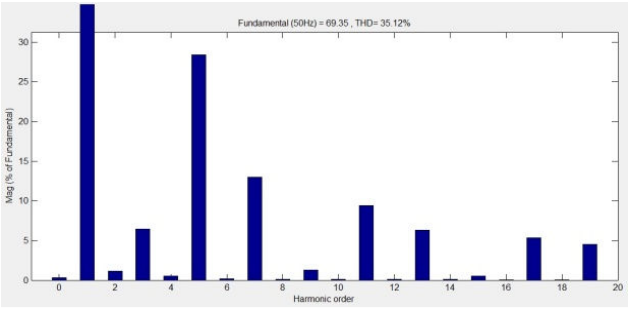


Figure 9. Phase current harmonic spectrum at the charging station input terminals (Mode 4).

IV. Discussion

A. Mode 3 operation

In the case of mode 3 operation of the battery charging station results show a low voltage distortion and a high current one (Fig. 4).

Total harmonic distortion of phase voltage is THDU=3.8%, while distortion from individual harmonics is less than 2% (Fig. 5). For comparison purposes, the IEEE Standard 519 and IEC 61000/EN61000 standard series are considered.

By IEEE STD 519 standard voltage THD for MV systems less than 60 kV must be under 5% and all individual harmonics should be fewer than 3% [18]. It can be seen that only 7th harmonic have magnitude close to 2%, while the rest are well under 2%.

By IEC 61000-2-4 standard voltage THD for MV systems is limited to 8% (Class 2) [19]. For individual harmonics, different limits have been stipulated between 1% and 6% (Fig.10). It can be seen that all voltage harmonics are below given limits.

As THDU=2% in MV grid was assumed as existing harmonics distortion, it may be concluded that the effect of charging station operation regarding voltage harmonics is in line with existing standards.

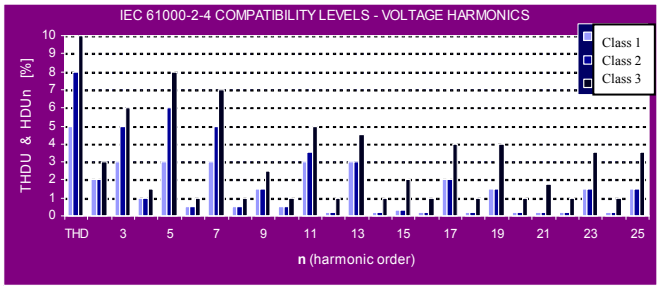


Figure 10. IEC 61000-2-4 Compatibility levels for MV voltage harmonics [19]

Fig. 6 shows that current harmonics are high. Total harmonic distortion is THDI=38.28%, while individual one lies between 5% and 43% for odd harmonics.

By the IEEE STD 519 standard, current harmonics are above limits. All odd harmonics under 11th should be less than 12% for systems with short circuit/maximum load (first harmonic) current ratio between 100 and 1000 which is here the case. Thus, current harmonics are well above the desired level. The biggest problems are the 5th, 7th, 9thand 11thharmonics.

If IEC 61000-3-4 standard is considered (Table II), it can be seen that limits are satisfied only for ratio of short circuit power and load power 450 and above [20].

B. Mode 4 operation

In the case of mode 4 operation of the battery charging station results again show a low voltage distortion and a high current one (Fig. 7). It is obvious that voltage distortion is similar to mode 3 case (THDU=3.05%), while input current is less distorted (THDI=35.12%; HDI=5%-28%). The same conclusions are valid for this mode, except that IEC 61000-3-4

standard limits are satisfied for ratio of short circuit power and load power of 250 and above.

TABLE II IEC/EN 61000-3-4 CURRENT HARMONICS LIMITS (ST 2) [20]

	HDIn – ODD HARMONICS ONLY [%]					THDI [%]
R_{SCE}= S_{SC}/S_L	n=5	n=7	n=11	n=13	n≥15	non-sim. / 3~ simetr.
<66	14	11	10	8	/	25 / 16
120	16	12	11	8	/	29 / 18
250	30	18	13	8	/	39 / 35
450	50	35	20	15	/	51 / 58
>600	60	40	25	18	/	- / 70
- Even harmonics are limited to: $I_n/I_1 = 16/n$						
- Harmonics multiply of 3 are not present						
- Assumption is that the load is balanced						

Regarding effect of harmonics on the line transformer, some researchers indicated that the current THD should be limited to 25–30% in order to have a reasonable transformer life expectancy [7]. In both considered cases obtained current THD is above this recommendation. This means that some problems with transformer overheating may be expected or some level of de-ratting should be applied.

The problem of deteriorated power quality in the local distribution network could be solved by connecting the chargers to another substation with higher power (higher short circuit power), by using harmonic filters or by implementing some more sophisticated chargers topologies.

V. CONCLUSIONS

Two types of battery charger stations have been investigated regarding input current and voltage harmonics, i.e. effects of their operation on power quality. A Matlab/Simulink simulation model have been developed and used for testing. The model parameters are selected in such a way to reflect real conditions in public grid. The results show that voltage harmonics are significantly below existing limits in IEEE and IEC/EN standards. On the other hand, the current harmonics are close to the limits or even higher, especially for some individual harmonics. This may have effect on distribution transformer overheating and reducing its life expectancy.

ACKNOWLEDGMENT

This research was partially co-funded by the by the Provincial Secretariat for Science and Technological Development of A.P. Vojvodina under contract No. 114-451-2248/2011-03 „Research and Development of Energy Efficient Power Supply and Propulsion Systems of Electric Vehicles”.

REFERENCES

[1] A.J. Morrison, “Global Demand Projections for Renewable Energy Resources”, IEEE Electrical Power Conf. – EPC 2007, Montreal (Canada), pp.537-542, Oct. 2007. DOI: 10.1109/EPC.2007.4520389

[2] M. Eshani, Y. Gao, A. Emadi, “Modern electric, hybrid electric and fuel cell vehicles-Fundamentals, Theory and Design”, 2nd Edition, CRC Press, Taylor & Francis Group, Boca Raton (USA), 2010.

[3] U.S. Department of Energy, “One Million Electric Vehicles By 2015”, Status Report, Feb.2011. http://www1.eere.energy.gov/vehiclesandfuels/pdfs/1_million_electric_vehicles_rpt.pdf

[4] U.S. Department of Energy, “Driving Ranges For Electric Vehicles”, Fact #797: September 16, 2013, http://energy.gov/eere/vehicles/fact-797-september-16-2013-driving-ranges-electric-vehicles

[5] http://www.teslamotors.com/goelectric#

[6] ***, “Number of Fast-Charging Stations for Electric Vehicles Set to Rise to Nearly 200,000 in 2020”, IHS Pressroom, Aug. 27, 2013. http://press.ihs.com/press-release/design-supply-chain-media/number-fast-charging-stations-electric-vehicles-set-rise-nea

[7] J.C. Gómez, M.M. Morcos, “Impact of EV Battery Chargers on the Power Quality of Distribution Systems”, IEEE Trans. On Power Delivery, vol. 18, no. 3, pp. 975-981, July 2003.

[8] M. Yilmaz, P. Krain, “Review of Battery Charger Topologies, Charging Power levels, and Infrastructure for Plug-In Electric and Hybrid Vehicles”, IEEE Trans. on Power Electronics, Vol.28, No.5., May 2013, pp. 2151-2169.

[9] IEC, “Electric Vehicle Conductive Charging System - Part 1: General Requirements”, 2.0 Edn., IEC 61851-1, Geneva, 2010.

[10] SAE International, “SAE Electric Vehicle and Plug in Hybrid Electric Vehicle Conductive Charge Coupler”, J1772-201210, 2012.

[11] A. Ayob, W.M.F.W. Mahmood, A. Mohamed, M.Z.C. Wanik, M.F.M. Siam, S. Sulaiman, A.H. Azit, M.A.M. Ali, “Review on Electric Vehicle, Battery Charger, Charging Station and Standards”, Research Jour. of Applied Scien., Eng. and Tech., Vol.7, No.2, 2014, pp.364-373.

[12] A.E. Demian Jr., C.A. Gallo, F.L. Tofoli, J. Batista Vieira Jr., L. Carlos de Freitas, V.J. Farias, E.A.A. Coelho, “A Novel Microprocessor-Based Battery Charger Circuit With Power Factor Correction”, Applied Power Electronics Conf. and Exp., APEC-04,Vol. 3, Feb.2004, pp.1407–1410.

[13] Y.-Ch. Chuang, Yu-L. Ke, “A Novel High-Efficiency Battery Charger With a Buck Zero-Voltage-Switching Resonant Converter”, IEEE Tran. on Energy Conversion”, Vol.22, No.4, Dec.2007, pp.848-854.

[14] L. Wang, J. Liang, G. Xu, K. Xu, Z. Song, “A novel battery charger for plug-in hybrid electric vehicles”, Inter. Conf. on Information and Automation (ICIA), Shenyang (China), pp.168-173. June 2012,

[15] Y.-Ch. Chuang, H.-Sh. Chuang, Y.-H. Liao, Ch.-H. Yang, Y.-Sh. Wang, “A Novel Battery Charger Circuit with an Improved Parallel-Loaded Resonant Converter for Rechargeable Batteries in Mobile Power Applications”, 23rd Int. Symp. on Ind. Electron.-ISIE, pp.353-359, 2014.

[16] ***, “Electric Vehicle Charging Infrastructure - Terra multi-standard DC charging station 53”, ABB Product Leaflet, ABB EV Charging Infrastructure, Rijswijk (The Netherlands), 2014, http://www05.abb.com/global/scot/scot344.nsf/veritydisplay/67fec26aa8fea552c1257d690039af2d/$file/4EVC204305-LFEN_Terra53C-CT-CJ-CJG.pdf

[17] V.A. Katic, “Power Quality – Harmonics”, Edition: Engineering Books – Monographs, No.6, University of Novi Sad - Faculty Technical Sciences, Novi Sad, 2002, (in Serbian).

[18] IEEE Std. 519-1992, "IEEE recommended practices and requirements for harmonic control in electrical power systems", IEEE Press, 1993.

[19] IEC/TR EN 61000-2-4 standard, "Electromagnetic compatibility (EMC) - Part 2-4: Environment - Compatibility levels in industrial plants for low-frequency conducted disturbances", IEC, Geneva, 2002.

[20] IEC/TR EN 61000-3-4 standard, "Electromagnetic compatibility (EMC) - Part 3-4: Limits - Limitation of emission of harmonic currents in low-voltage power supply systems for equipment with rated current greater than 16 A”, IEC, Geneva, 1998.

Multiple Battery Charger Stations Impact on Power Quality

Vladimir A. Katić, Ivan M. Todorović, Milan Pecelj, Zoltan Čorba, Boris Dumnić, Dragan Milićević
University of Novi Sad, Faculty of Technical Sciences
Novi Sad, Serbia

***Abstract*—This paper addresses the question of the impact of the multiple battery charger stations on the local distribution grid power quality by investigating grid current and voltage spectrum. The case of station with multiple battery chargers of two types (mode 3 and mode 4) is considered. A Matlab/Simulink model has been developed and used for testing. The model parameters are selected in such a way to reflect real conditions in public grid. The results show that level of voltage harmonics is increased. Further on, the current harmonics are also high and above IEEE standard 519-1992 limits and IEC/EN 61000-3-4 levels. Still, as a result of harmonic cancelation and attenuation effect, the THDI is lower than in case of single charger operation.**

Keywords- multiple battery charger station, power quality, THD values, harmonics.

I. INTRODUCTION

Recent market studies foresee a huge rise of number of plug-in hybrid (PHEV) and plug-in electric (PEV) vehicles on the roads in this decade and further on [1, 2, 3, 4, 5]. Only in U.S.A. about one million of HEVs/EVs are expected by the end of 2015 (about 400,000 new HEVs/EVs to be sold in 2015) [5]. Globally, by the end of the next year around 2,000,000 will be on the streets and roads [3]. By 2020 20 million are predicted worldwide, and by 2050 more than 60% of all vehicles in U.S. will be HEV/EV.

One of the main challenges in realization of this ambitious plan is providing the drivers adequate energy supply infrastructure. Such infrastructure needs to be built, to be available in wide areas and with services similar to existing gas stations. However, not all services are needed. The reasons are in different characteristics, which are actually advantages of battery charging stations (BChS) over the gas pumps, such as:

- no need for underground storage tanks,
- environmentally friendly - no risk of pollution, gasoline leaking and similar accidents,
- self-service is simple,
- charging may be easily controlled,
- billing may be simplified using bank cards,
- no need for regular stuff and
- easy for maintenance and distance control.

Still, there are some items to be taken care of, such as: danger of electric shock, EM radiation, different standards for plugs and sockets, safe placement of station-to-vehicle connecting cables (which may be heavy and bulky), anti-vandalism strategy and measures for prevention of stealing or robbing the station. Modern realizations provide solution to the most of these issues.

Another very important issue is connection with public or private grid. One aspect is electrical energy availability in electric power system and reliability of the supply. Many researchers have studied such problems. Different operating strategies, charging models and scheduling have been proposed. [6, 7, 8, 9, 10]. Further on, a possibility of EV´s application as electrical storage systems and grid support (V2G) has been discussed, too [11, 12, 13].

The next, very important aspect is quality of supply and interaction of the BChS with power system. The battery chargers are known as non-linear loads, which distort input current and produce harmonic "pollution" in the public grid. A lot of papers have addressed this problem regarding single BChS operation and its impact on the grid [14, 15, 16]. It has been reported that total harmonics distortion of input current is between 60% and 70% if uncontrolled AC/DC converters are used [14].

In this paper, impact of multiple battery charging stations (MBChS) on power quality in the public grid is considered. To investigate this problem, a computer model of single charger is used [17]. The model represent the most common modern charger topology, actually a train of power electronics converters applied for charging purposes [18]. This charger differs from the one used by other authors, when a cluster of simple chargers has been considered [19]. The MBChS is regarded as a load equipped with a number of modern battery chargers, which may operate simultaneously. The overall station model is developed and applied for investigation of the MBChS impact on power quality. The simulation results are shown and results are discussed and compared with existing power quality standards.

II. MODEL OF A MULTIPLE BATTERY CHARGING STATION

Single BChS is standardized in USA and EU in three or four different types depending on charging current (AC or DC) and power size (input voltage and current) and consequently on duration of charging time: chargers of Levels 1 to 3 (SAE J1772 standard, USA) or Modes 1 to 4 (IEC 61851-1 standard, EU) [20, 21]. Their main electrical characteristics and charging time are listed in Table I and Table II.

Nowadays, the BChSs, either as single or multiple units, are spreading fast in the Europe. Typical public station may be located on open road, or more likely in a city centre (business

and shopping areas) and industrial zones. Suitable locations are also university campus and shopping moles parking sites, while residential chargers are to be expected in a city suburbs and villages. A map presented on Fig. 1 shows current position and operation status of BChSs in Europe, according to [22]. Although the map may not be detailed enough and probably does not show all types of chargers (especially in big city areas), still it indicates that the public and high-power/fast chargers are in majority. To serve several customers simultaneously, a number of BChS are needed in one location. Therefore, in this paper an MBChS operating in the mode 3 (moderate speed charging) and the mode 4 (fast-charging) is considered.

TABLE I TYPE OF CHARGERS ACCORDING TO SAE J1772 [20]

Type	AC Charging	DC Charging
Level 1	120V, 1~, max 16A max 1.9 kW	200 – 450 V, max. 80A 19.2 kW
Charging time	12h - 14h	~20 min
Level 2	240V, 1~, max 80A max 19.2 kW	200 – 450 V, max. 200A 90 kW
Charging time	3h - 4h	~15 min
Level 3	*TBD* (3~)	*TBD* (200-600V, max 400A max 240kW)

TABLE II TYPE OF CHARGERS ACCORDING TO IEC 61851-1 [21]

Type	AC Charging 1 phase	AC Charging 3 phase	DC Charging
Mode 1	230V, max 16A	NA	NA
Charging time	Slow (8 - 12h)	NA	NA
Mode 2	NA	400V, max 16A	NA
Charging time	NA	Slow (8 - 12h)	NA
Mode 3	230V, max 32A	400V, max 32A	NA
Charging time	Moderate (2 – 4h)	Moderate (1 – 2h)	NA
Mode 4	NA	400V, max 63A	600V, 400A
Charging time	NA	Fast (20-30 min)	Fast (15-20 min)

Figure 1. Available battery charging stations in Europe (Oct.2014) [22]

A modern BChS consists of several power electronics converters to enable adequate power processing, control and galvanic isolation. Fig.2 shows a three-phase AC topology applied in this paper, which enables mode 3 or mode 4 operations.

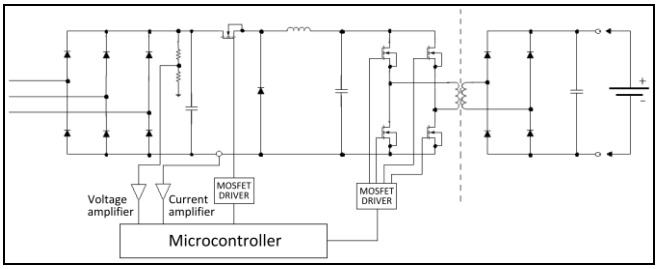


Figure 2. Topology of the modelled BChS.

The model is configured in such a way to reflect a real condition in the field, as much as it is possible. Standard distribution transformer of rated power 630 kVA and the transfer ratio of 10kV/0.4kV is used. The charging station is connected to LV grid, 0.4 kV and the connecting cable parameters are taken into consideration. Input AC/DC converter is an uncontrolled three-phase bridge diode rectifier, which is followed by buck-converter for the DC voltage levelling and a full-bridge DC/DC unit to provide galvanic isolation and output control. Output of such a converter is rated to 450 V. Taking into account the type of the rectifier and DC/DC converter used it is obvious that only one direction of energy is assumed in this paper, i.e. the possibility that the batteries are used as the temporary energy bank for distribution grid is excluded. At the end a lithium-ion battery was chosen. More details of the model are explained in [17].

Figure 3 depicts organization and content of the MBChS, which is modeled in the paper. It can be seen that it consists of six chargers – two fast and four moderate ones. It also shows how chargers can be connected to the distribution network. By turning on/off various switches, different chargers could be connected and put into operation. There are a number of combinations of possible usage of these chargers. All of them are modeled, but results of 14 of them will be presented here. For easier understanding, these cases are labeled as: 0F1M meaning zero fast and one moderate charger are connected (all switches, except switch 6 are open) or 2F4M meaning two fast and 4 moderate chargers are in operation (all switches are closed).

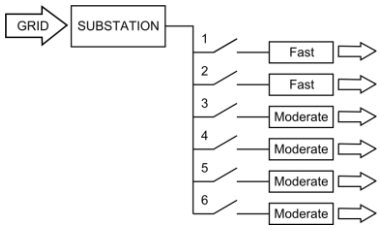


Figure 3. The MBChS - Energy flow from grid to batteries.

III. SIMULATION RESULTS

The Matlab/Simulink model of the MBChS is used to track input current and voltage harmonics for all 14 cases of station operation. Results are presented in Tables III and IV.

TABLE III INDIVIDUAL AND TOTAL HARMONIC DISTORTION OF INPUT VOLTAGE FOR DIFFERENT CASES OF CHARGER OPERATION

State of chargers	HDU [%] for harmonic order									THDU [%]
	3	5	7	9	11	13	15	17	19	
0F1M	0.42	0.89	1.99	0.15	0.65	0.51	0.08	0.55	0.52	2.87
0F2M	0.7	0.87	1.97	0.24	1.44	0.85	0.14	1.17	0.95	3.80
0F3M	0.73	1.67	1.96	0.47	1.98	1.07	0.33	1.53	1.16	4.75
0F4M	0.87	2.47	1.96	0.75	2.41	1.21	0.53	1.76	1.27	5.55
1F0M	0.2	0.72	1.53	0.14	0.97	0.74	0.12	0.86	0.79	3.05
1F1M	0.4	0.66	1.81	0.14	1.26	1.02	0.06	1.03	1.12	3.59
1F2M	0.49	1.09	1.89	0.15	1.62	1.21	0.03	1.26	1.33	4.16
1F3M	0.72	1.59	1.91	0.13	1.99	1.39	0.03	1.51	1.48	4.77
1F4M	0.78	2.09	1.87	0.33	2.28	1.44	0.18	1.7	1.51	5.29
2F0M	0.29	1.17	1.63	0.13	1.9	1.52	0.01	1.46	1.61	4.53
2F1M	0.5	1.75	1.82	0.12	2.16	1.7	0.06	1.57	1.78	5.05
2F2M	0.64	2.43	1.82	0.23	2.59	1.83	0.07	1.84	1.79	5.87
2F3M	0.66	2.88	1.81	0.38	2.84	1.84	0.18	1.98	1.74	6.29
2F4M	0.77	3.28	1.88	0.38	2.97	1.94	0.17	2.01	1.77	6.56

TABLE IV INDIVIDUAL AND TOTAL HARMONIC DISTORTION OF INPUT CURRENT FOR DIFFERENT CASES OF CHARGER OPERATION

State of chargers	HDI [%] for harmonic order									THDI [%]
	3	5	7	9	11	13	15	17	19	
0F1M	18.3	43.51	17.4	2.7	10.46	6.75	0.61	5.62	4.8	52.89
0F2M	21.13	43.95	15.59	1.81	11.62	6.13	0.36	6.03	4.64	53.82
0F3M	14.06	38.16	12.33	2.4	10.37	4.87	0.93	5.2	3.66	44.92
0F4M	11.95	33.65	9.92	3.38	9.0	3.86	1.47	4.28	2.8	39.10
1F0M	6.44	28.41	12.99	1.27	9.39	6.33	0.52	5.33	4.54	35.12
1F1M	7.59	30.64	12.49	1.16	9.07	5.86	0.25	4.81	4.38	36.58
1F2M	10.93	33.23	12.99	1.2	9.59	5.78	0.2	4.92	4.28	39.80
1F3M	11.7	33.9	12.69	1.0	9.58	5.61	0.16	4.73	4.12	40.36
1F4M	12.21	34.47	12.21	1.47	9.74	5.2	0.42	4.7	3.68	40.76
2F0M	5.33	27.4	12.04	0.94	8.6	5.66	0.24	4.36	4.13	32.94
2F1M	5.89	28.16	11.67	0.86	8.37	5.38	0.15	3.99	3.86	33.31
2F2M	9.2	30.79	12.13	1.03	9.04	5.29	0.25	4.22	3.62	36.61
2F3M	8.28	30.68	11.38	1.32	8.75	4.91	0.4	3.98	3.26	35.79
2F4M	9.31	31.0	11.19	1.53	8.79	4.72	0.51	3.87	2.98	36.18

Due to limited space, only the two of them will be regarded here, which represents two extremes – minimum and maximum charger loads. These two cases in comparison may show the main trend for voltage, current and power quality for all other cases. In this paper two types of MBChS operation are simulated:

1. Single unit in operation (switch 6 closed) - 3 phase AC, 0.4 kV, Mode 3 (moderate charging speed) charger and
2. All units in operation (all switches closed) - 3 phase AC, 0.4 kV, Mode 4 (fast charging) charger.

Grid voltage, current and harmonic content of both voltage and current have been observed and calculated. Results are presented in Figs. 4-9.

A. Simulation results in case of one charger in mode 3 operation

Figs. 4, 5 and 6 show simulation results obtained for a single battery charger station operated as mode 3 charger and connected to the substation (switch 6 closed). Fig. 4 shows line voltage, phase voltage and phase current at the charging station input terminals. Fig. 5 shows harmonic spectrum of the one phase voltage, while Fig. 6 shows harmonic spectrum of the phase (line) current, both at the MBChS input terminals.

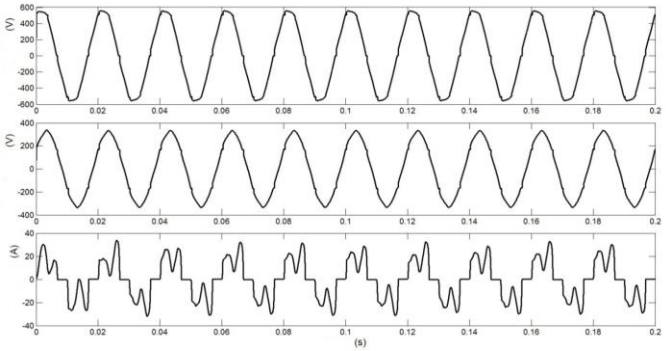


Figure 4. Line voltage, phase voltage and phase current at the charging station input terminals (one charger connected - mode 3).

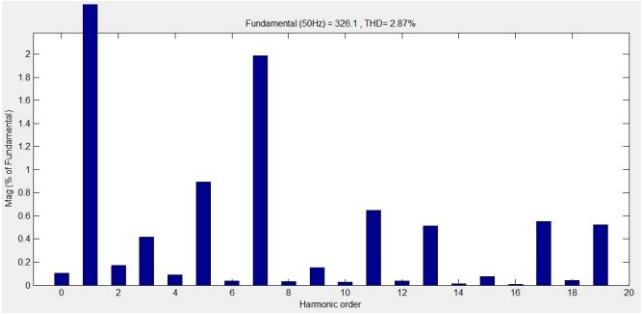


Figure 5. Phase voltage harmonic spectrum at the charging station input terminals (one charger connected - mode 3).

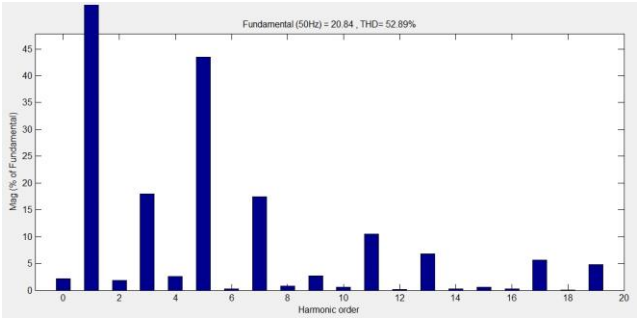


Figure 6. Phase current harmonic spectrum at the charging station input terminals (one charger connected - mode 3).

B. Simulation results in case of all chargers in operation

Figs. 7, 8 and 9 present simulation results obtained for simultaneous operation of all battery chargers (all switches closed), both mode 3 (moderate charging time) and mode 4 (fast charging time), and connected to the substation. Fig. 7 shows line voltage, phase voltage and phase current at the MBChS input terminals. Fig. 8 shows harmonic spectrum of the phase voltage, while Fig. 9 shows harmonic spectrum of the phase (line) current, both at the MBChS input terminals.

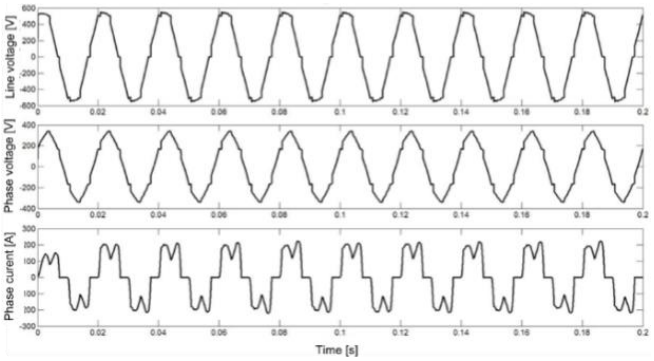


Figure7. Line voltage, phase voltage and phase current at the MBChS input terminals (all chargers connected).

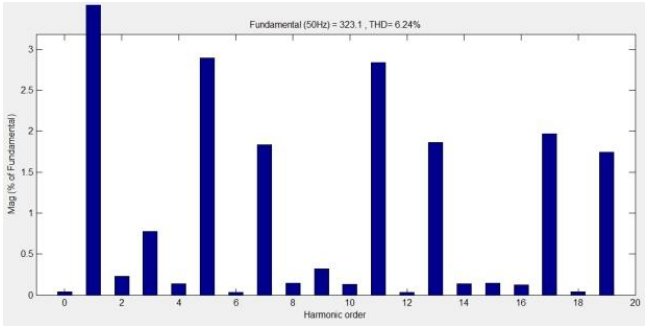


Figure 8. Phase voltage harmonic spectrum at the MBChS input terminals (all chargers connected).

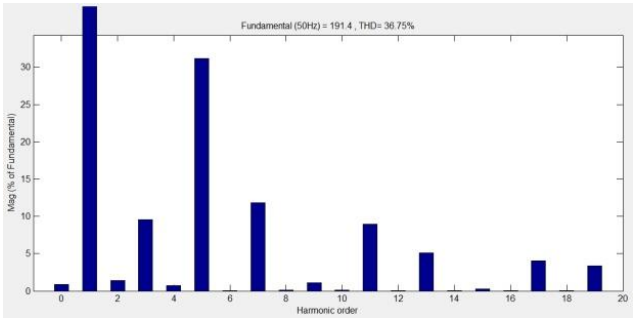


Figure 9. Phase current harmonic spectrum at the MBChS input terminals (all chargers connected).

IV. DISCUSSION

A. Harmonic compatibility levels

For comparison purposes, the IEEE Standard 519 and IEC 61000/EN61000 standard series are considered.

By IEEE STD 519 standard voltage THD for MV systems less than 60 kV must be under 5% and all individual harmonics should be fewer than 3% [23, 24]. Fig. 10 shows limits stipulated by IEC 61000-2-4 standard for voltage THD and

MV systems. It can be seen that the THDU is limited to 8% for Class 2 loads [23, 25].

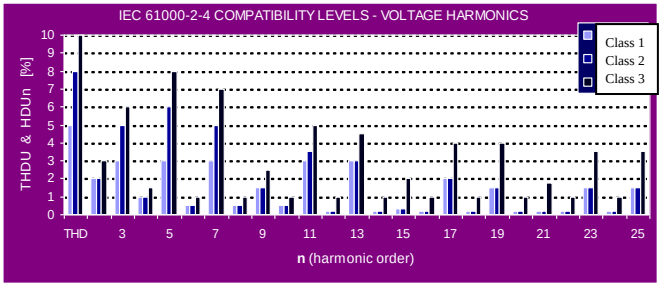


Figure 10. IEC 61000-2-4 Compatibility levels for MV voltage harmonics

The level of current harmonics limits depends on ratio of short-circuit and load current (IEEE standard 519-1992) or short-circuit and load power (IEC 61000-3-4) at the point of common coupling of a harmonic producing device. Table V and VI shows the limits levels stipulated in IEEE standard 519-1992 and IEC 61000-2-4, respectively.

TABLE V IEEE 519-1992 - CURRENT HARMONICS LIMITS

	HD In – ONLY ODD (%)					THDI (%)
I_{SC}/I_L	n<11	11<n<17	17<n<23	23<n<35	n>35	
<20*	4,0	2,0	1,5	0,6	0,3	5
20-50	7,0	3,5	2,5	1,0	0,5	8
50-100	10,0	4,5	4,0	1,5	0,7	12
100-1000	12,0	5,5	5,0	2,0	1,0	15
>1000	15,0	7,0	6,0	2,5	1,4	20

- Even harmonics- 25% of odd harmonics limits.
- * All generators are limited on above values regardless on actual value of I_{SC}/I_L
- Legend: I_{SC} = Maximum short circuit current at PCC
I_L = Maximum load current (fundamental harmonic)
- For PSSs from 69 kV up to 138 kV limits are 50% of above values.
- For PCCs above 138 kV case by case estimation is required.

TABLE VI IEC/EN 61000-3-4 CURRENT HARMONICS LIMITS (STAGE 2)

	HDIn – ODD HARMONICS ONLY [%]					THDI [%]
R_{SCE}= S_{SC}/S_L	n=5	n=7	n=11	n=13	n≥15	non-sim. / 3~ simetr.
<66	14	11	10	8	/	25 / 16
120	16	12	11	8	/	29 / 18
250	30	18	13	8	/	39 / 35
450	50	35	20	15	/	51 / 58
>600	60	40	25	18	/	- / 70
- Even harmonics are limited to: I_n/I_1 = 16/n						
- Harmonics multiply of 3 are not present						
- Assumption is that the load is balanced						

B. *Single charger, mode 3 operation*

In the case of single charger, mode 3 operation of the battery charging station results show a low voltage distortion and a high current one.

Total harmonic distortion of phase voltage is THDU=2.87%, while distortion of individual harmonics is less than 1% (Fig. 5). Both distortions are well below the limits of IEEE STD 519 standard. If IEC 61000-2-4 standard is considered, it is clear that THDU is below compatibility levels, as well as individual harmonics (all odd harmonics are below 1%). As MV grid was assumed with 2% of existing voltage harmonics distortion, it may be concluded that the impact of single charger operation regarding voltage harmonics is in line with existing standards.

On the other hand, Fig. 6 shows that current harmonics are high. Total harmonic distortion is THDI=52.89%, while individual ones lie between 5% and 43% for odd harmonics. By the IEEE STD 519 standard (Table V), current harmonics are above limits. All odd harmonics under 11th should be less than 12% for systems with short circuit/maximum load (first harmonic) current ratio between 100 and 1000 which is here the case. Thus, current harmonics are well above the desired level. The biggest problems are the 5th, 7th, 9th and 11th harmonics. If IEC 61000-3-4 standard is considered (Table VI), it can be seen that limits are satisfied only for ratio of short circuit power and load power 450 and above.

C. *All chargers in operation (Mode 3 and Mode 4)*

In the case of simultaneous operation of all battery chargers in the MBChS (mode 3 and mode 4), results show relatively low voltage distortion, while the current one is high.

Voltage distortion in this case is significantly higher than in previous case when only one of the mode 3 chargers were in operation. The THDU is now up to 6.24%, which is above IEEE Standard 519-1992 limit. Also, individual harmonics are high, close to the stipulated values, especially the 5th and the 11th ones (HDU_5=2.8%, HDU_{11}=2.7%). If IEC 61000-3-4 is considered, the conclusions are a little bit different – THDU and all individual harmonics are inside the limits, except the 19th. Therefore, in this case we may conclude that MBChS operation will result in high voltage distortion, especially if background voltage distortion of 2% is already present in the system.

Current harmonics are lower than in previous case, but still high. The results show that total harmonic distortion is THDI=36.75% and that individual current harmonics are between 3% and 11%. Again, they are higher than IEEE 519 limits in all cases, but acceptable for IEC 61000-3-4 if ratio of short circuit power and load power is 250 and above.

The obtained result of decreasing of the THDI in case of a number of chargers operating at the same time is similar to effects of a PC cluster operation. It was shown that if a large number of PCs are operating simultaneously, the THDI level is significantly lower, than in case of single PC operation. This phenomenon is explained with harmonic cancelation and attenuation effects. The THDI vs number of PCs relation has been proposed [26].

V. CONCLUSION

The multiple battery charger station is analyzed regarding possible impact on power quality in public grid. The case of station with multiple battery chargers of two types (mode 3 and

mode 4) is considered. Effects of their operation on power quality are investigated. A Matlab/Simulink simulation model has been developed and used for testing. The model parameters are selected in such a way to reflect real conditions in public grid. The results show that in case of all chargers simultaneous operation voltage harmonics are increased – they surpassed the IEEE standard 519-1992 compatibility levels and even some of the IEC/EN 61000-3-2 standard ones (the 19th harmonic). Further on, the current harmonics are also high and above IEEE standard 519-1992 limits and IEC/EN 61000-3-4 levels if ratio of short circuit power and load power is less than 250. Still, as a result of harmonic cancelation and attenuation effect, the THDI is lower than in case of single charger operation.

ACKNOWLEDGMENT

This research was partially co-funded by the by the Provincial Secretariat for Science and Technological Development of AP Vojvodina under contract No. 114-451-2248/2011-03 „Research and Development of Energy Efficient Power Supply and Propulsion Systems of Electric Vehicles".

REFERENCES

[1] B. Berman, J. Gartner, "Plug-in Electric Vehicles Battery Electric and Plug-in Hybrid Electric Vehicles: OEM Strategies, Demand Drivers, Technology Issues, Key Industry Players, and Global Market Forecasts", Pike Research LLC, Boulder (USA), 2012. http://www.navigantresearch.com/wp-content/uploads/2012/06/PEV-12-Executive-Summary.pdf

[2] C. Zhu, N. Nigro, "Plug-In Electric Vehicle Deployment In The Northeast - A Market Overview and Literature Review", Transportation and Climate Initiative, Georgetown Climate Centre, and New York State Energy Research and Development Authority, Georgetown (USA), Sep. 2012. http://www.c2es.org/docUploads/pev-northeast.pdf

[3] ***, "Global EV Outlook - Understanding the Electric Vehicle Landscape to 2020", Clean Energy Ministerial - Electric Vehicles Initiative - International Energy Agency, OECD/IEA, Paris (France) April 2013, http://www.iea.org/publications/globalevoutlook_2013.pdf

[4] M. Contestabile, G. Offer, R. North, "Electric Vehicles: A Synthesis of the Current Literature with a Focus on Economic and Environmental Viability", LCA Works, London (Great Britain), June 2012. http://www.lcaworks.com/EV%20Lit%20Rev%20FINAL.pdf

[5] U.S. Department of Energy, "One Million Electric Vehicles by 2015", Status Report, Feb.2011. http://www1.eere.energy.gov/vehiclesandfuels/pdfs/1_million_electric_vehicles_rpt.pdf

[6] R. Garcia-Valle, J.A. Peças Lopes, Editors, "Electric Vehicle Integration into Modern Power Networks", Springer, New York, 2013.

[7] H. Turker, S. Bacha, D. Chatroux, "Impact of Plug-in Hybrid Electric Vehicles (PHEVs) on the French Electric Grid", Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 11-13 Oct. 2010, Gothenburg (Sweden), DOI: 10.1109/ISGTEUROPE.2010.5638948

[8] Y. Zhou A. Vyas, "Keeping plug-in electric vehicles connected to the grid - Patterns of vehicle use", IEEE PES Innovative Smart Grid Technologies (ISGT), 2012 Washington, Jan.2012, DOI: 10.1109/ISGT.2012.6175805

[9] M. Alonso, H. Amaris, J.G. Germain, J.M. Galan, "Optimal Charging Scheduling of Electrical Vehicles in Smart Grids by Heuristic Algorithms", Energies, Vol. 7, April 2014, pp.2449-2475; DOI: 10.3390/en7042449

[10] L.E. Bremermann, M. Matos, J.A. Pecas Lopes, M. Rosa, "Electric vehicle models for evaluating the security of supply", Electric Power Systems Research, Vol.111, June 2014, pp.32–39, DOI: 10.1016/j.epsr.2014.02.001

[11] J. Tomic, W. Kempton, Using fleets of electric-drive vehicles for grid support, Journal of Power Sources, Vol. 168, Issue 2, June 2007, pp.459–468. DOI: 10.1016/j.jpowsour.2007.03.010

[12] W. Kramer, S. Chakraborty, B. Kroposki, A. Hoke, G. Martin, T. Markel, "Grid Interconnection and Performance Testing Procedures for Vehicle-To-Grid (V2G) Power Electronics", World Renewable Energy Forum 2012, Denver, Colorado, May 13-17, 2012.

[13] E. Sortomme, M.A. El-Sharkawi, Optimal scheduling of vehicle-to-grid energy and ancillary services, IEEE Trans. on Smart Grid, Vol.3, Issue 1, March 2012, pp.351–359. DOI: 10.1109/TSG.2011.2164099

[14] J.C. Gómez, M.M. Morcos, "Impact of EV Battery Chargers on the Power Quality of Distribution Systems", IEEE Trans. on Power Delivery, vol. 18, no. 3, pp. 975-981, July 2003.

[15] C. Jiang, R. Torquato, D. Salles, W. Xu, "Method to Assess the Power-Quality Impact of Plug-in Electric Vehicles", IEEE Trans. on Power Delivery, Vol. 29, No. 2, April 2014, pp.958-965.

[16] P.S. Moses, S. Deilami, A.S. Masoum, M.A.S. Masoum "Power Quality of Smart Grids with Plug-in Electric Vehicles Considering Battery Charging Profile", Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 11-13 Oct. 2010, Gothenburg (Sweden), DOI: 10.1109/ISGTEUROPE.2010.5638983

[17] V.A. Katić, M. Pecelj, I. Todorović, "Effects of Individual Battery Charger Station on Power Quality", 10th International Symposium on Industrial Electronics – INDEL 2014, Nov. 6-8, 2014, Banja Luka (Bosnia and Herzegovina).

[18] M.Yilmaz, P. Krain, "Review of Battery Charger Topologies, Charging Power levels, and Infrastructure for Plug-In Electric and Hybrid Vehicles", IEEE Trans. on Power Electronics, Vol.28, No.5., May 2013, pp. 2151-2169.

[19] J.A. Orr, A.E. Emanuel, K.W. Oberg, "Current Harmonics Generated by a Cluster of Electric Vehicles Battery Chargers," IEEE Trans. Power Apparatus Systems, Vol. PAS-101, No. 3, pp.691–700, Mar. 1982.

[20] SAE International, "SAE Electric Vehicle and Plug in Hybrid Electric Vehicle Conductive Charge Coupler", J1772-201210, 2012.

[21] IEC, "Electric Vehicle Conductive Charging System - Part 1: General Requirements", 2.0 Edn., IEC 61851-1, Geneva, 2010.

[22] www.plugshare.com

[23] V.A. Katić, "Power Quality – Harmonics", Edition: Engineering Books – Monographs, No.6, University of Novi Sad - Faculty Technical Sciences, Novi Sad, 2002, (in Serbian).

[24] IEEE Std 519-1992, "IEEE recommended practices and requirements for harmonic control in electrical power systems", IEEE Press, 1993.

[25] IEC/TR EN 61000-2-4 standard " Electromagnetic compatibility (EMC) - Part 2-4: Environment - Compatibility levels in industrial plants for low-frequency conducted disturbances", IEC, Geneva, 2002.

[26] S. Mujović, V.A. Katić, J. Radović, "Improved Analytical Expression for Calculating Total Harmonic Distortion of PC Clusters", *Electric Power Systems Research*, Vol. 81, No.7, July 2011, pp.1317-1324, [Online]. Available: http://dx.doi.org/10.1016/j.epsr.2011.01.023

Reference Analysis of the Analogous Models for Photovoltaic Cells by Comparison with the Real Photovoltaic Modules

Duško Lukač
Rheinische Fachhochschule Köln gGmbH
University of Applied Sciences
Vogelsanger Strasse 295, 50825 Köln, Deutschland
e-mail: lukac@rfh-koeln.de

Miona Andrejević Stošović, Dragiša Milovanović and Vančo Litovski
University of Niš, Faculty of Electronic Engineering
Aleksandra Medvedeva 14, 18000 Niš, Serbia

***Abstract*— This paper is based on the research project between Rheinischen Fachhochschule in Köln - University of Applied Sciences, Germany, and the University of Niš, Serbia in regard of modeling of photovoltaic cells of different technologies. Standard models used for modeling of photovoltaic cells give partially satisfactory results, which are mostly depending on the physical-mathematical model and on the technology of the photovoltaic cells used. In this work we compare simulation results based on simplified diode model and one-diode model with the characteristics of three different real photovoltaic modules of different cell type based on different cell technology. We show the dependency of accuracy of the physical models by the choice of ideality factor and reverse saturation current, which at least often led to calculation of different efficiency factors of PV cells. Thus, the calculated data given by manufacturer can distinguish significantly in regard of cell efficiency compared with the measured data. By varying of ideality factor and reverse saturation current, the curve course of the simulated characteristic curve can be adjusted to show a very good correspondence with the curve course of the real cell. Furthermore, we show the differences in the quality factor according to the model used and give a reference in conclusion to the limits of the current based on different cell technology.**

Keywords — Modeling, Photovoltaic cells, Simulation component.

I. INTRODUCTION

Within the scope of the common research project between the Rheinischen Fachhochschule in Köln - University of Applied Sciences, Germany as well as the University of Niš, Serbia the analogous model of the current photovoltaic technologies, in particular the silicon-based mono-crystalline as well as polycrystalline solar cells are investigated in regard of their exactness and compared with the real photovoltaic (PV) cells. In previous publications the topographic influence as well as the inverter-feedback on the PV cells were presented, which up to now were not taken into consideration in the current diode models [1, 2, 3, 4]. In order to test the so called diode models on a real system and show the dependency of the physical-mathematical models on the variable parameters depending on the cell type, as well as on saturation current and ideality factor, technical PV cells data of the manufacturer must be once investigated. In order to estimate the manufacturer's data of the PV cells special precision stamped measuring instruments are to be used to validate the basic manufacturer's data about the used PV cells. Furthermore simulation applications, which allow the variation of the saturation current and ideality factor, can be used to draw a comparison between the own developed model and a real PV cell. In this publication simulation application for a test of the simplified diode model as well as one-diode model are presented. By creating the PV characteristic curves it is possible to get values about the nominal power, open-circuit voltage and short-circuit current of the module which are calculated under the so called standard test conditions (STC). By STC one understands the performance of the PV module with an irradiation of 1000 W / m² at module level with a certain spectrum of the light. Furthermore the temperature may amount only 25˚C. By the STC standardization of all measuring data, also different modules can be compared with each other by using different irradiation and temperature conditions and relations. In this paper we use symbols according to IEC 60617-2:1996 and EN 60617-2:1997. Symbols used in Figure 4 and 8 for diodes are symbol 05-03-01 out from the EN-60617-range and for ideal current source symbol 02-16-01 out from the EN-60617-range.

II. MEASUREMENT PRINCIPLES AND REQUIREMENTS

For the measurement of the current and voltage characteristic curves a module is loaded with a defined electric load and therefore the voltage and the electrical current are measured as a value pair, transfer analogy on the y and x axis, defining the I/U PV model. During the measurement, the environmental parameters play an essential role. The condition of the environment is to be documented inevitably during the measurement, because it is difficult to catch a few moments in the year (at least in central Europe) when optimum measurement conditions exist. An alternative exists by using of the so called solar spectral simulators. In order to make a projection from the environmental conditions on the STC terms, the irradiation for the most PV instruments, should amount to at least 700 W/m². Due to the respective measuring technology, environmental temperature, irradiation angle and spectral responsivity of the sensor, the measurements of the irradiation strength can become different. Even the determination of the exact PV module temperature is connected

with bigger problems because the relevant temperature is not measurable in the solar cell. The simplest, still reliable method is the use of an exact external temperature sensor, which is appropriated to the back side of the PV module. In this work we will firstly present how to compare the PV diode models with the real PV cells of a certain manufacturer device by using the simulations application. In further work we will explain the development of the real test device for the measurement of the specific PV data required by the model of the PV cells.

A. Characteristic curves of three real PV modules

The modeling of the characteristic curve of PV-modules with the help of analogous models is a usual and proven method [5, 6, 7]. To simulate the PV modules, the simulation application "PV-Teach" was used [8]. PV-Teach is simulation tool which presents the possibilities to use the single analogous models by varying of diverse parameters. The program allows also the import of the real module characteristics and generation of the real curves according to data sheet of the manufacturer. Also it allows simulation by using of analogous models. First of all technical data of real PV modules is read in, in order to construct the characteristic curves. Then, comparison of the real PV module with one of the PV analogues models as e.g. simplistic model, one-diode model, two-diode model or the model of the effectual characteristic curve can be carried out. In addition, an optimization of the saturation current and ideality factor parameters takes place. In this work the so called simplified diode model and one-diode model will be presented and compared using the PV-Teach. In order to test the PV models, characteristic curves of 3 different real PV cells are used. The used PV cells are: SW 165 Wp by Solarworld, Sunrise SRM 185 dp by Solartec and ATF 43 solar cell by Antec. The three modules have following main technical characteristics:

TABLE I. TECHNICAL DATA OF SOLARWORLD SW 165 WP

Solarworld SW 165 Wp	
Nominal power	165 Wp
Nominal voltage	35,4 V
Nominal current	4,70 A
Open-circuit voltage	43,3 V
Short-circuit current	5,10 A
Cell type	polycrystalline

By using of PV Teach following measured (real) characteristic curve is characterized for Solarworld SW 165 Wp module:

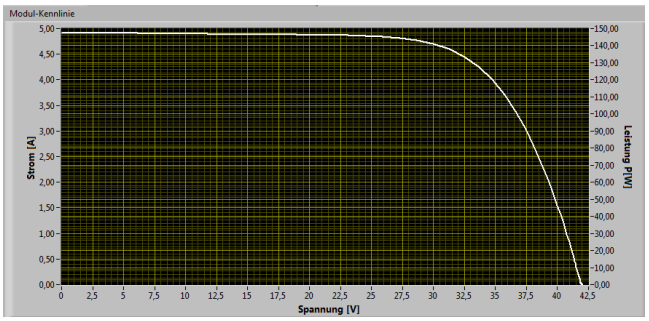


Fig. 1. Measured characteristic curve Solarworld SW 165

TABLE II. TECHNICAL DATA OF SUNRISE SRM-185

Solartech Sunrise SRM-185	
Nominal power	185 Wp
Nominal voltage	36,3 V
Nominal current	5,10 A
Open-circuit voltage	44,2 V
Short-circuit current	5,51 A
Cell type	monocrystalline

By using of PV Teach following characteristic curve is characterized for Solartech Sunrise SRM-185 module:

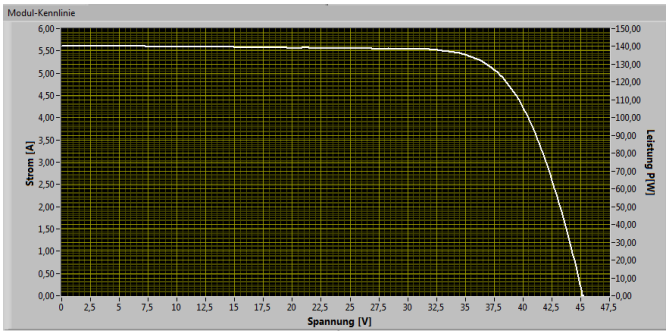


Fig. 2. Measured characteristic curve Sunrise SRM-185

TABLE III. TECHNICAL DATA OF MODUL ATF 43

Anatec Thin Layer Modul ATF 43	
Nominal power	43 Wp
Nominal voltage	53 V
Nominal current	0,81 A
Open-circuit voltage	81 V
Short-circuit current	1,07 A
Cell type	Cds/CdTe

By using PV Teach, following characteristic curve is characterized for Anatec Thin Layer Modul ATF 43 module:

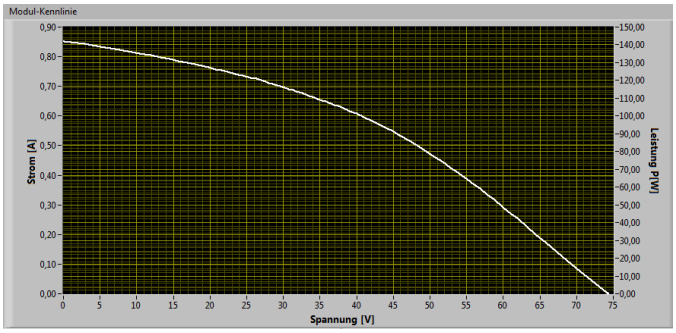


Fig. 3. Measured characteristic curve ATF 43

Physical simulation models describe the PV module mathematically or physically and offer a relation to the real component. An unirradiated PV module is nothing else than one large-area diode. Hence, this electric structural element can be described by the diode equation according to Shockley. In

addition, a numerical simulation model is described with the help of the model of the actual characteristic curve. With the empiric simulation model no relation to the real component exists. Despite of it, the model can deliver actual characteristic curve with very good results.

B. *Simplified Diode-Model vs. Simulation Model by using PV Teach*

The unlit solar cell is nothing else than one solid-state diode [5, 6]. It behaves physically in the similar way. By irradiation of the solar cell, so called photoelectric current occurs. The photoelectric current is referred to as I_{Ph}. With an equivalent circuit diagram from an ideal current generator and a diode, these relations can be exposed.

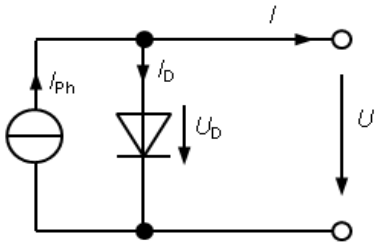


Fig. 4. Simplified diode model

The equation for the characteristic curve is given with:

$$I = I_{Ph} - I_D = I_{Ph} - I_s \cdot (e^{\frac{U}{m \cdot U_T}} - 1) \qquad (1)$$

I_S is so-called reverse saturation current according to the Shockley-Equation. It depends on diffusion voltage (junction built-in voltage), the length of the diffusion of electrons and holes, density of donator atoms, Richardson-Coefficient and the surface of PN-transition area. The *Is* is in the case of germanium based diodes *Is* ≈ 100nA and for silicon diodes $I_S \approx$ 10pA [5, 6, 7]. In order to reproduce better a real solar cell characteristic curves, one more factor, so called ideality factor *m* (also known as emission coefficient) is introduced in the exponent. Ideality factor *m* has value between 1 and 2 and it represents the measure of the divergence of the characteristic curve compared with the ideal diode character [7, 9]. In order to define the holistic solar cell characteristic curves, short-circuit current I_K and open-circuit voltage U_L need to be defined. Based on the simplified diode-model these are defined as follows:

$$I_K = I(U = 0) = I_{Ph} - I_s \cdot (e^0 - 1) = I_{Ph} \qquad (2)$$

$$U_L = U(I = 0) = m \cdot U_T \cdot \ln\left(\frac{I_K}{I_S} + 1\right) \qquad (3)$$

$$U_L = m \cdot U_T \cdot \ln\left(\frac{I_K}{I_S}\right) \text{ for very small currents} \qquad (4)$$

By using the simulations software PV-Teach the measured characteristic curves for PV modules SW 165 Wp by Solarworld, Sunrise SRM 185 dp by Solartec and ATF 43 solar cell by Antec will be compared with simulated simplified diode-model. Following starting values for real module characteristics are used (Table IV):

TABLE IV. REAL MODULE CHARACTERISTICS

	Solarworld SW 165 Wp	Solartech Sunrise SRM-185	Anatec Thin Layer Modul ATF 43
Photoelectric current	4,97 A	5,61 A	0,99 A
Saturation current	2,113 µA	1,287 nA	17,53 µA
Diode factor	2,87 V	2,04 V	7,05 V

The simulated characteristic curves (red line) compared with the original measured curves (white line) given by manufacturer are presented in the following figures:

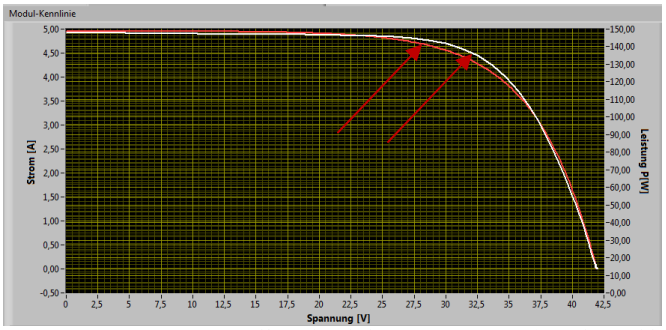


Fig. 5. Comparison: measured/simulated characteristic curve for Solarworld SW 165

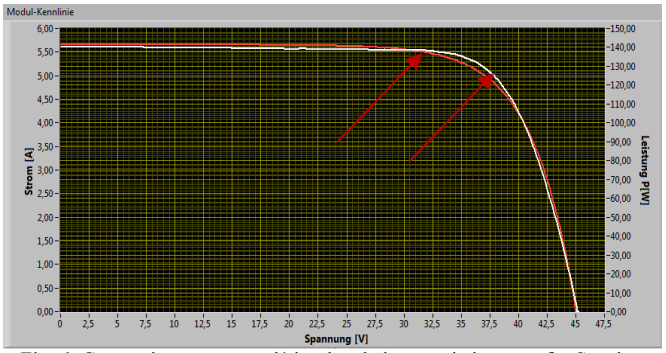


Fig. 6. Comparison: measured/simulated characteristic curve for Sunrise SRM-185

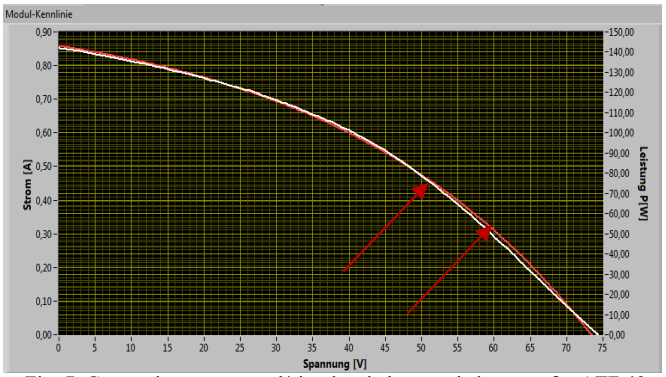


Fig. 7. Comparison: measured/simulated characteristic curve for ATF 43

The already mentioned parameters m and I_S are varied until the curve course of the simulated characteristic curve shows a very good correspondence with the curve course of the real cell. Indeed, it is obvious that even after optimization of the parameters still a clear divergence between measured characteristic curve and the simulated characteristic curve exists. It means that depending on the choice of the parameters, different efficiency data of the PV cell can arise. The achieved goodness (quality factor) amounts for the polycrystalline PV module of Solarworld to G = 98.33, for the mono-crystalline module Solartec the achieved goodness is G = 98.39, and at least the goodness for the thin layer-module ATF 43 amounts to G = 98.62. The divergence or the inaccuracy occurs because the simplified diode-model is only an idealized model for presentation of the diode characteristics [10].

C. One Diode-Model vs. Simulation Model by using PV Teach

For exact consideration of the electrical losses in the solar cell, it is inevitable to connect the ideal diode with farther components which describe the deviations from the reality. One possible analogous model of a real solar cell is shown with the so-called one-diode model.

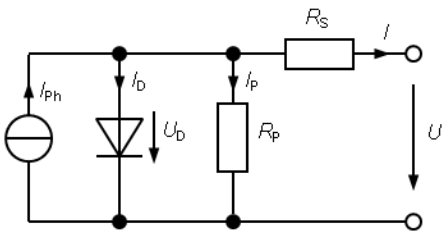


Fig.8. One-diode model of the solar cell

Compared with the simplified diode model, the one-diode model includes standard serial resistance R_S which describes the ohm losses in the front contacts of the solar cell and in the metal semiconductor and parallel resistance R_P which describes the leakage currents. I_D is the saturation current which exists because of the minority bearers available in the barrier layer. The equation which describes the one-diode model is given with:

$$I = I_{Ph} - I_s \cdot (e^{\frac{U+I\cdot R_S}{m\cdot U_T}} - 1) - \frac{U + I \cdot R_S}{R_P} \quad (5)$$

The parameters in one-diode model must be optimized in that way, that the model shows the almost identical electric behavior as the real PV cell. Theoretically it concerns a more dimensional non-linear optimization problem and for it the Levenberg Marquardt algorithm for the optimization can be used. The parameters are varied by the algorithm, as long until the difference of the real and simulated characteristic curve reached the smallest value. To solve the characteristic curve equation a numerical procedure must be used like the Newtonian iterative method, because the one-diode model owns no explicit solution [10]. This means that the following condition $f(I, U)$=0 must be defined. The iteration is described with:

$$I_{i+1} = I_i - \frac{f(I,U)}{\frac{\partial f(I,U)}{\partial I}} \quad (6)$$

The iteration is carried out until $|I_i - I_{i+1}| < \sigma$. Thereby I_i is calculated current in the iteration step i, I_{i+1} is calculated current in the iteration step i+1 and σ is lower barrier of the terminating condition of iteration. For the function $f(I,U)$ applies:

$$f(I,U) = I_{Ph} - I_s \cdot e^{\frac{U}{m\cdot U_T}} \cdot e^{\frac{I\cdot R_S}{m\cdot U_T}} + I_S - \frac{U}{R_P} - \frac{I \cdot R_S}{R_P} - I = 0 \quad (7)$$

And for the derivation applies:

$$\frac{\partial f(I,U)}{\partial I} = -\frac{I_S \cdot R_S}{m \cdot U_T} \cdot e^{\frac{U+R_S}{m\cdot U_T}} - \frac{R_S}{R_P} - 1 \quad (8)$$

The influence of standard serial resistance R_S on the solar characteristic curve is presented in the following Fig. 9:

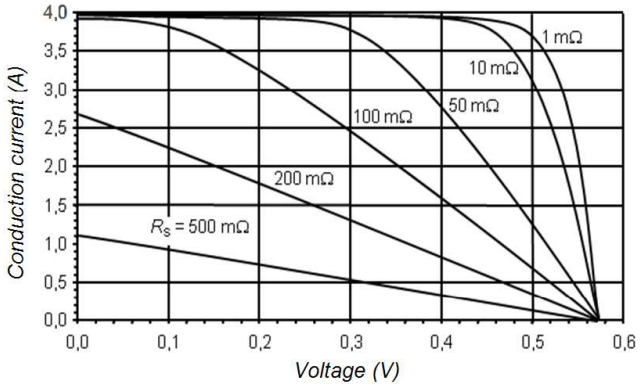


Fig.9. Influence of the standard serial resistance R_S

If the value of R_S rises, the curve flattens and the fill factor - the ratio of a photovoltaic cell's actual power to its maximal power if both current and voltage are at their maxima, gets much smaller. Influence of resistance R_P on the solar characteristic curve is given:

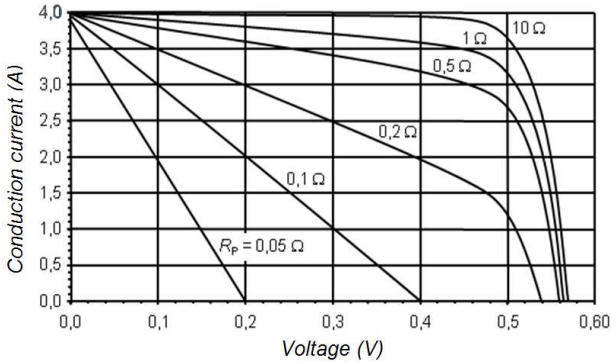


Fig.10. Influence of the parallel resistance R_P

By the falling of the value of the parallel resistance R_P the situation behaves similarly. The rising parallel current I_P lets the diode voltage descend and influences even the open-circuit voltage. In the following the measured characteristic curves are

compared with simulated for one-diode models. Following starting values are used (Table V):

TABLE V. REAL MODULE CHARACTERISTICS

	Solarworld SW 165 Wp	Solartech Sunrise SRM-185	Anatec Thin Layer Modul ATF 43
Photoelectric current	4,97 A	5,61 A	0,99 A
Saturation current	2,113 µA	1,287 nA	17,53 µA
Series resistance	0,55 Ω	0,53 Ω	34,70 Ω
Parallel resistance	270,25 Ω	702,21 Ω	213,84 Ω
Diode factor	2,87 V	2,04 V	7,05 V

The simulated characteristic curves by using the one diode-model compared with the original measured curves are presented in following (Figs. 11, 12 and 13):

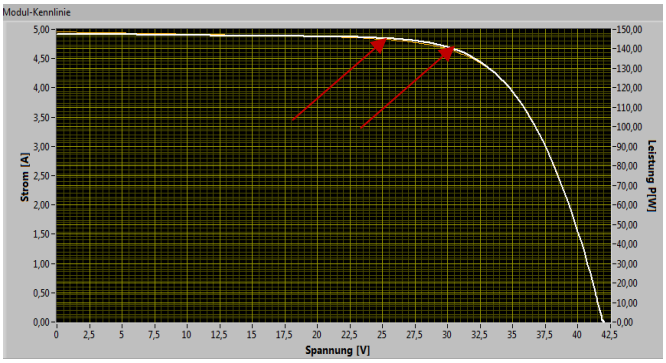


Fig.11. Comparison: measured/simulated characteristic curve for Solarworld SW 165

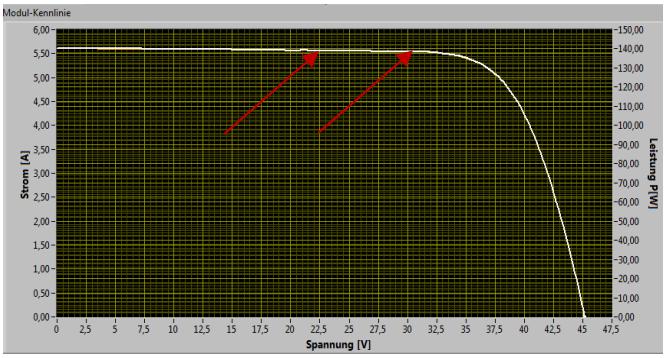


Fig.12. Comparison: measured/simulated characteristic curve for Sunrise SRM-185

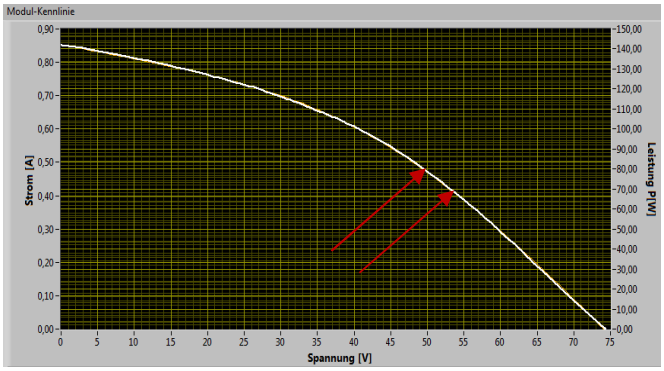


Fig.13. Comparison: measured/simulated characteristic curve for ATF 43

By using of the one-diode model a better approximation goodness of the simulated characteristic curves compared with the simplified diode-model is recognizable because the impact of the parameters saturation current and ideality factor is lower. Here a very high goodness is achieved. For the polycrystalline module of Solarworld the goodness amounts to $G = 99.57$, for the mono-crystalline module of Solartech the goodness amounts to $G = 99.86$ and for the thin layer module of Antec ATF 43 the goodness amounts to $G = 99.56$. With the extend of the simplified diode-model with the resistances R_S and R_P the losses are described in the solar cell and with it, the real behavior of the cell is more exactly expressed. Even if the existing analogues models offer good simulation results, many important aspects, which have impact on the function of the PV cell are not included in the models and try to be compensated with the ideality factor m. So the influence of the topography on the PV cells in the diode models or the influence of the inverter feed-back creating harmonic components influencing the behavior of the PV cells are still not included in the diode models [8, 11]. Therefore more precise and more universal models of PV cells need to be presented.

III. CONCLUSION

All simulation models with equivalent circuit diagrams try to reproduce the physical processes in the solar cell as good as possible. The notion of the modeling of the characteristic curves with equivalent circuit diagrams lies in the calculability of adaptation problems between photovoltaic-solar generators and consumer loads. Hence, certain requirements for the simulation method based on calculation are given and those can be an explicit calculation of the current-voltage characteristic curve equation or the explicit calculation of the characteristic curve equation parameters from the characteristic value. The simulation results presented in this work show the dependency of the simulation of the right choice of the saturation current and ideality factor of the cells, as well as the choice of the PV cell model. Also the models are of statically nature which is not including further parameters which are influencing the characteristic curves of solar cells. The diode models are further useful only for specific PV technologies as for instance silicon based amorphous silicon or CdTe technology. For more exact models and more universal models of PV cells, other more universal analogues models need to be developed.

REFERENCES

[1] M. Andrejević Stošović, D. Lukač, V. Litovski, "Quantification of Power Quality Issues at the PV Panel-ConverterInterface", INDEL Conference, Banja Luka, Bosnia-Herzegovina, pp. 256-262, November 2012.

[2] M. Andrejević Stošović, D. Lukač, V. Litovski, "A view to the modelling and simulation of the PV system - Converter interface", TECHNOPORT Renewable Energy Research Conference 2012, Trondheim, Norway, April 2012.

[3] D. Lukač, M. Andrejević-Stošović, V. Litovski, S. Vukomirović, "Technology and Light Source as Challenge for proper Modeling of PV Cells", XI International Scientific – Professional Symposium INFOTEH, Sarajevo, Bosnia and Herzegovina, pp. 173-176, March 2012.

[4] K. Mertens, “Photovoltaik- Lehrbuch zu Grundlagen, Technologie und Praxis”, PV-Teach, Avaliable <http://www.lehrbuch-photovoltaik.de/software/pv-teach.zip>, 28.12.12, 12:06h.

[5] K. Mertens, “Photovoltaik”, 2nd Edition, Fachbuchverlag Leipzig, pp.64-74, 2013.

[6] J. Siegl, “Schaltungstechnik”,4th Edition, Springer Verlag Heidelberg, pp. 128-129, 2011.

[7] M. Krystek, “Berechung der Messusicherheit”, Betuh Verlag GmBH, Berlin, pp.44-46, 2012.

[8] D. Lukač, M. Andrejević-Stošović, V. Litovski, “Operating Points and Topographic Dependence of the Thin Layer-Photovoltaic Cells as Relevant Characteristics for Modeling of the PV Cells”, Proceedings of the 4th Small Systems Simulation Symposium, Niš, pp. 20-23, 2012.

[9] M. Hofmann, “Time-distraught photoemission spectroscopy in Au-GaAs Schottky contacts”, Thesis, Faculty of Physics and Astronomy, Bavarian Julius-Maximilians University, pp.34-35, 2011.

[10] D. Lukač, H. Kücük, “Validation of the Photovoltaic (PV) - cell models for PV cells of the different cell technologies”, Theis, pp. 1-93, Rheinische Fachhochschule Köln, 2011.

[11] M. Andrejević Stošović, D. Lukač, V. Litovski, “Modeling and circuit simulation of photovoltaic cells – an overview”, 7th Int. Symp. Nikola Tesla, Belgrade, Serbia, pp. 83-92, November 2011, Invited paper.

Improved model for estimating PV system production

Marko Ikić, Jovan Mikulović, Željko Đurišić
Faculty of Electrical Engineering
University of Belgrade
Belgrade, Serbia
marko.ikic@etf.unssa.rs.ba, mikulovic@etf.rs, djurisic@etf.rs

***Abstract*— In this paper is presented improved model for estimating the electricity production of photovoltaic (PV) systems. In the literature is known two kinds of models, theoretical one and one based on the measured data on horizontal solar insolation. In second one, solar insolation on panel surface, beside of other parameters, is determinate based on average value of slope factor. This paper gives improvements in model for estimating panel production, determining solar insolation on panel surface based on available data of average ten-minute solar insolation on horizontal surface and ambient temperature, and calculating slope factor for each ten-minute period. Experimental results obtained by the real photovoltaic system mounted on a roof of Faculty served as a verification of described improved model.**

Keywords- output power; photovoltaic system; solar insolation

I. INTRODUCTION

Photovoltaic (PV) system is an integrated set of photovoltaic modules and other components, designed so that the primary solar energy directly transform into electricity, which ensures the work of a number DC and/or AC consumers. In photovoltaic systems connected to the distribution grid, the direct current produced by the solar modules is converted to AC power by the inverter, which is connected to the grid, allowing, besides supplying the consumer, energy exchange with the grid [1].

Design of photovoltaic systems is usually done based on their annual energy production, which is also a good parameter for monitoring the long-term performance of a photovoltaic system. To predict the annual energy production of photovoltaic systems, reliable models and methods are needed, considering the stochastic nature of solar radiation and the large number of influencing factors (environmental conditions and system performance) [2]. In the literature [3-6] there are presented models for predicting the power and energy of solar radiation per unit of area (irradiation and insolation).

This paper describes the improved model for estimating PV system production based on measured data of solar insolation on horizontal surface and ambient temperature. Verification of the model was carried out on a real photovoltaic system mounted on the roof of Faculty of Electrical Engineering East Sarajevo. Analytics results and practical measurements are given in following of paper.

II. PHOTOVOLTAIC SYSTEM PRODUCTION ANALYSIS

In order to design a grid connected PV systems, solar energy resources, environmental conditions, and characteristics of all elements of system must be well acquainted. Estimate of solar energy resources is based on measurements and calculations based on solar radiation at the surface on which it is planned to set up the panel. In order to estimate system performance, the rated DC power output of an individual module under standard test conditions (irradiation of 1 kWh/m^2, air mass ratio AM 1.5, cell operating temperature 25°C, modules completely clean) can be used at the beginning of process. In real operating conditions, output power of PV system delivered to the grid P_{AC} is less than the DC output modules at standard conditions $P_{DC(STC)}$ for due to losses regarding conversion efficiency:

$$P_{AC} = P_{DC(STC)} \cdot \eta_Z \cdot \eta_N \cdot \eta_T \cdot \eta_{inv} \quad (1)$$

where η_Z, η_N and η_T define efficiency reduction due to dirty of panels, modules mismatch, differences in ambient conditions, and inverter efficiency η_{inv}. The impact of these losses can reduce the output power by 20–40% [1].

The inverter efficiency varies according to the load. Grid connected inverters have efficiency over 90%, except at very low loads. Modules mismatch causes a decrease in the output power of parallel connected modules, because their current-voltage characteristics are not identical.

Another factor that has significant effect on the reduction of output power of panel below the rated value is cell operating temperature. To be able to determine the module efficiency under different ambient conditions it is necessary to calculate the temperature of the module. On the module temperature dominant influence has solar radiation and air-cooling conditions (wind). For each module manufacturer defines the temperature at nominal operating conditions (*NOCT - Nominal Operation Cell Temperature*). The *NOCT* is cell temperature in a module at ambient temperature of 20°C, solar irradiation of 800 W/m^2 and wind speed of 1 m/s. Based on *NOCT* parameter module temperature can be estimated by:

$$T_{cell} = T_{amb} + \left(\frac{NOCT - 20}{0.8} \right) \cdot I_c \quad (2)$$

where are: T_{cell} module temperature, T_{amb} ambient temperature, and I_c solar irradiation on the surface of the module.

Due to an increase in temperature of solar cells above the standard value (25°C), and module efficiency reduction of ΔP = -0.5%/°C, output power on DC connectors of the system is given by (3).

$$P_{DC(PTC)} = P_{DC(STC)}\left(1-0.005\left(T_{cell}-25^0\mathrm{C}\right)\right) \quad (3)$$

Taking into account the above mentioned losses and solar insolation, output power that PV system delivering to grid can be estimated by (4).

$$P_{AC} = P_{DC(STC)}\left(1-0.005\left(T_{cell}-25^0\mathrm{C}\right)\right)\cdot I_c \cdot \eta_Z \cdot \eta_N \cdot \eta_{inv} \quad (4)$$

Determination of insolation on the panel surface can be done in two ways, as it described in [1]. The first method is the determination of insolation on the surface of the panel based on extraterrestrial irradiation on a clear day (*Clear day* model - theoretical model), while the second method determines the insolation on the surface of the panel in real conditions based on measured insolation on a horizontal surface. Review of these two methods is given as following.

A. Determination of insolation using „Clear Day" model

Solar radiation reaches the solar panels in the form of beam (direct), diffused and reflected radiation. These components of solar radiation on the solar panel in a clear day can be calculated on the basis of extraterrestrial solar insolation [1]. Solar radiation in the form of beam radiation that reaches the earth's surface is less than extraterrestrial radiation due to absorption and scattering in the atmosphere. Evaluation of beam irradiation of the earth's surface is given Bouquer-Lambert law:

$$I_B = Ae^{-km} \quad (5)$$

where I_B represent beam radiation on the earth's surface, A extraterrestrial flux that entering in the atmosphere, k is coefficient of attenuation of solar radiation in the Earth's atmosphere (optical depth), and m is air mass ratio. Calculation of the values present in the previous equation is done according to the relations given in [1]. The total insolation I_c on the surface of the panel is given by:

$$I_C = I_{BC} + I_{DC} + I_{RC} \quad (6)$$

where I_{BC} represent beam insolation on the panel surface, I_{DC} diffuse radiation of the panel, and I_{RC} reflected radiation by surfaces in front of the panel, respectively. These components of total solar insolation can be calculated according to the following relationship:

$$I_{BC} = I_B \cdot \cos\theta = I_{BH} \cdot R_B \quad (7)$$

$$I_{DC} = I_{DH} \cdot \left(\frac{1+\cos\Sigma}{2}\right) \quad (8)$$

$$I_{RC} = \rho\left(I_{BH} + I_{DH}\right)\left(\frac{1-\cos\Sigma}{2}\right) \quad (9)$$

where are: θ incidence angle, Σ tilt angle of the panel, ρ ground reflectance, R_B slope factor, I_{BH} beam insolation on a horizontal surface, I_{DH} diffuse insolation on a horizontal surface.

The beam insolation on a horizontal surface I_{BH} and diffuse insolation on a horizontal surface I_{DH} is calculated as follows:

$$I_{BH} = I_B \sin\beta \quad (10)$$

$$I_{DH} = C \cdot I_B \quad (11)$$

where C represents sky diffuse factor.

B. Determination of insolation using model based on measured insolation at a horizontal surface

In order to extrapolate measured data of the total horizontal insolation I_H, which were collected for a given period, to arbitrarily oriented solar module it is necessary to decompose the total horizontal insolation at appropriate direct-beam I_{BH} and diffuse I_{DH} component [1,6,7,8]. For the determination of the total horizontal insolation components it is necessary to calculate the clearness index K_T. Clearness index can be calculated for each day or as average monthly index. It is define as the ratio of the average horizontal insolation I_{HAV} and extraterrestrial insolation on a horizontal surface I_{0AV}.

$$K_T = \frac{I_{HAV}}{I_{0AV}} \quad (12)$$

Higher value of clearness index means that the sky and atmosphere is clear and lower value indicates overcast conditions. The average daily on a horizontal surface I_0 can be obtained by integrating total extraterrestrial insolation from sunrise to sunset and its projection on a horizontal surface.

The diffuse portion of horizontal insolation can be estimated by using Liu-Jordan's empirical relation who gives correlation between the diffuse component of insolation on a horizontal surfaces and clearness index [1,9].

$$\frac{I_{DHAV}}{I_{HAV}} = 1.39 - 4.027 \cdot K_T + 5.531 \cdot K_T^2 - 3.108 \cdot K_T^3 \quad (13)$$

The direct-beam component of insolation on the surface of the panel is calculated based on the relation (14), while other two components can be obtained based on (8) and (9).

$$I_{BC} = I_{BH} \cdot \frac{\cos\theta}{\sin\beta} = I_{BH} \cdot R_B \tag{14}$$

Incidence angle θ, between the collector and beam, depends on the orientation of the panel, panel inclination, altitude angle β and the solar azimuth, so that slope factor is changed during the day. If it is known only the information of the average monthly (or daily) values of horizontal insolation, it is necessary to calculate the average value of the slope factors [1,9].

$$\overline{R}_B = \frac{\cos(L-\Sigma)\cdot\cos\delta\cdot\sin H_{SRC} + H_{SRC}\cdot\sin(L-\Sigma)\cdot\sin\delta}{\cos L\cdot\cos\delta\cdot\sin H_{SR} + H_{SR}\cdot\sin L\cdot\sin\delta} \tag{15}$$

The average value of the slope factor is estimated by averaging the value of cosθ over those hours of the day in which the sun is in front of the panel and dividing that by the average value of sinβ over those hours of the day when the sun is above the horizon. Expression (15) used in [1] represents a closed-form solution for those averages.

In this paper, it will be shown that the calculation of the slope factor by procedure of averaging results in an error in the assessment of irradiation on the PV panel, and consequently, an error in the estimation of power production. It be clearly highlighted the difference between instantaneous and average slope factor and their impact on the estimation of power production.

III. EXPERIMENTAL RESULTS AND DISCUSSION

In this chapter, practical realization of one PV system with weather measurement equipment, mounted on a roof of Faculty of Electrical Engineering East Sarajevo is given.

Figure 1. PV panels and weather measurement equipment mounted on the roof of Faculty

Test results, obtained by the simulation of two models described before (in Matlab/Simulink environment) and practical measurement of horizontal irradiation and PV system output power are described.

In Fig.1. is shown *Prostar Solar Model. PR75Wp/24V* four panels connected by two in series ($P_{DC(STC)}$ = 2x75 = 150 Wp), oriented on the south and fixed by angle of 33^0 (optimal angle for region of Sarajevo). Measurement equipment with data logger iBOX-EKO21N-v7 is also installed on the roof measuring the wind speed, wind direction, horizontal solar irradiation and environmental temperature. Detailed information about measurement equipment can be found in [10].

PV panels are connected to the load via converter with MPPT function (with P&O algorithm implemented). This converter was developed as prototype at the Faculty, represents a Buck-Boost converter controlled by dual current mode control and efficiency of 93%. The lack of equipment for grid connection, system was connected to the active load of R = 15 Ω (Fig.2.). Measurement of the output current and voltage were carried out every ten minutes during period of 11 a.m. to 1 p.m. at Wednesday May 7th 2014, covering the solar noon at that way.

Results of measurements of solar radiation on horizontal surface and ambient temperature, as well as measurement of PV system output power are given in Table I.

As can be seen from Table I, average solar irradiation on horizontal surface was from 800 to 900 W/m^2 during the period of measurement. But, in some time interval solar irradiation on horizontal surface was less than average (point at 11:10 and 11:20), that cause lower production of PV system. Also, from experimental results of PV system output power is verified that ambient temperature has influence of energy production. At point of 12:40 and 12:50 output power of PV system has the same amount 112.8 W, although the solar insolation is higher in second case. This is result of the increase the temperature in second case and its negative influence on energy production.

Figure 2. Test equipment with measuring instruments and MPPT converter

TABLE I. MEASURED PARAMETERS OF SOLAR INSOLATION ON HORIZONTAL SURFACE, AMBIENT TEMPERATURE AND PV SYSTEM OUTPUT POWER

Time period	Solar irradiation [W/m^2]	Ambient temperature [$^{\circ}$C]	PV system output power [W]
11:00	873.9	20.53	113.8
11:10	501.3	21.24	67.3
11:20	724.3	20.96	99.2
11:30	831.7	20.99	109.8
11:40	839.1	23.23	108.4
11:50	839.1	23.14	109.1
12:00	842.4	24.27	107.7
12:10	816.4	24.56	106.1
12:20	867.1	23.58	111.6
12:30	873	23.92	112.6
12:40	878.2	24.24	112.8
12:50	882.6	24.51	112.8
13:00	903.5	24.7	114.5

In Fig.3. and Fig.4. the components of the solar radiation on panel obtained by two described models and measured solar insolation on horizontal surface are shown. “Clear day” model as theoretical model does not respect the terms of cloudiness, and as such gives estimated solar insolation components (Ic, Ibc, Idc, Irc) different from measured solar insolation (Ih), as it is shown on Fig.3. A better estimation of solar irradiation components is achieved using the realistic model (including suggested adjustment) with respect the terms of cloudiness. As it is shown on Fig.4., total solar irradiation on panel and its components have the same shape and follow behavior of measured solar irradiation on horizontal surface. Difference between total solar irradiation on panel and measured solar irradiation on horizontal surface, is in additional amount of solar irradiation on panel due to diffuse and reflection of solar radiation, which brings more PV system output power production.

Difference in PV system production based on two models and experimental results is shown on Fig.5. It is obvious that estimating PV system production based on “Clear day” model could give an incorrect value of energy production with which is plan to design PV system. Verification of the realistic model, based on measured parameters of solar insolation on horizontal surface and ambient temperature, shows that estimated PV system production is very close to the one with experimental measured results. Minor variations in the graphs (Pac-Realistic model and Pacm) may be due to wrong assessment of efficiency factors due to panel dirty and modules mismatch as well as acceptable error of measurement instruments.

On the other hand, influence of calculating slope factor on determining solar insolation on panel surfaces as well as estimating panel production, can be seen on Fig.6. and Fig.7.

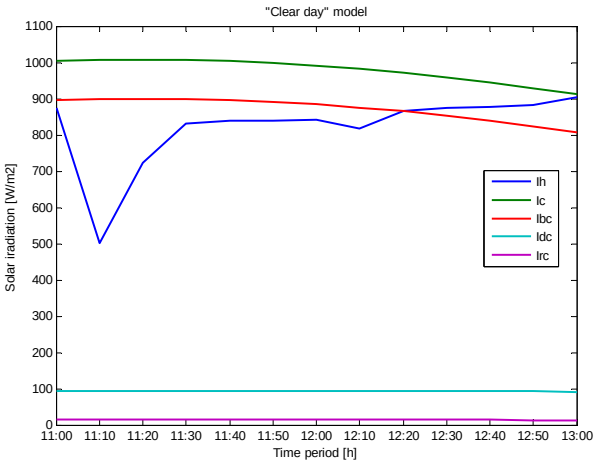


Figure 3. Measured solar insolation on horizontal surface and total solar insolation on panel with its component obtained with “Clear day” model.

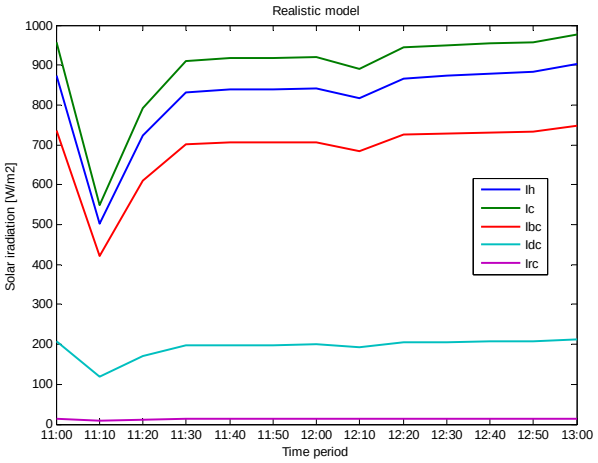


Figure 4. Measured solar insolation on horizontal surface and total solar insolation on panel with its component obtained with realistic model.

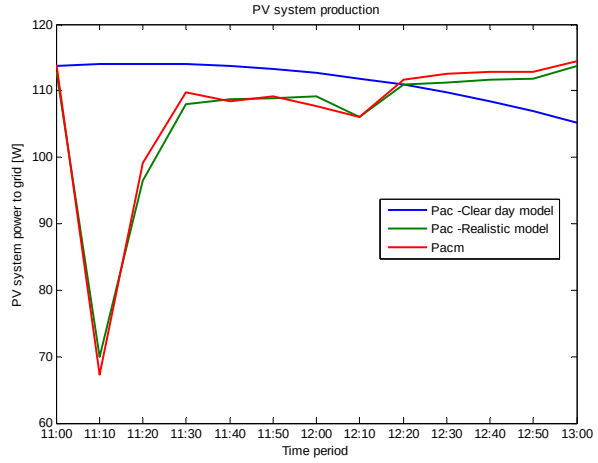


Figure 5. Experimental results of PV system output power with estimating one obtained with given models.

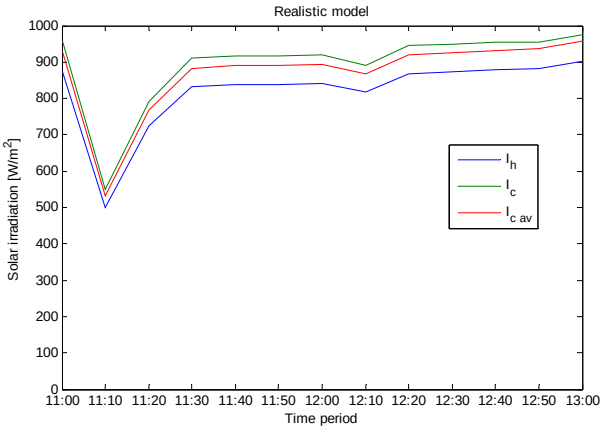


Figure 6. Measured solar irradiation on horizontal surface and total solar irradiation on panel determined based on average and instantaneous slope factor.

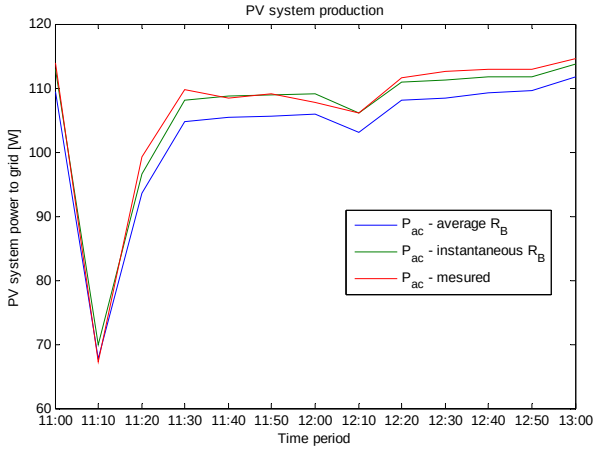


Figure 7. Experimental results of PV system output power compared with estimating ones based on different calculation way of slope factor.

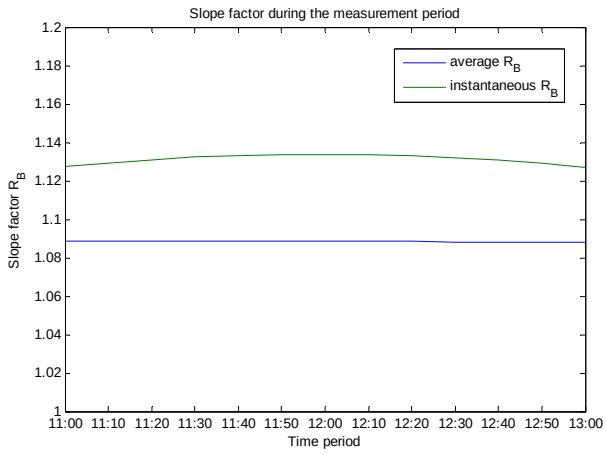


Figure 8. Difference between average and instantaneous slope factor during the measurement period.

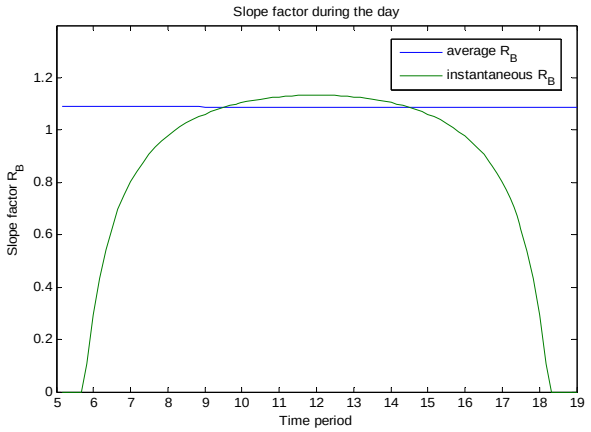


Figure 9. Difference between average and instantaneous slope factor during the day.

Simulation results from Fig.6. shows that the solar insolation on panel surface calculated with average slope factor (referred to the [1]) has the smaller values in comparison to the solar insolation calculated based on instantaneous slope factor. Therefore, the estimated PV panel production (shown on Fig.7.) is lower when the average slope factor is used. Error in power production obtained with average slope factor calculation relative to measured output power is -3.14%, while the error obtained with instantaneous slope factor calculation is -0.47% (for the measurement period). In Fig.8. and Fig.9. the average and instantaneous slope factor during the measurement period and during the day are drawn, respectively. It is clear that from 9 a.m. until 3 p.m. instantaneous slope factor has higher value than average one, so the solar insolation, as well as power production, is higher, but the question is what happens beyond this time period related to power production and determination of solar insolation in cases of use the instantaneous or average slope factor. The answer lies in fact that the solar irradiation in that period is much lower than in period when instantaneous slope factor has higher value relative to average, so the impact on the power production will be small.

IV. CONCLUSION

Considering the stochastic nature of solar radiation and the large number of influencing factors (environmental conditions and system performance), to estimate the PV system production it is suitable to use a model based on knowledge of the solar insolation on horizontal surface and ambient temperature. Results of estimated PV system production with improved model realized in Matlab/Simulink environment well matched with experimental results measured from real PV system installed on a roof of Faculty of Electrical Engineering East Sarajevo. From previous, it can be concluded that the difference in the calculation of the slope factor (instantaneous and average), has an impact on the estimation of power production. By using the instantaneous slope factor for estimation of PV system production, a minor error occurs while designing the PV system.

REFERENCES

[1] Gilbert M. Masters, "Renewable and Efficient Electric Power Systems", Stanford University, A John Wiley & Sons, Inc., Publication 2004

[2] D. L. King, W. E. Boyson, J. A. Kratochvil, "Analysis of Factors Influencing the Annual Energy Production of Photovoltaics Systems", Photovoltaic Specialists Conference, 2002.

[3] M. J. Ahmat, G. N. Tiawari, "Solar radiation models – review", International Journal of Energy and Environment, Volume 1, Issue 3, 2010, pp. 513-532.

[4] S. Kaplanis, E. Kaplani, "A model to predict expected mean and stochastic hourly global solar radiation I(h, nj) values", Renewable Energy 32, 2007, pp. 1414-1425.

[5] P. Bacher, "Short-term Solar Power Forecasting", Kongens Lyngbay 2008, IMM-2008-13.

[6] J. Mikulović, Ž. Đurišić, "Statistička metoda za procenu proizvodnje električne energije iz fotonaponskog sistema ", Infoteh, Mart 2012.

[7] J. Mikulović, Ž. Đurišić, R. Kostić " Određivanje optimalnih nagibnih uglova fotonaponskih panela ", Infoteh, Mart 2013.

[8] O. Perpinan, E. Lorenzo, M. A. Castro, "On the calculation of energy produced by PV grid-connected system", Progress in Photovoltaics Research and Applications (2007), Volume: 15, Issue: 3, Publisher: Wiley Online Library, Pages: 265-274.

[9] Lj. Sinjeri, P. Kulišić, " Solar radiation on variously oriented collectors in Croatia", Renewable Energy Vol. 4, No. 2, 1994, pp. 235-240.

[10] www.ekopower.nl, October 2014.

Session TO-10
TELECOMMUNICATION TECHNOLOGIES

The Personalization of a Cloud Ecosystem: Adding Dimensions to Situational Awareness

Invited Paper

Gordana Velikic, Milan Bjelica, Nemanja Ignjatov, Mica Cetkovic
RT-TK Institute for Computer Based Systems
Serbia, Novi Sad

Ivan Kastelan
FTN, University of Novi Sad
Novi Sad, Serbia

***Abstract*—The lack of easy to use environment to connect and manage the things is one of the obstacles that hold back multimodal personalized use of Internet of Things. We present a novel idea and a cloud framework with original elements that enable development of ecosystem for the user tailored deployment of things in the context of a personalized situational awareness. The benefits of the infrastructure are fortified with several use case scenarios that provide an insight to the platform's innovative possibilities.**

Keywords-Cloud, Internet of Things, Situational Awareness

I. INTRODUCTION

Internet of Things (IoT) [1] paradigm has been extensively promoted in recent years, however, the overall impression is that the number of end-user applications benefiting on the generic IoT platforms, such as Xively [2] or lately Apple's HomeKit [3] is still low.

Several problems are holding back IoT to evolve a step further from the basic eco-system of things. The major problem is lack of easy to use environment to connect and manage the things, and to navigate through the sea of data to extract desired user-centric information. Instead of searching through data and configuring devices, users are rather interested in gaining an easy access to information and devices, and to easily tailor parameters of applications according to their needs.

This leads us to situational awareness concept that can be easily utilized with easy to use user tailored platform to create ad-hoc applications for use case scenarios in different domains of interest, such as but not limited to security, wellness, or ambient adjustment. Situational awareness refers to knowledge about the environment and its changes over time or some other predefined parameter, like for example energy consumption [4]. The situational awareness should not be confused with a context awareness which is mostly associated with location awareness and activity recognition [5]. Context awareness is mostly applied to cases where user's behavior needs to adapt to physical environment. Situation awareness models users' environments and help them to be aware of current situation. The concept embedded in our platform is that user's can define and personalize specific processed situational awareness captured by the system for novel uses, such as but not limited to: home energy consumption in security or housekeeping monitoring, guest behavior in wellness, or TV channel status in ambient settings or parental control. Novel elements – user-friendly script agents enable users to individually define situational awareness use case scenario and create a script agent that will execute desired application.

II. PERSONAL CLOUD AGENT FRAMEWORK

Each object in our surrounding can be a building block for an IoT platform. Organized hierarchically, some objects may be able to detect raw physical parameters, some can devise state of the environment based on rules, and the most complex objects can provide alerts and guidance [6]. In this paper a concept of a personal cloud agent framework is proposed, which allows the creation of scriptable cloud agents. Using aggregated information in a customer profile, processed collected raw data acquired from devices, or both with additional information from other accessible sources, the customer can use the scriptable cloud agents to create personalized applications in home settings, or to stimulate actions on various devices available around customers in diverse commercial applications or settings.

The scriptable agents are part of a self-growing cloud-based platform for different content, devices, and environments (context based applications that can evolve into awareness based applications). The user, either an individual or industry, can tailor scripts to adapt or create novel scenarios and uses. For example, a hospitality industry, e.g. hotels, can create a specific personal cloud agent to send suggestions related to customer wellness back to all devices and applications in physical realm of IoT. The intention behind the creation of such agent is to enhance customers' experiences, their well-being, and overall satisfaction during a visit to vacation venue.

The proposed framework provides a technical medium for creation of objects for user-specific ad-hoc uses, such as above mentioned personal agents which provide guidance in wellness. A generic agent platform for IoT allows the implementation of IoT applications in which agents aggregate information sourced from various devices and protocols, and pass on the information among themselves to create target applications [7]. Script-based agents, written in Lua, allow for the implementation of complex logic and device self-management

[8]. Personal cloud agent framework is based on Insight IoT platform [8][9], which enables communication with end devices via TR-069 and XMPP and exposes appropriate API to above layers (see Figure 1).

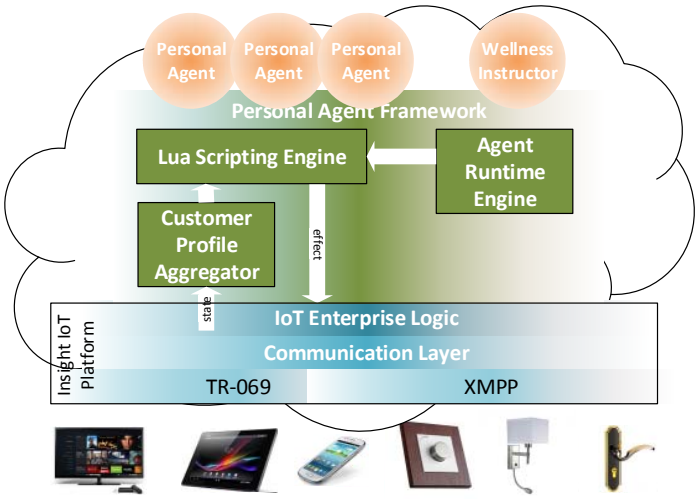


Figure 1. Personal Cloud Agent Framework Architecture: "personalized" for wellness aplication for use in hospitality industry.

Each device is presented with its *data model*, in the form of an object/parameter tree as defined in [10]. Each parameter can be set or read, and thus deploy a command to a device that either configures the device or queries its status. The *Customer Profile Aggregator* maps each device to an individual *customer profile*. Data model of each customer device is periodically processed to extract information relevant to the agent framework (as shown in Table I). *The Lua Scripting Engine* enables the execution of agents in the form of Lua scripts, while enabling extension APIs for Lua. This allows agents to access customer profiles and to manipulate with targeted device's data model from script language (e.g. to send deduced knowledge to a user application). The *Agent Runtime Engine* maintains the lifecycle of an agent instance, and links the script code of an agent with the customer profile and the device list that the agent is associated with. Additionally, the *Personal Agent Framework* provides an API for the top level web based applications, and allows them to manipulate the customer profiles and agents, and to make assignments.

TABLE I: INFORMATION RELEVANT TO A CUSTOMER PROFILE

Data model parameter	*Extracted information*
Outlet.1.Consumption.Value	Customer presence in the room
Dimmer.1.Mode	Customer activity / sleeping / bathroom
Time.CurrentLocalTime	Estimated daylight / Wellness availability / Expected biorhythm
STBService.Components. Service.CurrentService	Customer activity / Watched TV show / Customer mood
LAN.Stats.TotalBytesSent	Customer activity on phone / laptop

Personal agents are programmed in Lua scripting language, which is extended to use exposed Insight IoT APIs.

III. USE CASE SCENARIOS

A. *Wellness instructor*

Wellness instructors are tasked to detect user behavior which has bad impact on wellbeing (e.g. long watching of TV program or inappropriate content, bad habits such as long stay indoors), or to detect favorable conditions for well-being enhancement (e.g. nice weather, sauna or pool availability, diet menu suggestion, vitamin pill intake time). Based on these detections, these agents send appropriate messages to devices which may be able to communicate to customers (e.g. TV or mobile applications), or to apply corrective actions without human interventions, such as for example switching light off.

For the purpose of evaluation, a wellness instructor agent was programmed in Lua. Android set-top box was used as a Smart TV in a hotel room simulation, what would allow the customer in the room to review suggestions set by the agent in the specifically designed web-based wellness application, accessible via remote controller. The same set-top box was a member of IoT realm, disclosing information about the usage of TV. Additionally, the room was equipped with one smart outlet, two smart dimmer switches and one dimmable lamp, all reporting their state to the IoT platform. In addition to Android application, feedback from wellness instructor was also in the form of automatic control over lighting conditions in the room.

TABLE II: EXCERPT FROM THE DECISION TREE OF WELLNESS INSTRUCTOR

If this is observed…	*…then do the following*
Wellness working hours, using laptop for more than 2 hours, not watching TV, lights off	Suggestion: Swim in the pool to relax Action: Turn on lamp to 20% dim level to prevent eye fatigue
Just arrived in the room, evening, used light in the bathroom for 10 minutes, turns bathroom light off	Suggestion 1: Visit sauna before bed Suggestion 2: Go to bed earlier Action: Set bedtime reading lights
Morning, breakfast time, TV on time history recorded	Action: Turn TV on to music channel, set low volume, switch on wake up lighting scenario
Watching news on TV for more than 2 hours, late evening	Suggestion 1: Switch to comedy Suggestion 2: Go to sleep

An excerpt from the agent's decision tree is shown in Table II.

For the illustration purposes, the evaluation was performed regarding to functionality and to potential to enhance customer well-being. Functionality was validated in lab conditions with real devices and human user. The potential for well-being enhancement was evaluated through a simulated customer behavior with IoT devices based on the observation of family behavior in a hotel room during a vacation. According to the probability of customer hotel room occupancy during a day estimated as in [11], and the typical opening hours of hotel amenities (restaurant, pool, spa and gym), several time windows were identified that are appropriate for the use of wellness amenities (Figure 3). Most of the suggestions made by the wellness instructor agent (~90%) in the simulation matched these time windows. Also, wellness instructor was dimming lights to pleasant levels when a customer was believed to be sleeping, what matched to expected sleeping hours in 80% of test cases. Another possible scenario for this agent is to detect possible health hazard situation if the guest enters additional health data.

```
local step = 1000*5;
local limit = step*24;
local counter =0;

    setParameterValue("Device.EnergyManagementService.RoomSituation.1.TotalVacuumingTime",0,"key");
    setParameterValue("Device.EnergyManagementService.RoomSituation.2.TotalVacuumingTime",0,"key");
    setParameterValue("Device.EnergyManagementService.RoomSituation.3.TotalVacuumingTime",0,"key");
    setParameterValue("Device.EnergyManagementService.RoomSituation.1.VacuumedToday","false","key");
    setParameterValue("Device.EnergyManagementService.RoomSituation.2.VacuumedToday","false","key");
    setParameterValue("Device.EnergyManagementService.RoomSituation.3.VacuumedToday","false","key");

while counter <= limit do
   local room = math.random(1,1000);
   room = (room % 3) + 1;
   local vacuumpw1;
   local vacuumpw2;
   local vacuumpw3;

    vacuumpw1 = getParameterValue("Device.EnergyManagementService.EnergyMeter.1.CurrentConsumption");
    vacuumpw2 = getParameterValue("Device.EnergyManagementService.EnergyMeter.2.CurrentConsumption");
    vacuumpw3 = getParameterValue("Device.EnergyManagementService.EnergyMeter.3.CurrentConsumption");

    if(vacuumpw1 > "0") then
        local time1 = getParameterValue("Device.EnergyManagementService.RoomSituation.1.TotalVacuumingTime");
        setParameterValue("Device.EnergyManagementService.RoomSituation.1.TotalVacuumingTime",time1 + step,"key");
        setParameterValue("Device.EnergyManagementService.RoomSituation.1.VacuumedToday","true","key");
    end

    if(vacuumpw2 > "0") then
        local time2 = getParameterValue("Device.EnergyManagementService.RoomSituation.2.TotalVacuumingTime");
        setParameterValue("Device.EnergyManagementService.RoomSituation.2.TotalVacuumingTime",time2+step,"key");
        setParameterValue("Device.EnergyManagementService.RoomSituation.2.VacuumedToday","true","key");
    end

    if(vacuumpw3 > "0") then
        local time3 = getParameterValue("Device.EnergyManagementService.RoomSituation.3.TotalVacuumingTime");
        setParameterValue("Device.EnergyManagementService.RoomSituation.3.TotalVacuumingTime",time3+step,"key");
        setParameterValue("Device.EnergyManagementService.RoomSituation.3.VacuumedToday","true","key");
    end

    if((vacuumpw1 == "1000") and (vacuumpw2 == "1000")) then
            setParameterValue("Device.EnergyManagementService.CurrentAlerts","Room 1 and Room 2 is currently being vacuumed.","key");
    elseif((vacuumpw2 == "1000") and (vacuumpw3 == "1000")) then
       setParameterValue("Device.EnergyManagementService.CurrentAlerts","Room 2 and Room 3 is currently being vacuumed.","key");
    elseif((vacuumpw1 == "1000") and (vacuumpw3 == "1000")) then
       setParameterValue("Device.EnergyManagementService.CurrentAlerts","Room 1 and Room 3 is currently being vacuumed.","key");
    else
        setParameterValue("Device.EnergyManagementService.CurrentAlerts","","key");
    end


   sleep(step);
   counter = counter + step;
end
```

Figure 2. “Raw” situational awareness script for the use case scenario B.

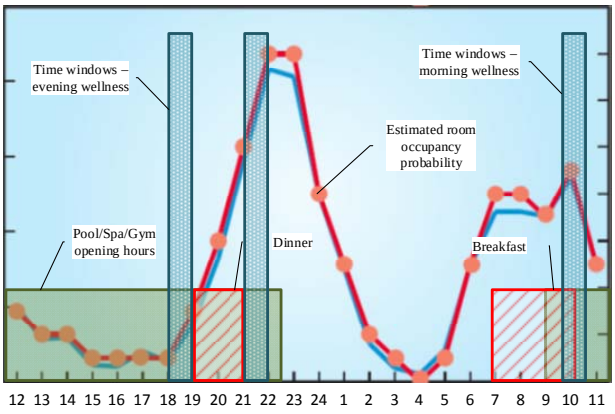


Figure 3. Time windows appropriate to use the hotel wellness amenities.

B. *Alert of possible intruder, monitoring house keeping*

Another example of situational awareness scenario demonstrates the utilization of a seemingly “useless” sensor, such as the smart outlet with energy measurement capability. In the scenario, it is possible to detect if a chore has been correctly accomplished for a given week. For example, a housemaid is tasked to perform vacuum cleaning of the whole home. Situational awareness script, as shown in Figure 2, enables the detection of the task fulfillment based on energy consumption from outlets in the vacuumed rooms. Assuming that only one vacuum cleaner is available and just one housemaid, script assumes that no two rooms can be simultaneously vacuumed and sets alerts accordingly. In the given script, energy consumption is being continuously acquired from smart outlets in each of the three rooms. If the consumption is non-zero, total vacuuming time for the respective room is being increased. If consumption is detected on more than one outlet at any given time, script alerts the homeowners of the inconsistency. In addition, homeowners are able to see final vacuuming times which are stored to output parameters, and therefore validate if the vacuuming was performed for each of the rooms in the adequate duration.

Quick Parameter View

Device Name	RoomMonitoring	Set
CurrentAlerts	Room 1 and Room 2 is current…	Set
CurrentConsumption	1000	Set
CurrentConsumption	0	Set
CurrentConsumption	1000	Set
VacuumedToday	true	Set
TotalVacuumingTime	70000	Set
TotalVacuumingTime	10000	Set
VacuumedToday	true	Set
TotalVacuumingTime	20000	Set
VacuumedToday	true	Set

Figure 4. Web application snap shot for the use case scenario B.

```
local step = 1000*5;
local limit = step*24;
local counter =0;

setParameterValue("Device.STBService.1.Components.UserActivityMonitoring.CurrentAlerts","","key");


while counter<=limit do
    setParameterValue("Device.STBService.1.Components.UserActivityMonitoring.CurrentTime",os.date("%c",),"key");
    local lastServiceChange = getParameterValueAndWait("Device.STBService.1.Components.UserActivityMonitoring.LastZapClk");

    print("Time elapsed since last service change "..os.time()-lastServiceChange);

    if((os.time()-lastServiceChange) > 4*60*60 ) then
        setParameterValue("Device.STBService.1.Components.UserActivityMonitoring.CurrentAlerts","Entering sleep mode in progress!");
    end

    sleep(step);
    counter = counter + step;

end
```

Figure 5. "Raw" situational awareness script for the use case scenario C.

Figure 4 shows the snapshot from a demonstration web application, which can be accessed by homeowners. In this snapshot, alerts can be accessed as well as current consumption on each of the outlets and, more importantly, a conclusion from the awareness script, stating if the room was finally vacuumed or not. Although a number of improvements with regard to edge cases and false diagnoses can be applied during the script refinement, the given example indicates the feasibility of situational awareness with the given platform even in the case of availability of only the most simplistic sensing. Examples of other uses with the same script may be intruder alert during vacation, or unexpected kid's party during parents night out. The alerts may be forwarded to phone or e-mail.

C. *Ambient settings*

The following situational awareness script enables the remote detection of idleness in TV watching. The script (see Figure 5) can efficiently detect if the channel was not changed for an excessive time (e.g. 4 hours), and suggest the sleep mode to the TV device via a specific parameter setup (as in Figures 6 and 7). The demonstration of the operation is shown in snapshots from the web application, showing the currently watched channel, last zapping time and alert field, indicating the sleep mode situation, if detected. Additional information from passive infrared sensor would further enhance this scenario, and enable more accurate detection of sleep mode, mitigating the edge case in which the user excessively watches one particular channel without zapping. Integration of situational awareness system with hybrid, connected devices, such as set-top boxes, is a natural step and the presented example shows one potential usage. Another possible scenario that can be built in this script is ambient mode settings for particular genre or channel.

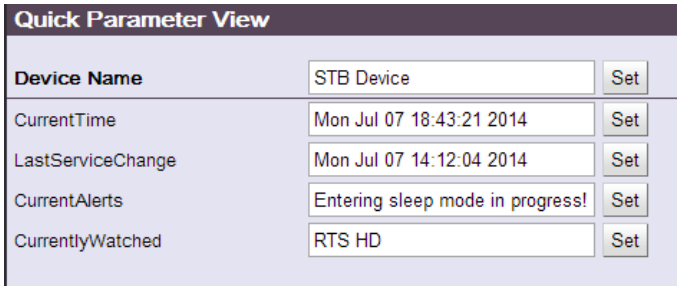


Figure 6. Web application snapshots for the use case scenario C. Given plot shows time span almost ten minutes and no alerts generated.

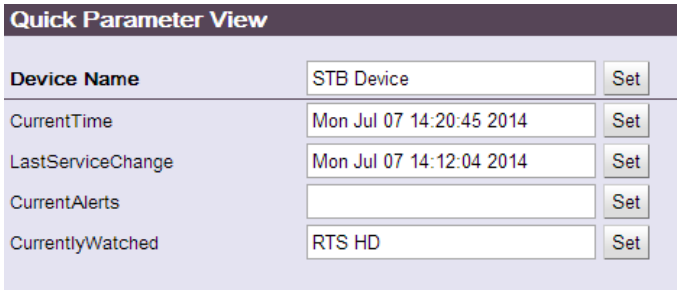


Figure 7. Web application snapshot for the use case scenario C. Given plot shows generated alert that will set preset actions since no activity is registered in defined time frame. In this case it is about four hours.

CONCLUSION

The presented cloud based platform cancels the need for user to have extra computational power, ads to wider accessibility, mobility, and lower maintenance cost for application developers. Modular and device agnostic approach makes the platform adaptable to future technology developments, extendible, and scalable. These are all features that most good platforms have. However, with our scriptable agent, it delivers an easy way for an individual customer to take control of personalization and turn data collected from day to day tasks, like cleaning, or home automation, into desired use cases without a need for professional technical help. Although the platform needs finishing touches, like intuitive and "fancy" graphical user interface to deploy commercially, it is ready and "fully loaded" to add new dimensions to situational awareness, as we showed in three examples.

REFERENCES

[1] K. Ashton, "That "Internet of Things" Thing," RFiD Journal 2009.

[2] Xively, https://xively.com/, 2014.

[3] Apple HomeKit Framework, https://developer.apple.com/homekit/, 2014.

[4] Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In Proceedings of the human factors society 32nd annual meeting, Santa Monica, CA (Vol. 1, pp. 97–101).

[5] Schilit, B., & Theimer, M. (1994). Disseminating active map information to mobile hosts. IEEE Network, 8, 22–32.

[6] G. Kortuem, F. Kawsar, D. Fitton and V. Sundramoorthy, "Smart objects as building blocks for the Internet of things," *Internet Computing, IEEE*, vol. 14, no. 1, pp. 44-51, 2010.

[7] A. Inmaculada, M. Amor and L. Fuentes, "An agent platform for self-configuring agents in the Internet of Things," *Proceedings of ITMAS 2012*, pp. 65-78, 2012.

[8] M. Z. Bjelica, N. Ignjatov, I. Papp and N. Teslic, "Device Cloud Platform with Script-Based Agents for Anywhere Access Applications Development," *Proceedings of MIPRO*, 2014.

[9] M. Z. Bjelica, G. Golan, S. Radovanovic, I. Papp and G. Velikic, "Adaptive Device Cloud for Internet of Things Applications," *Proceedings of ICCE China*, 2014.

[10] "TR-106 Data Model Template for TR-069-Enabled Devices," *Broadband Forum*, Issue 1, Amendment 3, 2009.

[11] "Energy Conservation for Hotels," *The Energy Conservation Center, Japan*, 2009.

TLM Modeling of Emissions from Printed Circuit Board of Power Amplifier Matching Circuits

Jugoslav Joković, Tijana Dimitrijević, Aleksandar Atanasković, Nataša Maleš-Ilić, Bratislav Milovanović
Department of Telecommunication
University of Nis, Faculty of Electronic Engineering,
Nis, Serbia
e-mail: jugoslav.jokovic@elfak.ni.ac.rs,

***Abstract*— The paper considers an electromagnetic (EM) emission from a printed circuit board (PCB) representing impedance matching circuits on microwave amplifier. The analysis is based on Transmission Line Matrix (TLM) method including the basic physical features of the input and output impedance matching circuits realized using the microstrip. The ports are described through the TLM wire compact model while a simple equivalent transistor model based on S parameters is applied to account for the connection between the PCB elements. Since a rectangular metallic enclosure is typical closed environment for microwave amplifier, the EM emissions inside and outside the enclosure with aperture are compared with respect to engineering purposes.**

Keywords - EM emissions, PCB, Enclosure, Power Amplifier, Aperture, TLM method

I. INTRODUCTION

In recent years, a great effort has been dedicated to development and utilization of advanced digital techniques for information processing and transmission. A number of complex components and devices, usually in high-density packaging, can be found in modern communication systems resulting in a very challenging electromagnetic (EM) field environment. Therefore, electromagnetic compatibility (EMC) [1] has become one of the major issues in design of these systems, especially some of their parts such as printed circuit boards (PCBs) and integrated circuits (ICs).

Clock rates that drive PCBs are now in the GHz frequency range in order to increase dramatically processing speed. Therefore, consideration of even a few higher harmonics of clock rates takes design of such circuits well into the microwave regime. PCBs are increasingly becoming more complex and, consequently, quantifying their EM presence is more difficult. In the microwave frequency range, PCBs have dimensions of the order of several wavelengths and thus become efficient radiators and receivers of EM energy. In addition, high-density packaging, widely applied to the PCB design, could cause a significant level of EM interference (EMI) between neighbors PCBs, particularly when they are placed in an enclosed environment. These effects in combination with the driving down of device switching voltage levels are making signal quality/integrity and emission/ susceptibility critical EMC issues in next generation high-speed systems.

Differential numerical techniques, such as the finite-difference time-domain (FD-TD) method [2] and the transmission line-matrix (TLM) method [3], are common tools for computational analysis of numerous EM and EMC problems. However, a full-wave three dimensional (3D) numerical simulation to accurately reproduce the EM field around a PCB usually requires substantial computing power and simulation run-time. Therefore, one efficient technique based on the equivalent principle [4], providing simplified equivalent dipole models to accurately predict radiated emissions without reference to the exact details of the PCB has been recently proposed [5]. The model has been deduced from experimental near-field scanning and it includes not only the excitation but also physical features of PCB such as its ground plane and a dielectric body, both very important in closed environment. However, when such model is incorporated into conventional calculation algorithms of FD-TD or TLM methods, it can be very complex and time consuming.

For some of the geometrically small but electrically important features (so-called fine features), such as wires, slots and air-vents, few enhancements to the TLM method have been developed [6-8]. These compact models have been implemented either in the form of an additional one-dimensional transmission line network running through a tube of regular nodes or in the form of an equivalent lumped element circuit, allowing to account for EM presence of fine features without applying a very fine mesh around them. Compared to the conventional approach, these models yield a dramatic improvement in computer resources required. Similar compact model could be developed for the PCB allowing for an efficient implementation into the TLM algorithm procedure and accurate representation of EM emissions and coupling of the PCB. Development of such model has assumed that an extensive full-wave analysis has to be conducted in order to fully characterize EM presence of the PCB either in the free space or in an enclosed environment.

Numerical TLM results of EM emissions from basic L-shaped PCB board inside the enclosure are verified with reference results based on equivalent dipole simulations and measurements [5,9,10]. In this paper, we consider the PCB consisting of two L-shaped microstrip tracks placed on FR4 substrate, representing the input and output impedance matching circuits of single stage power amplifier operating at 1GHz [11]. The ports at the ends of tracks are realized through the wire elements while the transistor is represented by an

equivalent model based on S parameters at the operating frequency. PCB is placed in a rectangular metallic enclosure as a typical closed environment for a power amplifier. In addition, an aperture on the top enclosure wall is also taken into account. The impact of radiated emission of this PCB structure on EM field distribution is investigated. Numerical TLM results of EM field at resonances are compared with corresponding results inside enclosure based on simulations and measurements [5]. Also, the EM field patterns inside and outside the enclosure at amplifier operating frequency are analyzed in terms of radiated EM emissions from PCB elements.

II. TLM MODELING

In the TLM method, 3D EM field distribution in a PCB structure in an enclosure is modelled by filling the space with a network of transmission lines and exciting a particular field component in the mesh by a voltage source. EM properties of substrate and the air in the enclosure are modelled by using a network of interconnected nodes. A typical node structure is the symmetrical condensed node (SCN), which is shown in Fig. 1. To operate at a higher time-step, a hybrid symmetrical condensed node (HSCN) [3] is used. An efficient computational algorithm of scattering properties, based on enforcing continuity of the electric and magnetic fields and conservation of charge and magnetic flux, is implemented to speed up the simulation process. For accurate modelling of this problem, a finer mesh within the substrate and cells with arbitrary aspect ratio suitable for modelling of particular geometrical features, such as microstrip track, are applied. External boundaries of an arbitrary reflection coefficient of enclosures are modelled in TLM by terminating the link lines at the edge of the problem space with an appropriate load.

In the TLM compact wire model, wire structures are considered as new elements that increase the capacitance and inductance of the medium which they are placed in. Thus, an appropriate wire network needs to be interposed over the existing TLM network to model the required deficit of electromagnetic parameters of the medium. In order to achieve consistency with the rest of the TLM model, it is most suitable

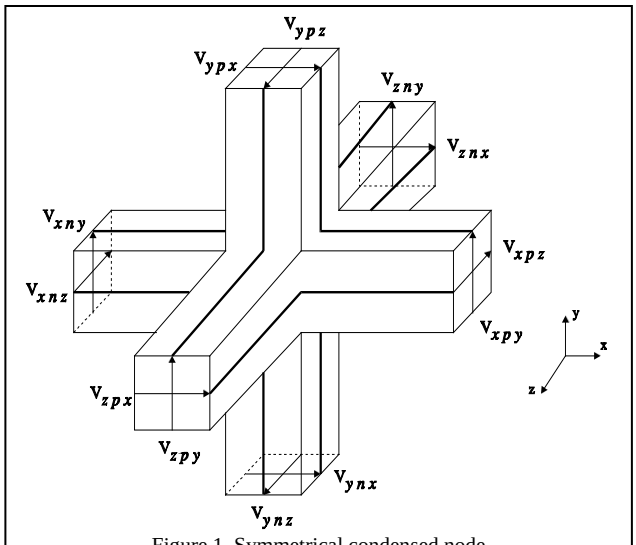


Figure 1. Symmetrical condensed node

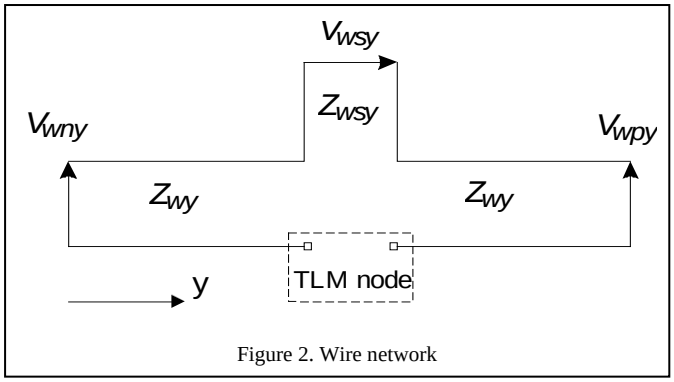


Figure 2. Wire network

to form wire networks by using TLM link and stub lines (Fig. 2) with characteristic impedances, denoted as Z_{wy} and Z_{wsy}, respectively. An interface between the wire network and the rest of the TLM network must be devised to simulate coupling between the EM field and the wire.

In order to model wire elements, wire network segments pass through the center of the TLM node. In that case, coupling between the field and the wire coincides with the scattering event in the node which makes the scattering matrix calculation, for the nodes containing a segment of wire network, more complex. Because of that, an approach proposed in [6], which solves interfacing between arbitrary complex wire network and arbitrary complex TLM nodes without a modification of the scattering procedure, is applied to the modelling of wire segments.

The single column of TLM nodes, through which wire conductor passes, can be used to approximately form the fictitious cylinder which represents capacitance and inductance of wire per unit length. Its effective diameter, different for capacitance and inductance, can be expressed as a product of factors empirically obtained by using known characteristics of TLM network and the mean dimensions of the node cross-section in the direction of wire running [6].

III. RESULTS AND ANALYSES

TLM simulations are carried out to determine the EM emissions from input and output matching circuits of a single stage amplifier, in form of microstrip lines printed on the dielectric substrate and placed in a metallic enclosure with an aperture. The layout of matching circuits is designed for single stage power amplifier based on LDMOSFET operating at 1GHz [11].

The PCB representing matching circuits of single stage power amplifier consists of two L-shaped microstrip tracks, placed on one side of a $PCB_x \times PCB_y \times PCB_z = (250 \times 150 \times 1.5)mm^3$ board made from FR4 substrate with relative permittivity $\varepsilon_r = 4.2$. The width of tracks is $w = 3.1$mm, while lengths are calculated in order to achieve the impedance matching ($l_{in,1} = 78.2$mm, $l_{in,2} = 31.6$mm, $l_{out,1} = 84.7$mm, $l_{out,2} = 27.7$mm). LDMOSFET is represented by model based on S parameters at 1GHz ($S_{11}=0.841e^{-j143°}$, $S_{21}=6.01e^{j76°}$, $S_{12}=0.018e^{j11°}$, $S_1=0.728e^{-j64°}$). The PCB is powered by external RF signals via probe (with diameter of 0.5 mm) placed at one end of microstrip track representing the

input circuit. Following the design of the amplifier powered by external RF signals via SMA connectors in practice, numerical characterization of input and output ports can be done by introducing wire ports.

The PCB is mounted on the bottom of an enclosure in the form of rectangular metallic box with dimensions $a \times b \times c = (284 \times 204 \times 75)$ mm^3. In this structure, the enclosure walls are modeled through setting reflection coefficients, while coax ports are described by using the compact wire model and applying generator and loads in TLM wire ports at the ends of microstrip track. Also, the aperture with dimensions $a_1 \times b_1 = (250 \times 10)$ mm^2, placed on the top wall of the enclosure above the PCB, is incorporated into the TLM model. The geometry of the PCB representing power amplifier matching circuits in the enclosure is shown in Fig. 3.

When a PCB is placed inside an enclosure, it is of particular interest to investigate the behavior near the resonant frequencies of the enclosure. The impact of the aperture is not critical for resonances when its dimension is much smaller than the volume of the enclosure, thus not disturbing EM field distribution inside the enclosure. Since the PCB causes differences in frequency values and peak field magnitudes, the modeling of PCB elements is essential in enclosed environment simulations. Therefore, numerical results of resonant frequencies in the modeled closed environment structure with the PCB are analyzed. Fig. 4 presents the TLM simulation results of resonant frequencies obtained from the vertical electric field sampled above the PCB, at point z = 35mm, corresponding to the center of the aperture in xy plane (x = 142mm, y = 75mm). In order to illustrate effect of a presence of a PCB representing amplifier matching circuits, obtained results of resonances, in Table I, are compared with corresponding values based on simulations and measurements of basic L-track PCB in the same enclosure [5].

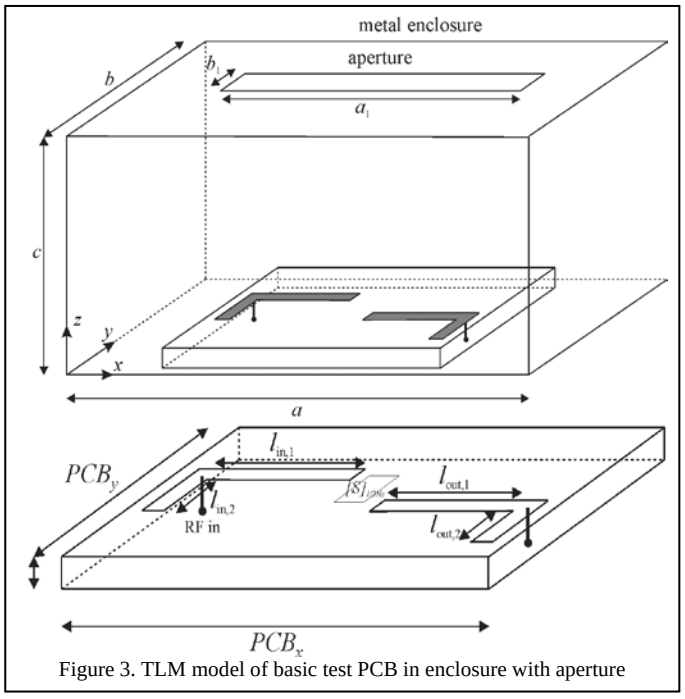


Figure 3. TLM model of basic test PCB in enclosure with aperture

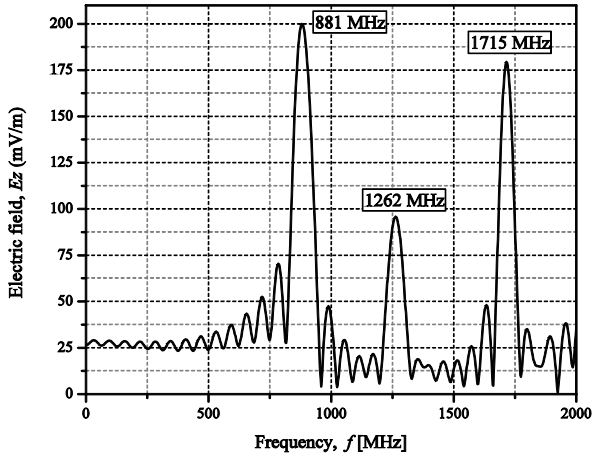


Figure 4. TLM numerical results of vertical component of electrical field from power amplifier PCB in enclosure with aperture

Fig. 5 shows the patterns based on the simulation results of an vertical electric field component at enclosure resonances, in horizontal xy plane, sampled at z = 35mm above the PCB representing EM emissions inside the enclosure. The obtained results illustrate change of EM field distribution of an enclosure due to the physical presence of the PCB elements, in respect to corresponding results based on equivalent dipole simulations and measurements [5].

Besides enclosure resonances, an attention is particularly put on EM emissions at the power amplifier operating frequency. Fig. 6 shows the simulation results of electric field component at 1 GHz frequency, in vertical xz plane, sampled at y = 142 mm (corresponds to aperture position) representing EM emissions inside and outside enclosure. Also, results of the vertical field component at the operating frequency are sampled in horizontal xy plane, at 15 mm above the aperture, representing EM emissions outside enclosure at 1 GHz. Since the results of the vertical field component at the operating frequency are sampled inside and outside the enclosure, TLM mesh is extended to the space above the enclosure top wall.

Generally, the patterns representing EM emissions inside an enclosure are dominantly determined by positions of wire ports and microstrip tracks. It can be seen from Fig. 6. that emissions outside of the enclosure are much smaller compared to corresponding levels inside the enclosure. Also, the emissions from lines representing output matching circuits are much higher than input, due to the level of a signal amplified by the transistor. Besides wire elements and microstrip lines, emissions outside the enclosure are also determined by the aperture position.

TABLE I. RESONANCES OF PCB IN THE ENCLOSURE

PCB in enclosure	**Measured Basic PCB [5]**	**TLM simulation**	
		Basic PCB	*Amplifier PCB*
Resonant frequencies (MHz)	900	903	881
	1290	1285	1262
	1740	1749	1715

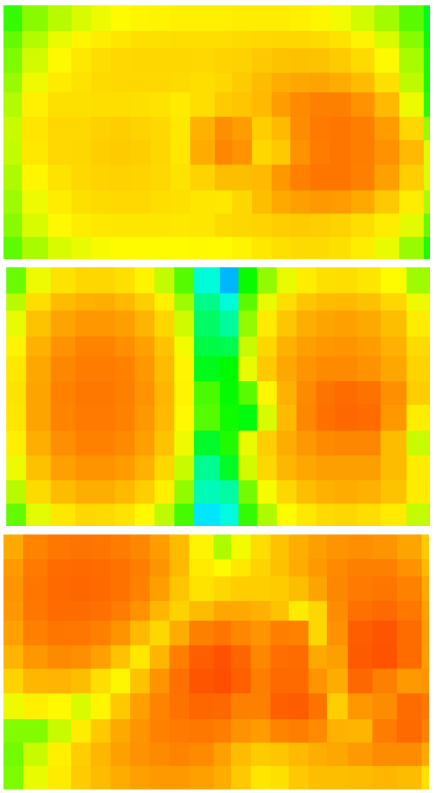


Figure 5. Patterns of Ez given by the TLM simulation of power amplifier PCB, at resonances of enclosure

IV. CONCLUSION

Starting that one of the main interests in EMC tests is the intensities and distributions of the radiated fields from equipment under test (EUT), results are presented here of the EM emissions from a PCB structure representing power amplifier matching circuits. The coax ports and transistor model are used to account for the interactions between the physical presence of the PCB elements. Also, an enclosure should be taken into account when outdoor emission EMC compliance test of PCB is conducted. A method applied to determine radiated emissions from a PCB is based on the TLM model of a board placed in an enclosure with an aperture.

The impact of presence of particular elements of PCB in an enclosure with an aperture is analysed through comparing values of resonances obtained using TLM simulation with reference values of enclosure resonances. The patterns of EM emissions at the amplifier operating frequency inside and outside the enclosure with the aperture are presented and an impact of particular elements on EM emissions is analysed.

The simulation results of basic amplifier boards show that the inclusion of basic features, such as the microstrip track of matching circuits and wire ports, in addition to transistor model, enables accurate prediction of emitted fields, inside the enclosure, that interacts with the PCBs inside. In overall, it is demonstrated that the TLM method have the potential to characterize emissions from PCB structures in realistic environments such as enclosures with an aperture and making it possible to perform system EMC studies.

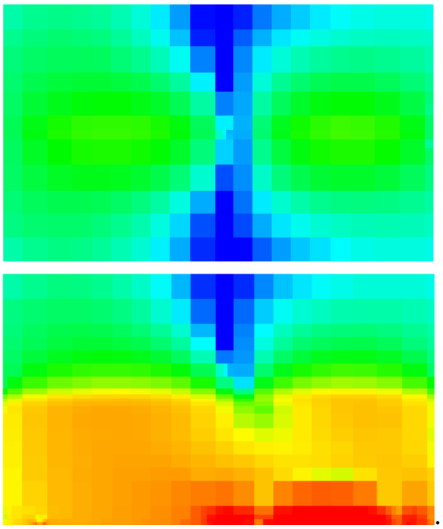


Figure 6. Patterns of Ez given by the TLM simulation of PCB in the enclosure with the aperture, at the operating frequency: a) in a horizontal plane outside the enclosure b) in a vertical plane

ACKNOWLEDGEMENT

This work was supported by Ministry of Education, Science and Technology development of Republic of Serbia, under the project III-44009.

REFERENCES

[1] Christopoulos, C., *Principles and Techniques of Electromagnetic Compatibility*, 2nd edition, CRC Press, Boca Raton, FL, 2007.

[2] Kunz, K. S., Luebbers, R. J., *The Finite Difference Time Domain Method for Electromagnetics*, CRC Press, Boca Raton, FL, 1993.

[3] Christopoulos, C., *The Transmission-Line Modelling (TLM) Method*, IEEE Press / Oxford University Press, Piscataway, NJ, 1995.

[4] Balanis, C. A., *Antenna Theory - Analysis and Design*, John Wiley and Sons, New York, 1997.

[5] Tong, X., Thomas, D.W.P., Nothofer, A., Sewell, P., Christopoulos, C., "Modeling Electromagnetic Emission From Printed Circuit Boards in Closed Environment Using Equivalent Dipoles", *IEEE Transactions on Electromagnetic Compatibility*, Vol. 52, No. 2, pp. 462-470, 2010.

[6] Wlodarczyk, A. J., Trenkic, V., Scaramuzza, R., Christopoulos, C., "A Fully Integrated Multiconductor Model For TLM", IEEE Transactions on Microwave Theory and Techniques, Vol. 46, No. 12, pp. 2431-2437, 1998.

[7] Trenkic, V., Scaramuzza, R., "Modelling of Arbitrary Slot Structures Using Transmission Line Matrix (TLM) Method", *International Symposium on Electromagnetic Compatibility*, Zurich, Switzerland, pp. 393-396, 2001.

[8] Dončov, N., Wlodarczyk, A. J., Scaramuzza, R., Trenkic, V., "Compact TLM Model of Air-vents", *Electronics Letters, Vol. 38, No. 16*, 2002, pp. 887-888.

[9] Bratislav Milovanović, Jugoslav Joković, Nebojša Dončov, "Modelling of Printed Circuit Boards in Closed Environment Using TLM Method", *Proc. of the SSSS 2012 Conference*, Niš, Serbia, pp. 93-96, February 2012.

[10] Jugoslav Joković, Nebojša Dončov, Bratislav Milovanović, Tijana Dimitrijević, "Analysis of Outdoor Emissions from Printed Circuit Board Enclosed in Metallic Box with Aperture", *Proc. of the SSSS 2014 Conference*, Niš, Serbia, pp. 47-50, February 2014.

[11] Aleksandar Atanasković, Nataša Maleš-Ilić, Bratislav Milovanović: "Linearization of power amplifiers by second harmonics and fourth-order nonlinear signals", *Microwave and Optical Technology Letters*, Wiley Periodicals, Inc., A Wiley Company, Vol.55, Issue 2, pp.425-430, 2013.

Level Crossing Rate of Macrodiversity System Operating over Gamma Shadowed Rician Fading Channel

Stefan Panic
Department of Information Technology
Faculty of Natural Science and Mathematics
Kosovska Mitrovica, Serbia
stefanpnc@yahoo.com

Djoko Bandjur, Branimir Jaksic
Department of Electrical and Computer Engineering
Faculty of Technical Sciences
Kosovska Mitrovica, Serbia
dbandjur@gmail.com, branimir.jaksic@pr.ac.rs

Ivana Dinic, Dejan Jaksic, Srboljub Zdravkovic
Department of Telecommunications
Faculty of Electronic Engineering
Nis, Serbia
i.dinic648@gmail.com, djaksic@yahoo.com

***Abstract*—Macrodiversity system with macrodiversity selection combining (SC) receiver and two microdiversity SC receivers is considered. Received signal is affected simultaneously to Gamma long term fading and Rician short term fading resulting in system performance degradation. Macrodiversity SC receiver reduces Gamma shadowing effects and microdiversity SC receivers reduce multipath fading effects on bit error probability. Closed form expression for average level crossing rate of macrodiversity SC receiver output signal envelope is evaluated. Numerical results are presented graphically to show the influence of Gamma long term fading severity, shadowing correlation and Rician factor on average level crossing rate.**

Keywords- fading channel; Gamma shadowed; Rician factor; selection combining (SC).

I. INTRODUCTION

Large scale fading and small scale fading degrade system performance and limit channel capacity wireless mobile communication. Reflections and refractions cause multipath propagation resulting in signal envelope variation and large obstacles cause shadowing resulting in signal envelope average power variation. Received signal experiences simultaneously long term fading and short term fading resulting in system performance degradation. There are several statistical models describing signal envelope variation in multipath fading channels depend on existence line-of-sight (LOS) components, nonlinearity of propagation channel, the number of clusters in propagation environment and variation of signal envelope power. The most frequently used statistical models which can be used to describe multipath fading are Rayleigh, Rician, Nakagami-m, Weibull, α-μ and Nakagami-q distributions. Rayleigh distributions can be used to describe small scale signal envelope variation in linear non line-of-sight multipath fading environments. In line-of-sight short term fading channels signal envelope variation can be described by using Rician distribution. Nakagami-m distribution describes multipath fading in environments with multipath scattering with large delay-time spreads and different clusters of reflected waves, while Weibull and α-μ distribution describe small scale fading in nonlinear channels. Large scale fading can be described by using log-normal distribution or Gamma distribution. When log-normal distribution describes large scale signal envelope average power, the expression for probability density function of macrodiversity receiver output signal can not be obtained in closed form. Theoretical and measured results are shown that Gamma distribution which is analytically better tractable, closely approximates the log-normal distribution in wide range of propagation conditions, and has a good fit to experimental details. When multipath fading is superimposed on shadowing, the instantaneous composite multipath/shadowed signal should be analyzed at the receiver, with Gamma distribution used for for shadowing model [1-3].

Macrodiversity system can be used to simultaneously reduce long term fading effects and short term fading effects on system performance. Macrodiversity system has macrodiversity receiver and two or more micoridiversity receivers. Macrodiversity receiver reduces long term fading effects and microdiversity receivers mitigate short term fading effects on bit error probability. The second order performance measures of wireless communication system are average level crossing rate and average fade duration. Average level crossing rate can be calculated as average value of the first derivative of random process. Average fade duration is equal to ratio of outage probability and average level crossing rate. Outage probability is defined as probability that output signal envelope falls below outage threshold [4].

There are several combining techniques which can be used to reduce long term fading effects and short term fading effects on average level crossing rate and average fade duration. The most frequently used combining techniques are maximal ratio combining (MRC), equal gain combining (EGC), selection combining (SC) and switch and stay combining (SSC). The MRC receiver enables the best results and it has highest implementation complexity. The SC receiver provides the least complexity of realization [5].

There are more works in open technical literature considering performance analysis of macrodiversity systems [6-8]. In [6], macrodiversity SC receiver with two microdiversity MRC receivers operating over Gamma shadowed Nakagami-m multipath fading environment is analyzed. Closed form expressions for average level crossing rate and average fade duration are calculated. Second order performance measures of macrodiversity system in the presence of Gamma shadowed and Rician multipath fading are analyzed.

In this paper macrodiversity system with macrodiversity SC receiver and two microdiversity SC receivers operating over composite shadowed multipath fading channel is considered. Received signal experiences correlated Gamma long term fading and Rician short term fading. Average level crossing rate of Rician random process and SC receiver output signal envelope are calculated. These expressions are used for evaluation average level crossing rate of macrodiversity SC receiver output signal envelope. To the best author's knowledge the second order statistics of macrodiversity SC receiver with two microdiversity SC receivers in the presence Gamma large scale fading and Rician small scale fading is not reported in open technical literature. Obtained results can be used in performance analysis of wireless communication system when received signal is subjected simultaneously to multipath fading and shadowing.

II. Rician Random Process Average Level Crossing Rate

Rician distribution can be used to describe large scale signal envelope variation in line-of-sight multipath fading channels with one cluster. Squared Rician random variable, x, is equal to sum of two squared independent Gaussian random variables with the same variance:

$$x^2 = x_1^2 + x_2^2, \tag{1}$$

where x_1 and x_2 are independent Gaussian random variable with the same variance σ^2. The first derivation of Rician random variable x is

$$\dot{x} = \frac{1}{x}\left(x_1\dot{x}_1 + x_2\dot{x}_2\right). \tag{2}$$

The first derivative of Gaussian random variable is Gaussian random variable. Thus $\dot{x}_1$ and $\dot{x}_2$ are zero-mean Gaussian random variables. Linear transform of Gaussian random variables is Gaussian random variable. Therefore, $\dot{x}$ follow Gaussian distribution. Mean of $\dot{x}$ is

$$\overline{\dot{x}} = \frac{1}{x}\left(x_1\overline{\dot{x}_1} + x_2\overline{\dot{x}_2}\right) = 0, \tag{3}$$

since $\overline{\dot{x}_1} = \overline{\dot{x}_2} = 0$.

The variance of the first derivate of Rician random variable is

$$\sigma_{\dot{x}}^2 = \frac{1}{x^2}\left(x_1^2\sigma_{\dot{x}_1}^2 + x_2^2\sigma_{\dot{x}_2}^2\right), \tag{4}$$

where [10]:

$$\sigma_{\dot{x}_1}^2 = \sigma_{\dot{x}_2}^2 = \pi^2 f_m^2 2\sigma^2, \tag{5}$$

where f_m is maximal Doppler frequency. After substituting (5) in (4), the variance of the first derivative of Rician random variable becomes

$$\sigma_{\dot{x}}^2 = \frac{1}{x^2}\pi^2 f_m^2 2\sigma^2\left(x_1^2 + x_2^2\right) = 2\pi^2 f_m^2\sigma^2. \tag{6}$$

The joint probability density function of Rician random variable and the first derivative of Rician random variable is

$$p_{x\dot{x}}\left(x\dot{x}\right) = p_x\left(x\right)p_{\dot{x}}\left(\dot{x}\right), \tag{7}$$

where $p_x(x)$ is probability density function of x:

$$p_x\left(x\right) = \frac{2\left(1+k\right)}{e^k\Omega}e^{-\frac{(1+k)x^2}{\Omega}}I_0\left(2\sqrt{\frac{k\left(1+k\right)}{\Omega}}x\right), \tag{8}$$

where k is Rician factor. Rician factor can be evaluated as ratio of dominant component power to scattering component power. The level crossing rate of Rician random process can be calculated as average value of the first derivative of Rician random variable:

$$N_x = \int_0^\infty d\dot{x}\cdot\dot{x}\cdot p_{x\dot{x}}\left(x\dot{x}\right) = p\left(x\right)\frac{1}{\sqrt{2\pi}}\sigma_{\dot{x}} = \\ = \frac{f_m\sqrt{2}\left(1+k\right)}{e^k\Omega^{\frac{1}{2}}}e^{-\frac{(1+k)x^2}{\Omega}}I_0\left(2\sqrt{\frac{k\left(1+k\right)}{\Omega}}x\right). \tag{9}$$

Expression for level crossing rate can be used for evaluation average fade duration of wireless communication system operating over Rician multipath fading channel. The cumulative distribution function of Rician distribution is

$$F_x(x)=\int_0^x p_x(t)\,dt= \\ =\frac{1}{e^k}\sum_{i_1=0}^{\infty}\frac{k^{i_1}}{(i_1!)^2}\gamma\left(i_1+1,\frac{(k+1)x^2}{\Omega}\right) \quad (10)$$

Obtained expression for cumulative distribution function of Rician random variable can be used in performance analysis of wireless communication system in the presence Rician short term fading.

The average level crossing rate of dual SC receiver operating over independent identical Rician multipath fading output signal envelope is

$$N_x=2F_{x_1}(x)N_{x_1}, \quad (11)$$

where N_{x1} is given with (9) and $F_{x1}(x)$ is given with (10). The expression can be used for calculation average fade duration of communication mobile system with SC reception operating over Rician multipath fading channel.

III. Level Crossing Rate of Macrodiversity SC Receiver Output Signal Envelope

Macrodiversity system with macrodiversity SC receiver and two microdiversity SC receivers operating over composite Gamma shadowed Rician multipath fading environment is considered. Microdiversity SC receiver is provided by using multiple antennas at base station resulting in reduction of Rician multipath fading effects and macrodiversity SC receiver uses signals from two or more base stations resulting in reduction Gamma long term effects on system performance. Signal envelopes at output of microdiversity SC receivers are denoted with x_1 and x_2. Macrodiversity SC receiver output signal envelope is denoted with x.

Signal envelope average power at inputs in microdiversity SC receivers are denoted with Ω_1 and Ω_2. Random variable Ω_1 and Ω_2 follow correlated Gamma distribution:

$$p_{\Omega_1\Omega_2}(\Omega_1\Omega_2)=\frac{1}{\Gamma(c)(1-\rho^2)\rho^{\frac{c-1}{2}}\Omega_0^{c+1}} \\ \times\sum_{i_2=0}^{\infty}\left(\frac{\rho}{\Omega_0(1-\rho^2)}\right)^{2i_2+c-1}\frac{1}{i_2!\Gamma(i_2+c)} \\ \times\Omega_1^{i_2+c-1}\Omega_2^{i_2+c-1}e^{-\frac{\Omega_1+\Omega_2}{\Omega_0(1-\rho^2)}}, \quad (12)$$

where c is order of Gamma distribution, ρ is correlation coefficient and Ω_o is average power of Ω_1 and Ω_2.

Macrodiversity SC receiver selects microdiversity SC receiver with higher signal envelope average power at inputs to provide service to users. Therefore, average level crossing rate of macrodiversity SC receiver output signal envelope is

$$N_x=\int_0^{\infty}d\Omega_1\int_0^{\Omega_1}d\Omega_2 N_{x_1/\Omega_1}p_{\Omega_1\Omega_2}(\Omega_1\Omega_2)+ \\ +\int_0^{\infty}d\Omega_2\int_0^{\Omega_2}d\Omega_1 N_{x_2/\Omega_2}p_{\Omega_1\Omega_2}(\Omega_1\Omega_2), \quad (13) \\ =2\int_0^{\infty}d\Omega_1\int_0^{\Omega_1}d\Omega_2 N_{x_1/\Omega_1}p_{\Omega_1\Omega_2}(\Omega_1\Omega_2)$$

where $N_{x1/\Omega1}$ and $N_{x2/\Omega2}$ is given with (11). By substituting (9), (10) and (12) in (13), the expression for average level crossing rate becomes

$$N_x=\frac{4\sqrt{2}f_m(1+k)}{e^{2k}\Gamma(c)(1-\rho^2)\rho^{\frac{c-1}{2}}\Omega_0} \\ \times\sum_{i_1=0}^{\infty}\frac{k^{i_1}}{(i_1!)^2}\sum_{i_3=0}^{\infty}\frac{(k(1+k))^{i_3}x^{2i_3}}{(i_3!)^2} \\ \times\sum_{i_2=0}^{\infty}\left(\frac{\rho}{\Omega_0(1-\rho^2)}\right)^{c+2i_2-1}\frac{(\Omega_0(1-\rho^2))^{i_2+c}}{i_2!\Gamma(i_2+c)} \\ \times\int_0^{\infty}d\Omega_1\Omega_1^{i_2-i_3+c-\frac{3}{2}}e^{-\frac{\Omega_1}{\Omega_0(1-\rho^2)}-\frac{(1+k)x^2}{\Omega_1}} \\ \times\gamma\left(i_1+1,\frac{(k+1)x^2}{\Omega_1}\right)\gamma\left(i_2+1,\frac{\Omega_1}{\Omega_0(1-\rho^2)}\right)= \\ =\frac{4\sqrt{2}f_m(1+k)}{e^{2k}\Gamma(c)(1-\rho^2)\rho^{\frac{c-1}{2}}\Omega_0}\sum_{i_1=0}^{\infty}\frac{k^{i_1}}{(i_1!)^2}\sum_{i_3=0}^{\infty}\frac{x^{2i_3}}{(i_3!)^2} \\ \times\sum_{i_2=0}^{\infty}\left(\frac{\rho}{\Omega_0(1-\rho^2)}\right)^{c+2i_2-1}\frac{(\Omega_0(1-\rho^2))^{i_2+c}}{i_2!\Gamma(i_2+c)}\times I, \quad (14)$$

where $\gamma(n,x)$ [9] is incomplete Gamma function and I is

$$I=\int_0^{\infty}d\Omega_1\Omega_1^{i_2-i_3+c-\frac{3}{2}}e^{-\frac{\Omega_1}{\Omega_0(1-\rho^2)}-\frac{(1+k)x^2}{\Omega_1}} \\ \left(\frac{1}{i_1+1}\left(\frac{(k+1)x^2}{\Omega_1}\right)^{i_1+1}\right. \\ \left.\times e^{-\frac{(k+1)x^2}{\Omega_1}}\sum_{j_1=0}^{\infty}\frac{1}{(i_1+2)_{(j_1)}}\left(\frac{(k+1)x^2}{\Omega_1}\right)^{j_1}\right)$$

$$\times\left(\frac{1}{i_2+1}\left(\frac{\Omega_1}{\Omega_0\left(1-\rho^2\right)}\right)^{i_2+1}\right.$$
$$\left.\times e^{-\frac{\Omega_1}{\Omega_0\left(1-\rho^2\right)}}\sum_{j_2=0}^{\infty}\frac{1}{\left(i_2+2\right)_{(j_2)}}\left(\frac{\Omega_1}{\Omega_0\left(1-\rho^2\right)}\right)^{j_2}\right)= \quad (15)$$
$$= I_1 - I_2 - I_3 + I_4$$

where $(a)_{(n)}$ denoting the Pocchammer symbol.

The integral I_1 can be solved as:

$$I_1 = \Gamma\left(i_1+1\right)\Gamma\left(i_2+1\right)$$
$$\times\left(\left(k+1\right)x^2\Omega_0\left(1-\rho^2\right)\right)^{\frac{i_2}{2}-\frac{i_3}{2}+\frac{c}{2}-\frac{1}{4}}. \quad (16)$$
$$\times K_{i_2-i_3+c-\frac{1}{4}}\left(2\sqrt{\frac{\left(k+1\right)x^2}{\Omega_0\left(1-\rho^2\right)}}\right)$$

The integral I_2 can be solved as:

$$I_2 = \frac{\Gamma\left(i_1+1\right)}{i_2+1}\sum_{j_2=0}^{\infty}\frac{1}{\left(i_2+2\right)_{(j_2)}}$$
$$\times\frac{1}{\left(\Omega_0\left(1-\rho^2\right)\right)^{i_2+j_2+1}} \quad . \quad (17)$$
$$\times\left(\frac{\left(k+1\right)x^2\Omega_0\left(1-\rho^2\right)}{2}\right)^{i_2-\frac{i_3}{2}+\frac{j_2}{2}+\frac{c}{2}+\frac{1}{4}}$$
$$\times K_{2i_2-i_3+j_2+c+\frac{1}{2}}\left(2\sqrt{\frac{2\left(k+1\right)x^2}{\Omega_0\left(1-\rho^2\right)}}\right)$$

The integral I_3 can be solved as:

$$I_3 = \frac{\Gamma\left(i_2+1\right)}{i_1+1}\sum_{j_1=0}^{\infty}\frac{\left(\left(k+1\right)x^2\right)^{i_1+j_1+1}}{\left(i_1+2\right)_{(j_1)}}$$
$$\times\left(2\left(k+1\right)x^2\Omega_0\left(1-\rho^2\right)\right)^{\frac{i_2}{2}-\frac{i_3}{2}-\frac{j_1}{2}+\frac{c}{2}-\frac{3}{4}}. \quad (18)$$
$$\times K_{i_2-i_3-j_1+c-\frac{3}{2}}\left(2\sqrt{\frac{2\left(k+1\right)x^2}{\Omega_0\left(1-\rho^2\right)}}\right)$$

The integral I_4 can be solved as:

$$I_4 = \frac{1}{i_1+1}\sum_{j_1=0}^{\infty}\frac{\left(\left(k+1\right)x^2\right)^{i_1+j_1+1}}{\left(i_1+2\right)_{(j_1)}}$$
$$\times\frac{1}{i_2+1}\sum_{j_2=0}^{\infty}\frac{1}{\left(i_2+2\right)_{(j_2)}}\frac{1}{\left(\Omega_0\left(1-\rho^2\right)\right)^{i_2+j_2+1}}, \quad (19)$$
$$\times\left(\left(k+1\right)x^2\Omega_0\left(1-\rho^2\right)\right)^{-\frac{i_1}{2}+i_2-\frac{i_3}{2}-\frac{j_1}{2}+\frac{j_2}{2}+\frac{c}{2}-\frac{1}{4}}$$
$$\times K_{-i_1+2i_2-i_3-j_1+j_2+c-\frac{1}{2}}\left(2\sqrt{\frac{4\left(k+1\right)x^2}{\Omega_0\left(1-\rho^2\right)}}\right)$$

where $K_n(x)$ is modified Bessel function of the second kind [9], order n and argument x.

IV. Numerical Results

In Figure 1 and Figure 2, normalized average level crossing rate of macrodiversity SC receiver output signal envelope is presented in the function of system parameters such are Gamma shadowing severity parameter c, correlation coefficient ρ and Rician factor K.

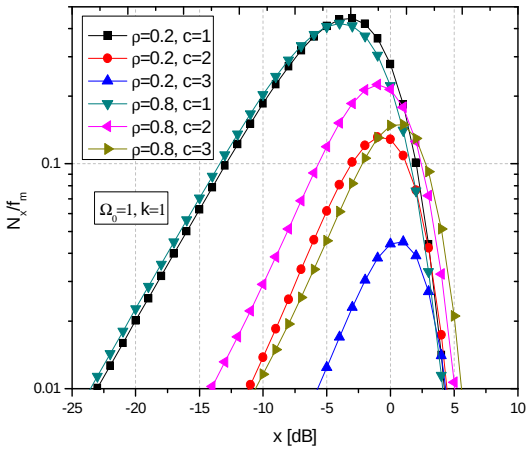


Figure 1. Level crossing rate of macrodiversity SC receiver output signal envelope for different values of Gamma shadowing severity parameter c and correlation coefficient ρ.

For lower values of SC receiver output signal envelope, average level crossing rate increases as signal envelope increases and for higher values of signal envelope, level crossing rate decreases as signal envelope increases.

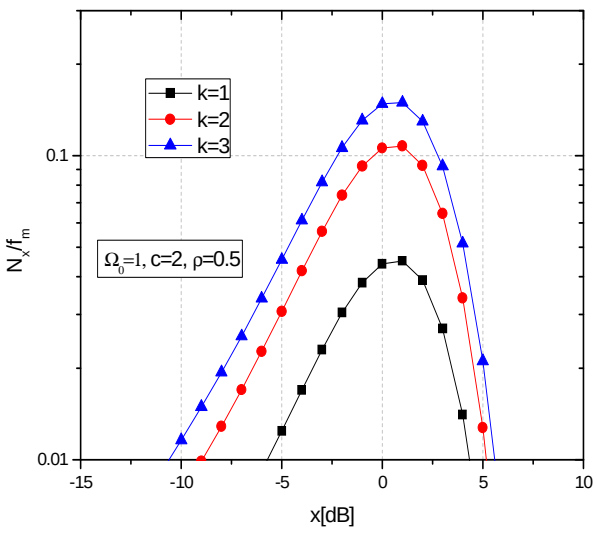


Figure 2. Level crossing rate of macrodiversity SC receiver output signal envelope for different values of Rician factor k.

V. CONCLUSION

SC macrodiversity system with dual SC microdiversity operating over composite shadowed Ricean/Gamma fading channel is considered. Expressions for LCR of the macrodiversity output are presented in closed-form. LCR dependence on observed szstem parameters, such are Gamma shadowing severity parameter c, correlation coefficient ρ and Rician factor K has also been graphicaly presented and discussed.

REFERENCES

[1] G. L. Stuber, Mobile communication, 2nd ed., Dordrecht: Kluwer Academic Publisher, 2003.

[2] M. K. Simon, M. S. Alouini, Digital Communication over Fading Channels, USA: John Wiley & Sons, 2000.

[3] S. Panic, M. Stefanovic, J. Anastasov, P. Spalevic, Fading and Interference Mitigation in Wireless Communications, USA: CRC Press, 2013.

[4] N. Sekulovic, M. Stefanovic, "Performance Analysis of System with Micro- and Macrodiversity Reception in Correlated Gamma Shadowed Rician Fading Channels," Wireless Personal Communications, (publisher: Springer), vol. 65, no. 1, pp. 143-156, 2012.

[5] E. Mekic, N. Sekulovic, M. Bandjur, M. Stefanovic, P. Spalevic, "The distribution of ratio of random variable and product of two random variables and its application in performance analysis of multi-hop relaying communications over fading channels," Przeglad Elektrotechniczny, vol. 88, no. 7A, pp. 133-137, 2012.

[6] D. Stefanovic, S. R. Panic, P. Spalevic "Second-order statistics of SC macrodiversity system operating over Gamma shadowed Nakagami-m fading channels," AEU - International Journal of Electronics and Communications, vo. 65, no. 5, pp. 413–418, 2011.

[7] D. Krstić, M. Stefanović, S. Panić, G. Stamenović, I. Petrović, "Second order statistics of selection macrodiversity system in the presence of Nakagami-*m* fading", XV International Symposium on Theoretical Electrical Engineering, ISTET'09, 22 – 24 June 2009, Lübeck, Germany

[8] Milosevic, B.; Spalevic, P.; Petrovic, M.; Vuckovic, D.; Milosavljevic, S. Statistics of Macro SC Diversity System with Two Micro EGC Diversity Systems and Fast Fading. Electronics and Electrical Engineering. 96, 8(2009), pp. 55-58.

[9] I. Gradshteyn, I, Ryzhik.Tables of Integrals, Series, and products, Academic Press, New York 1994.

[10] Dong, X. and Beaulieu, N. C.: Average Level Crossing Rate and Average Fade Duration of Selection Diversity, *IEEE Communication Letters*. vol. 10, No 5, pp.396-399, (2001)

Implementation of TR-069 connection request mechanism

Ivana Savić, Milan S. Savić, Gordana Velikić
RT-RK Computer Based Systems,
Novi Sad, Serbia
ivana.ostojic@rt-rk.com,
msavic@rt-rk.com
gordana.velikic@rt-rk.com

***Abstract*—This paper proposes and evaluates an approach for monitoring and controlling devices over a network based on TR-069 protocol. Protocol itself relies on communication between server and clients (devices). Main characteristic of this protocol is that all the communication is initiated from the device side. In order to allow server to initiate interaction, a connection request mechanism is defined. It serves as a notification system that informs the device to start interaction with the server. Depending on the network state, available resources at hand and security considerations, there are four types of connection request to choose from: connection request over TCP, connection request over TCP with port forwarding, Session Traversal Utilities for NAT (STUN) based connection request and Extensible Messaging and Presence Protocol (XMPP) based connection request.**

Keywords-TR069, connection request, STUN, XMPP

I. INTRODUCTION

The purpose of this paper is to propose a solution for device monitoring over a network with the emphasis on connection requests from server side. As shown in [1], TR-069 CPE WAN Management Protocol (CWMP) is protocol dedicated to monitoring and management of end-user devices. It enables communication between customer-premises equipment (CPE) and auto configuration server (ACS), as shown in Fig. 1.

Communication between ACS and CPE is initiated exclusively by CPE. Only way for ACS to initiate a connection with CPE is to notify CPE that it should engage in a session. This is achieved with connection request. It represents ping mechanism that allows the ACS to initiate connection. Connection requests are divided primarily into two groups depending on TR-069 protocol version used.

When TR-069 amendment 1 – 4 is used [1] – [4], then connection request that is going to be used is determined by network state. If CPE has a routable IP address TCP based connection request is used. Main characteristic of this kind of connection request is that it is used when ACS and CPE are inside the same address space. When CPE is behind Network Address Translation (NAT) gateway, STUN based connection request is used. Key note for this connection request is that it relies on tunneling mechanism which is created between each CPE and ACS.

Latest version of TR-069 protocol [5] proposes two new types of connection requests, TCP based connection request that relies on port forwarding and it is used in highly controllable network and it is not suited for requests over open internet. Alternative to previous request is XMPP based connection request which is usable in all network architectures with highest security of all proposed solutions.

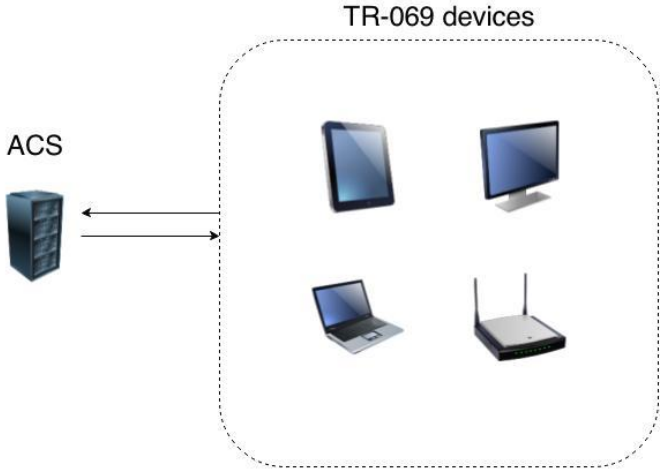


Figure 1. TR-069 Environment

Existing solutions that are based on CWMP protocol [6] don't follow the protocol by the letter, that way making devices that use it not compatible with other TR-069 solutions. A method for establishing connection through NAT is presented in [7], but instead of STUN based approach, TCP connection is used. Both solutions are not standardized and hard to implement in practice.

II. SOLUTION CONCEPT

When ACS has the need to inform CPE that it should engage in a session connection request process is started. Essentially it is an authentication process which depends on type of connection request implemented. In case of TCP connection request, HTTP Digest authentication is used. When STUN based approach is used, custom authentication defined in TR-069 protocol is used. If XMPP Connection request is

used, Transport Layer Security (TLS) authentication is in place.

A. *TCP Connection Request*

This mechanism is used when CPE is not behind NAT gateway as shown in Fig. 2. Authentication method used is HTTP digest [8]. CPE acts like HTTP server and ACS has the client role. Authentication challenge contains data which combined with data from TR-106 datamodel [9] represent all the information necessary for successful authentication. TCP protocol is used for communication between ACS and CPE and for that reason it is important that ACS and CPE are in the same address space, in order to be visible to each other. More important it is essential that CPE is visible from the ACS because server sends connection requests and thus it has to be able to reach client.

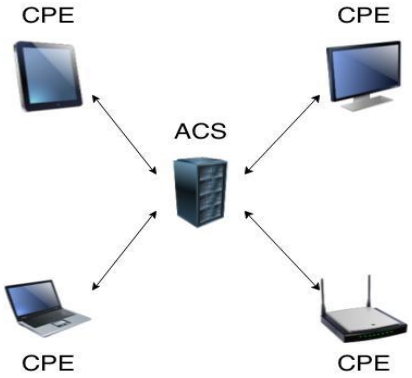


Figure 2. TR-069 environment with ACS and CPE in the same address space

As shown on Fig. 3, successful authentication serves as a signal for the CPE to start a session with CPE. Authentication is based on two steps. First empty GET request is send on which server responds with unauthorized response in which challenge for successful authentication is present. Client resends GET request, but this time with all the necessary data for successful authentication.

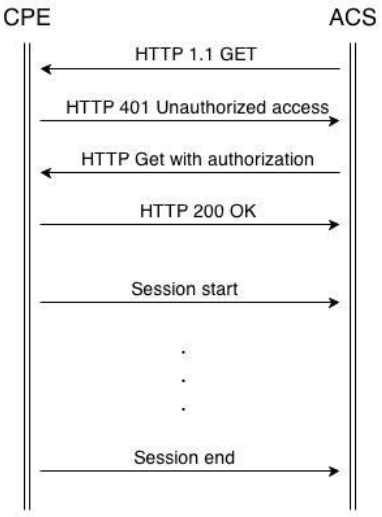


Figure 3. TR-069 connection request mechanism

B. *STUN based connection request*

When end-device is behind a gateway and if resources at hand do not allow use of latest TR-069 protocol version, STUN based connection request is used. It allows ACS to send requests to CPE even if end-device is in private address space. As shown on Fig. 4, STUN based connection request comprises out of:

- STUN server
- STUN client
- UDP server
- UDP client

STUN server is located on ACS side and STUN client on CPE side. Their role in connection request is to make a tunnel between ACS and CPE and that way make direct communication possible.

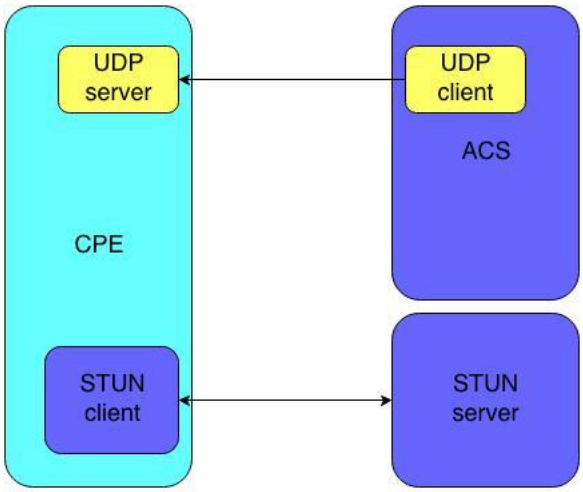


Figure 4. STUN based connection request

The purpose of STUN mechanism is to discover network and port translation and maintain that binding over time [10]. Mechanism is fairly simple, client sends a request to server and in response address and port on which response is send is conveyed. That way client is aware if port and address translation is in progress. If client determines that it is behind a NAT gateway, keep alive mechanism is activated. In this mode client sends messages periodically in order to keep the binding. For that purpose secondary port in STUN mechanism is defined and all the messages in keep alive process are sent from it. In keep alive message address and port on which response should be sent is present and in case that response arrives on that specific port, binding is in place.

Second part of STUN based TR-069 connection request is UDP server and client. It is similar to implementation of TR-069 connection request over TCP with couple of exceptions:

- Instead of TCP, UDP connection is used
- Authentication mechanism isn't HTTP digest, instead custom authentication mechanism is used
- Client doesn't get a response from server, only requests are sent and no response is generated.

C. TCP connection request with port forwarding

Connection request with port forwarding is used when gateway, behind which CPE devices reside, and ACS are in the same address space. That allows ACS to reach end devices through designated port on gateway which is mapped to a port on CPE as shown on Fig. 5.

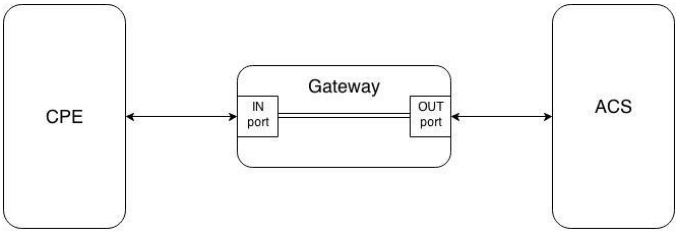


Figure 5. TR-069 connection request with port forwarding

Except this modification concerning network architecture everything is the same as in case of regular TR-069 connection request over TCP without port forwarding.

D. Connection request over XMPP

Connection request over XMPP protocol [11] enables the ACS to reach end device no matter the network architecture as shown on Fig. 6. All the data necessary for successful communication is conveyed through TR-157 datamodel [12]. Connection request architecture can be divided into three parts:

- XMPP client on CPE
- XMPP client on ACS
- XMPP server

Architecture defers from previous connection requests where CPE had the role of server and ACS the role of client. Now both ACS and CPE are clients and their messages are conveyed through XMPP server which is responsible for enabling communication between clients.

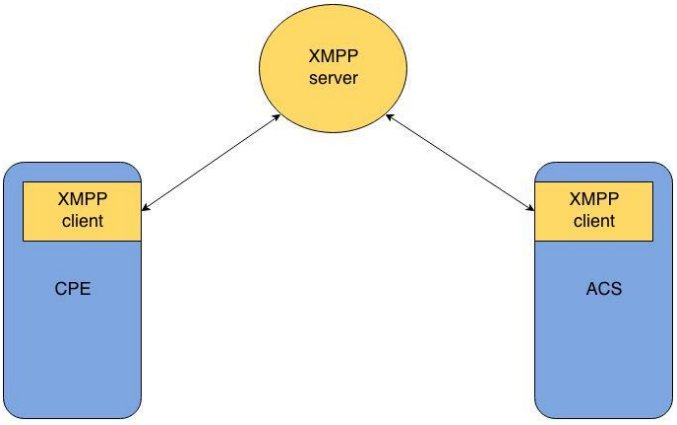


Figure 6. TR-069 XMPP connection request

Connection between XMPP clients and server is based on TCP. Once client is connected on server, connection is kept alive by periodic ping messages from server to clients. That way clients are always reachable and client state (offline or online) can easily be determined. Knowing the client state at any time makes this connection request robust which is not the case for some previous implementations. This kind of connection has it downside in higher data consumption.

III. Experimental Results

The proposed concept was validated on Linux based set-top-box (STB) on client side and for the ACS side Intel based server was used. With connection request ACS gets almost instance response from the CPE, that way making TR-069 system more responsive and agile.

TR-069 connection request over TCP and connection request with port forwarding have the same authentication process and thus have same experimental results, as shown in Table I. Although the network architecture is different, the complexity itself remains almost the same. Only difference is that port forwarding is used when gateway is between CPE and ACS. That doesn't add up much to network complexity and for that reason it doesn't add any delay to the system response.

TABLE I. TR-069 Connection Request

Test	Value
Channel change	~2.8 s
Authentication time	~50 ms

TR-069 request over STUN is used when gateway or numerous gateways are between CPE and ACS. This kind of connection request is best suited when connection between ACS and CPE is conveyed through open internet. Due to custom authentication process which is simpler than HTTP digest mechanism used in connection request over TCP, system response time is faster, as shown in Table II.

TABLE II. TR-069 STUN Connection Request

Test	Value
Channel change	~2.1 s
Authentication time	~18 ms

Connection request over XMPP can be used on any network configuration. Because of the architecture of the XMPP server/client communication, authentication between client and server is done only once upon authentication with XMPP server, that way every time the connection request is sent over XMPP there is no authentication process. There is only the time it takes a message to be sent from one XMPP client to another. As shown in Table III, system with XMPP connection request is highly responsive and gives the best results out of all proposed connection request mechanisms.

TABLE III. TR-069 XMPP CONNECTION REQUEST

Test	***Value***
Channel change	~2 s
Authentication time	0 s
Message transfer time	2-7 ms

IV. CONCLUSION

Connection request enables instantaneous system response from end-device. No matter what network architecture is, end-device is always reachable from the ACS. Connection request mechanism is highly usable when TR-069 protocol is used in home automation. In smart house solutions user has the need to interact with his household over the Internet. In order to achieve comfortable usage system has to be responsive and enable the user to experience almost real time monitoring with data available right. Only concern regarding TR-069 protocol is privacy. User data resides on server which is outside of users reach and therefore making private data vulnerable. The solution might be in small household systems where the ACS would reside inside a house and would be used to monitor and control household appliances. Without a third person having access to your personal data, privacy issue is basically non-existing.

ACKNOWLEDGMENT

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, under grant number TR 32034.

REFERENCES

[1] DSL Forum, "TR-069 CPE WAN Management Protocol v1.1", November 2006.

[2] DSL Forum, "TR-069 CPE WAN Management Protocol v1.2", December 2007.

[3] DSL Forum, "TR-069 CPE WAN Management Protocol v1.3", November 2010.

[4] DSL Forum, "TR-069 CPE WAN Management Protocol v1.3", July 2011.

[5] DSL Forum, "TR-069 CPE WAN Management Protocol v1.4", November 2013.

[6] Tiago Cruz, Paulo Simões, Patrício Batista, João Almeida, Edmundo Monteiro, "CWMP Extensions for Enhanced Management of Domestic Network Services", LCN 2010, Denver, Colorado.

[7] Jeffrey L. Eppinger," TCP Connections for P2P Apps: A Software Approach to Solving the NAT Problem", Institute for Software Research International, School of Computer Science, Carnegie Mellon University, 2005.

[8] "HTTP Authentication: Basic and Digest Access Authentication", IETF RFC 2617, 1999.

[9] DSL Forum, "TR-106 Amendment 1 Data Model Template for TR-069 –Enabled Devices", November 2006.

[10] "STUN – Simple Traversal of User Datagram Protocol through Network Address Translators", IETF RFC 3489, March 2003.

[11] "Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence", IETF RFC 6121, March 2011.

[12] DSL Forum, "TR-157 Amendment 5 Component Objects for CWMP", November 2011.

Comparison of Models for Self-Similar Network Traffic Generation

Igor Tomić, Nebojša Maletić
Faculty of Electrical Engineering
Banja Luka, Bosnia and Herzegovina
igor.tomic@etfbl.net, nebojsa.maletic@etfbl.net

***Abstract*—This paper gives an overview of five different models for self-similar network traffic generation (iterated chaotic maps, fractional Gaussian noise model, Pareto model and finite and infinite Markov chain model). The models are compared on the basis of Hurst parameter and mean value of generated sequences, and also on the basis of algorithm efficiency. R/S plot, Variance-Time plot and Periodogram method are used for the Hurst parameter estimation. According to simulation results, the model which gives sequences whose parameters are close enough to the given ones is fronted.**

Keywords-Hurst parameter; long-range dependence; self-similarity; traffic model

I. INTRODUCTION

In modern packet-oriented networks, telecommunication traffic is considered to be self-similar process. Unlike the Poisson process (used for classic telephone traffic modelling), which get smooth when averaged over large timescales, self-similar process retains its burstiness over a wide range of time scales [1].

There are several factors contributing to self-similarity of network traffic, such as link bandwidth [2], file size distribution, reliability and flow control mechanisms in the transport layer [3], VBR (Variable Bit Rate) video streaming [4] and others, so that the degree of traffic burstiness differs between network segments.

A number of papers have studied the impact of self-similarity on network performance in terms of packet loss rate, queueing delay or throughput [2], [5], [6]. Synthesized self-similar traffic is important in evaluating the performance of various switch architectures under realistic conditions. A good traffic model may lead to a better designing routers and network devices which handle long packet bursts. According to differences in the traffic statistic within different parts of the network, the generator is recquired to be flexible enough.

A lot of self-similar traffic generator models are developed, but none of them simulates all aspects of real traffic. This paper provides comparison of some of the most referred models.

II. PARAMETERS OF SELF-SIMILAR PROCESSES

In this section, an overview of basic concepts and terms of self-similar network traffic is given.

As a measure of self-similarity the *Hurst* parameter (H) is used, where higher H ($H\in(1/2,1)$) implies higher self-similarity.

The packet traffic trace can be represented by binary sequence, where one stands for a packet and zero for an interpacket gap. If the sequence x is divided into non-overlapping adjacent blocks of size m and then blocks are averaged, m-th order aggregation of the sequence is given as

$$x_t^{(m)} = \frac{1}{m}\sum_{i=tm-m+1}^{tm} x_i \ . \qquad (1)$$

Two traffic traces with the aggregation scale m=1000 for different values of H are shown in Fig. 1. It can be noted that process with lower H becomes smooth when aggregated over a large scale.

A stochastic process is considered to be self-similar with parameter β, 0<β<1, if, for all m=1, 2, ..., the following applies:

$$Var\left(x^{(m)}\right) \sim \frac{Var(x)}{m^{\beta}} . \qquad (2)$$

The parameter β is related to H by H=1–β/2. The variance of the process with higher H decays more slowly with the increase of aggregation degree.

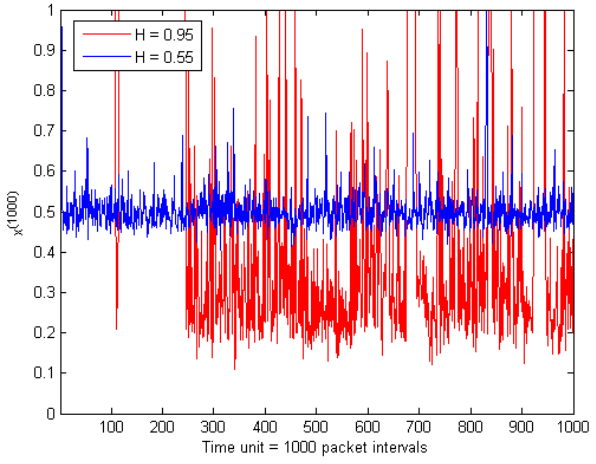


Fig. 1. Traces with the aggregation scale m=1000 for different values of H.

If distant time samples of the process are correlated, it is said to be LRD (*Long-Range Dependence*) process. The LRD process has a hyperbolically decaying autocovariance function:

$$C(\tau) \sim |\tau|^{-\beta}, \text{ as } |\tau| \to \infty, 0 < \beta < 1. \quad (3)$$

The important characteristic of the self-similar process is that packet train and interpacket gap duration have a heavy-tailed distribution. One of the most frequently used heavy tailed distributions is the Pareto with cumulative distribution function (CDF) given by:

$$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}, (x \geq \beta, \alpha > 0). \quad (4)$$

Parameter α, 1<α<2, denotes the tail index.

III. SELF-SIMILAR TRAFFIC GENERATION

In this section, five models for self-similar traffic generation, frequently found in literature, are presented.

A. Iterated Chaotic Maps

The iterated chaotic map (ICM) model is described by traffic load $d \in (0, 1)$ and parameters m_1, $m_2 \in (1.5, 2)$, which determine the slope of curve plotted in Fig. 2.

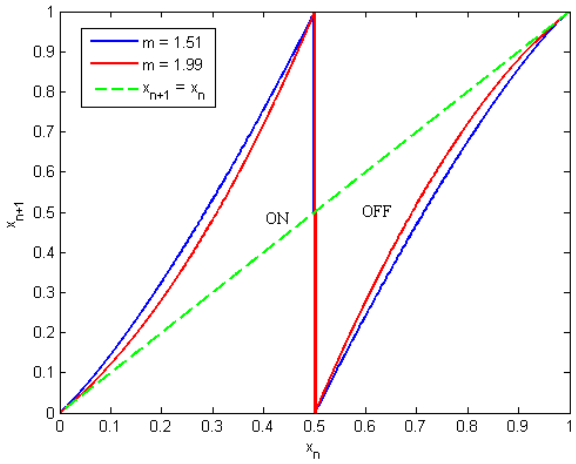


Fig. 2. Graph of an iterated chaotic map for d=0.5.

Based on initial value $x_0 \in (0, 1)$, the next one is given by:

$$x_{n+1} = \begin{cases} x_n + \dfrac{1-d}{d^{m_1}} x_n^{m_1}, & 0 < x_n < d \\ x_n - \dfrac{d}{(1-d)^{m_2}} (1 - x_n)^{m_2}, & d < x_n < 1 \end{cases}. \quad (5)$$

The sequence x takes values between 0 and 1. The binary time series $\{y_n: n \in N\}$ is generated by applying the following rule: $y_i = 1$ if $x_i < d$ and $y_i = 0$, otherwise. If $m_1 = m_2 = m$, the relation with H is given by $H = (3m–4)/(2m–2)$. Thus, parameters d and H describe the model.

When close to the mean value d, short sequences of zeros and ones are generated. Long sequences are generated when x_n is found near points zero or one, by which traffic burstiness is obtained. A detailed explanation about the model can be found in [7].

B. Pareto Traffic Generation

The starting model hypothesis is that the time of packet train and interpacket gaps have a Pareto distribution (Fig. 3).

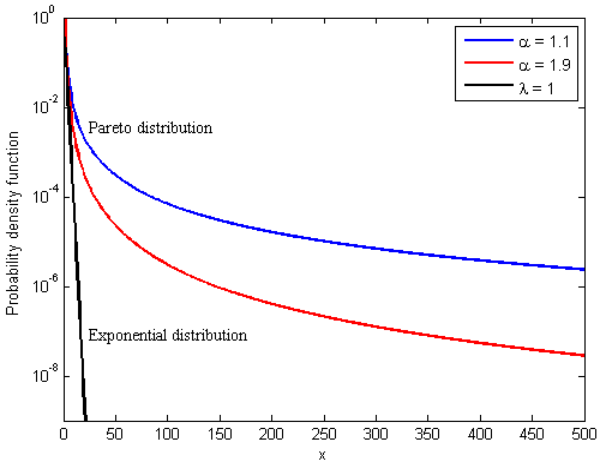


Fig. 3. Pareto vs exponential density distribution function.

Two random variables are required to generate binary sequences. One variable stands for packet sequence duration, and the other for interpacket gap duration measured in time intervals equal to packet duration. As $x \geq \beta$ (eq. 4), parameter β presents the shortest length of the packet train. Tail index, α (1<α<2), defines how fast CDF decays. As $\alpha \to 1$, the rate of decay is low, indicating that appearance of longer bursts is more probable. Poisson distribution (used for classic voice traffic modelling) decays much faster (exponential) compared to Pareto distribution.

The parameter H depends on both parameter value, α and β. Larger β, as well as α closer to unity, causes longer bursts. Anyway, it was showed that the effect of β is not so influential on the degree of self-similarity as α is. The relation with H is H=(3–α)/2. Details of the algorithm can be found in [6].

C. FGN Traffic Generation

The Fractional Gaussian noise (FGN) traffic generator produces the streams of variable-length packets with self-similar statistics, unlike Pareto generator, which gives fixed length packets [6].

The model is based on features of LRD process in the frequency domain. The first step in synthesis of self-similar sequence is the estimation of FGN power spectrum with given Hurst parameter (Fig. 4). The Inverse Fourier Transform (IFFT) is used to obtain the time trace which looks like noise. It is necessary to modify time trace to obtain the desired variance and mean value. Every sample then represents a flow in number of packets per time unit.

To obtain a sequence of packets of variable lengths, it is necessary to know pakcet distribution in the observed network segment. More details on the model can be found in [6], [8], [9].

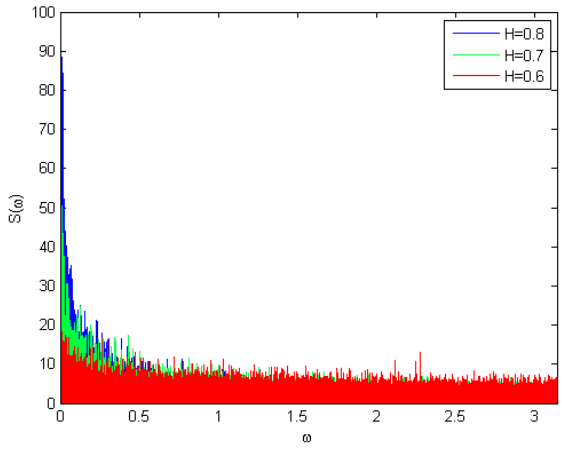


Fig. 4. The power spectrum of FGN process.

This method requires that the size of sequence which needs to be generated to be defined.

D. Markov Chain Based Models

An infinite Markov chain (IMC) can be used to generate a time series exhibiting LRD (see Fig. 5). If the chain is in the state i, $i\neq0$, it moves to state i–1 in next time instant generating one in the output sequence. Only in the state i=0 (zero state), zero is generated in the output sequence. Any state, i, can be reached from the zero state with the transition probability f_i. In brief, if the chain changes its state from 0 into n ($n\neq0$), then the sequence of n ones will be generates. The sequence of zeros will be generated if the chain remains in the zero state.

In practical implementation, IMC does not have an infinitive number of states. It is specified to an maximum state $N_{\max}$. For finite Markov chain (FMC) model, the number of states, N, is fixed. It is also the maximum packet burst size. Since burstiness is the condition for the self-similarity, the sequence with large H cannot be obtained with low N.

In infinite Markov chain realization, the number of states is determined dynamically. The initial number of states, N, a relatively small number, is specified. If the probability f_i, generated in the zero state, does not correspond to any state in range from 0 to N, then states from N+1 to 2N are generated and the state corresponding to given probability is looked for. The procedure is repeated until the maximum state, N_{max}, is reached. Therefore, this chain is also constrained by N_{max}, but N_{max} could be very large number.

Details about these models can be found in [7],[10].

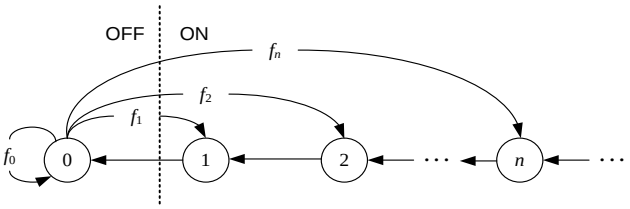


Fig. 5. An infinite Markov chain.

IV. Hurst Parameter Estimation Techniques

The properties of self-similar processes (see Section II) lead to the different methods to estimate H.

The R/S statistic is a method that estimates the parameter H as a slope of the straight line in plot of log(R/S) versus log(T), where R is a measure of the range of the process, S is the standard deviation of the sample and T is the time interval.

The Variance-Time plot (VTP) method estimates the Hurst parameter on the basis of the variance of the aggregated time series. Considering eq. (2), a plot of log[Var($x^{(m)}$)] versus log(m) will yield a straight line with slope of –β. The Hurst parameter is related to β by H=1–β/2.

The Periodogram plot is a frequency domain technique, unlike the previous two methods which are the time domain estimators. By plotting log(S_n) versus log(ω), where S_n denotes the power spectrum density of n-length time-sequence block, and ω is the frequency, the points of the periodogram scattered around a negative slope are obtained. An estimate of the Hurst parameter is given by H=(1–γ)/2, where γ is the slope.

More about these techniques can be found in [10].

V. Simulation Results

For traffic generator models presented in the Section III, H and traffic load values of generated sequences are estimated by simulation analysis.

For ICM model, parameters m_1 and m_2 are equal. The shortest packet burst length for Pareto generator is β=1. For FMC model, different chain lengths are specified for various H values in order to attain the appropriate burstiness [10]. Chain sizes are N=2^{13}, 2^{14}, 2^{15}, 2^{16}, 2^{17}, for H=0.55, 0.65, 0.75, 0.85 and 0.95, respectively. For IMC model, the initial chain length is N=256, and maximum one is $N_{\max}$=65636 for H<0.8 and N=2^{10}, $N_{\max}$=2^{20} for H>0.8.

The R/S plot, Variance-Time plot and Periodogram plot are used to estimate H from simulated sequences.

Each model, except the FGN, provides the binary sequence of ten million bits. The parameter H is estimated for the sequence aggregation of degree m=100. The FGN model generates a sequence of 100 000 points. The mean value H_{avg} and standard deviation σ_H are calculated for the Hurst parameter value estimated for 50 sequences. Traffic load is equal to ρ=0.5. The results are shown in Table I.

Based on the results, it is noticed that R/S method provides small standard deviation with lower H values when compared to other two methods. For higher H values, R/S method provides much higher standard deviation.

For H=0.55 and H=0.65, the FGN model generates series with H close enough to specified one. For higher values of H, the FMC model provides better results than FGN in terms of estimated H. Though, it should be bared in mind that the FGN has the lowest standard deviation for all specified H values. For ICM and Pareto models, the difference between specified and estimated H is large. Also, the variance of H is larger compared to other models.

TABLE I. ESTIMATION OF THE HURST PARAMETER.

Generator model	H	R/S $H_{avg} \pm \sigma_H$	Variance-Time $H_{avg} \pm \sigma_H$	Periodogram $H_{avg} \pm \sigma_H$
Iterated chaotic map	0.55	0.570 ± 0.025	0.567 ± 0.040	0.579 ± 0.038
	0.65	0.611 ± 0.037	0.634 ± 0.072	0.642 ± 0.043
	0.75	0.684 ± 0.114	0.729 ± 0.077	0.727 ± 0.033
	0.85	0.751 ± 0.117	0.805 ± 0.046	0.810 ± 0.024
	0.95	0.893 ± 0.192	0.890 ± 0.037	0.900 ± 0.017
Fractional Gaussian noise	0.55	0.561 ± 0.016	0.548 ± 0.010	0.538 ± 0.011
	0.65	0.655 ± 0.019	0.646 ± 0.011	0.629 ± 0.012
	0.75	0.743 ± 0.018	0.737 ± 0.010	0.723 ± 0.012
	0.85	0.824 ± 0.017	0.821 ± 0.013	0.811 ± 0.011
	0.95	0.895 ± 0.020	0.889 ± 0.011	0.899 ± 0.011
Pareto distribution	0.55	0.575 ± 0.019	0.561 ± 0.064	0.569 ± 0.049
	0.65	0.613 ± 0.044	0.626 ± 0.069	0.652 ± 0.053
	0.75	0.677 ± 0.060	0.712 ± 0.082	0.731 ± 0.036
	0.85	0.768 ± 0.089	0.809 ± 0.059	0.814 ± 0.029
	0.95	0.903 ± 0.252	0.894 ± 0.033	0.906 ± 0.017
Finite Markov chain	0.55	0.586 ± 0.023	0.577 ± 0.029	0.598 ± 0.037
	0.65	0.634 ± 0.043	0.644 ± 0.045	0.678 ± 0.051
	0.75	0.724 ± 0.058	0.753 ± 0.037	0.775 ± 0.036
	0.85	0.847 ± 0.081	0.836 ± 0.030	0.824 ± 0.025
	0.95	0.946 ± 0.068	0.922 ± 0.008	0.925 ± 0.018
Infinite Markov chain	0.55	0.569 ± 0.017	0.547 ± 0.032	0.559 ± 0.039
	0.65	0.616 ± 0.037	0.626 ± 0.064	0.652 ± 0.056
	0.75	0.664 ± 0.055	0.687 ± 0.060	0.712 ± 0.037
	0.85	0.750 ± 0.085	0.785 ± 0.048	0.797 ± 0.027
	0.95	0.808 ± 0.099	0.847 ± 0.030	0.872 ± 0.021

Results for IMC model are much lower than expected for H>0.7. None of the observed models can reach very high Hurst parameter values as H=0.95.

Performances of network devices are usually analysed under high traffic loads conditions. Also, higher values of H parameter are of interest. Fig. 6. plots traffic load ρ for three H parameter values. Resultant load is calculated as average value of loads estimated for 50 sequences.

FGN and FMC models provide traffic load which is close to the desired one. For other models, deviations from the specified value are much higher, especially for large traffic load and H parameter value.The IMC model is invalid for some combinations of ρ and H [10]. Values for these combinations are labeled in Fig. 6. with "x".

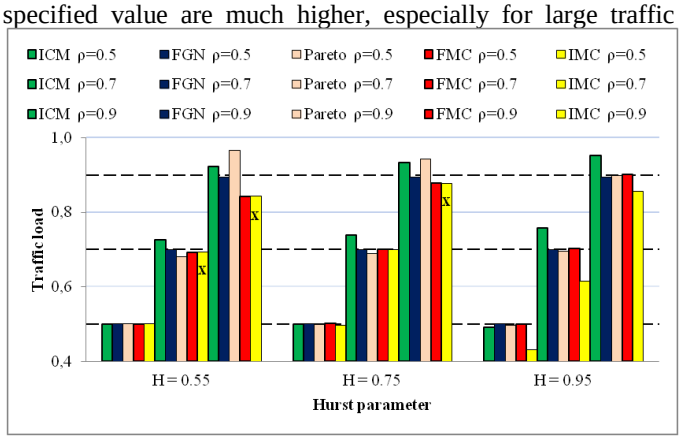


Fig. 6. Traffic load for different values of Hurst parameter.

All models are implemented in Matlab. To generate the binary sequence 10 million bits long, it took in average 0.3 seconds for FMC, 0.87 seconds for FGN, 3.01 seconds for ICM, 3.92 seconds for Pareto model and 4.45 seconds for IMC. The execution time of algorithms depends on values of parameters H and ρ, and other parameters, such as size of a Markov chain. Given values are determined on 750 runs with various combinations of H i ρ as in Table I.

VI. CONCLUSION

Considering all results, among analysed self-similar traffic generator models, according to performances referring to Hurst parameter and traffic load of generated series, and also to algorithm efficiency, FGN model and the one with finite Markov chain are emphasized. The advantage of FGN model is lower variance of the Hurst parameter and traffic load which is closer to expected. The difference between desired and obtained H is smaller when FGN is used with smaller H values, although finite Markov chain has proven to be better for higher values of H parameter. The advantage of Markov model is its efficiency and its capability to attain higher Hurst parameter values compared to FGN model. In applications where traffic generation in online manner is required, it is sometimes simpler to use Markov model because for FGN model the total length of sequence must be specified in advance.

REFERENCES

[1] V. Paxson, S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling", *IEEE/ACM Trans. Networking*, vol. 3, no 3, pp. 226-244, 1995.

[2] K. Park, G. Kim and M. Crovella, "On the effect of Traffic selfsimilarity on network performance", in *Proc. Of SPIE International Conf. on Performance and Control of Network System*, Nov. 1997., pp. 293-310.

[3] K Park, G Kim, M Crovella,"On the relationship between file sizes, transport protocols, and self-similar network traffic", *International Conference on Network Protocols*, Oct. 1996.,pp. 171-180.

[4] Garrett, M.W., Willinger, W., "Analysis, modeling and generation of self-similar VBR video traffic", In: Proc. ACM SIGCOMM 1994, London, UK, pp. 269–280

[5] A. Adas and A. Mukherjee. "On resource management and QoS guarantees for long range dependent traffic", In Proc. IEEE INFOCOM 1995, pages 779–787,

[6] S. Taebi Harandi, "Self-Similar Traffic Modeling and Clos-like Packet Switch Architectures," M.Sc. thesis, Electrical Engineering (Optical Networking) University of Ottawa, Ottawa, ON, Canada, (2005)

[7] R. G. Clegg and M. Dodson, "A Markov Chain Based Method for Generating Long-Range Dependence," Phys. Rev. E 72, 026118 DoI: 10.1103/PhysRevE.72.026118, (2005)

[8] S. Ledesma and D. Liu, "Synthesis of Fractional Gaussian Noise Using Linear Approximation for Generating Self-Similar Network Traffic," ACM Computer Communication Review, Volume 30, Number 2, pp 4-17, (2000)

[9] W. Paxon, "Fast, Approximate Synthesis of Fractional Gaussian Noise for Generating Self-Similar Network Traffic," Comp Comm. Rev. 27,5-18, (1997)

[10] R. G. Clegg, "The Statistics of Dynamic Networks," Ph.D. thesis, Dept. of Math., Uni. of York, York, (2004)

Overview of the HbbTV compliant browser upgrade on Android based DTV platform

Milena Milosevic, Krsto Lazic, Branimir Kovacevic, Nenad Jovanovic, Marko Kovacevic
Department of Computer Engineering and Communications
Faculty of Technical Sciences
Novi Sad, Serbia

***Abstract*—Upgrade to the new version of software is resource and time consuming process. Manufacturers try to optimize this process in any way. This paper presents upgrade of the HbbTV software solution as a system which relies on web technologies. Since web browser is one of the platform dependent components of the HbbTV software solution, the paper focuses on the web browser's update on Android based DTV platform. With the Android upgrade, web browser also changes. These changes can be major changes and to achieve faster releases and better performances, manufacturers have to choose which browser to choose, existing or the upgraded one. Two default Android web browsers are analyzed and compared. Web browsers' overview of supported features and performances related to the HbbTV is given.**

Keywords-digital television; HbbTV; Android; web browser

I. INTRODUCTION

Demand for digital entertainment is one of the fastest growing in the CE (Consumer Electronics) market. Compared to mobile devices, tablets and PCs, where user experience is interactive, traditional television offers static user experience. Nowadays, same user interface is expected on all connected devices, including DTV (digital television) devices (TVs and STBs (set-top boxes)). More interactions and social features are requested in addition to watching TV, and thus DTV devices are moving forward from static to interactive user experience. In the fast growing DTV market there are variety of devices, technologies and standards offering interactivity and entertainment.

Hybrid Broadcast-Broadband TV (HbbTV) standard [1] provides integration of broadcast DVB (Digital Video Broadcasting) services and broadband IP network. The integration provides better user experience through services like digital teletext, EPG (Electronic Program Guide), and program non-related services, such as video on demand, catch-up TV, interactive ads, voting, personalization, games or social networking.

The traditional TV software had small significant updates over time. With the appearance of the Smart and Connected TVs, TV market rapidly changed [2], with constant need for upgrades and new outcomes. For example, Android based DTV platform needs an upgrade each time new Android major version is available.

Software solution of DTV architecture is complex. Its integration with new Android version and upgrades are time and resource consuming processes for CE manufacturers.

The mentioned changes in DTV software development influenced emerges of traditional TV software solutions with rapid development model. Rapid software development model is also a choice in web browser development. Google's Chrome applied this model from the beginning, while Mozilla Firefox transitioned to this model [3] [4].

This paper presents and analyzes upgrade to the new version of the Android platform through upgrade of the HbbTV software solution as a complex software system which relies on web technologies. HbbTV applications which have become widespread are executed in the web browser. Web browsers are customized to support HbbTV standard requirements. There is more than one web browser on which HbbTV compliant web browser can be based. There are open source, proprietary and hybrid browsers (half open source, half proprietary). Some browsers have support for different operating systems.

Android platform has its default web browser. This work focuses on HbbTV solution based on the default web browser. With every new Android release, update of default web browser is present. Updates can be minor and customization of that browser for HbbTV purposes goes smooth, but sometimes updates lead to some major changes. One of such updates with major changes is transition from WebKit-based web browser to Chromium-based web browser which came about in Android release 4.4.x. These changes impact on the porting of the HbbTV solution and thus impact on the whole DTV solution upgrade in terms of time and resources.

Question is how to achieve rapid release development from one to another version, now and in the future? We propose choosing one web browser and making it HbbTV dedicated. Two Android default browsers will be discussed and analyzed as HbbTV candidate browsers: WebKit-based browser on which existing HbbTV solution relies, and Chromium-based browser which is going to have greater support in the future. Advantages and disadvantages of both solutions are presented.

The rest of this paper is organized as follows. Section 2 gives related work. HbbTV overview and features that should be supported by web browser are given in third section, while section 4 gives comparison of these two browsers regarding HbbTV requirements. Last section concludes the paper.

This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant 32034.

II. Related Work

Modern web browsers are compared by many criteria. Anand and Saxena [5] compare Chrome, Firefox and Internet Explorer in terms of CPU utilization, disk usage, multiple tab performances, compatibility with modern web technologies (HTML5 and CSS3) and execution time of JavaScript engines. Nielson, Williamson and Arlitt [6] give overview of the benchmark performance test results for four browsers (Internet Explorer, Safari, Firefox and Opera). Their focus is mostly on JavaScript, rendering and AJAX performances. They also emphasize the choice of the operating system which may affect the test results. All authors agree that there is no 'one browser for all'; which browser is better depends on its use and purpose.

Web browser is HbbTV compliant if it supports some predefined features. Features which need to be implemented on the Java based devices in order to support HbbTV functionality are described in [7]. Some of these features are web browser related: support for spatial navigation, additional DOM events, support for JavaScript plugins required by the HbbTV 1.1.1 standard [1] etc. More specific overview of these and additional changes necessary for the implementation of the HbbTV required functionality on the Android based DTV devices is given in [8]. HbbTV 1.2.1 standard [9] requires support for one more JavaScript object (application/oipfSearchManager). One implementation of this embedded object is described in [10], with accent on the search performances. HbbTV 2.0 specification is not yet published and should offer a solution for synchronized IP and DVB reception, companion screen feature and transition to some of the HTML5 features as new widely adopted technology. Technology development of HBB-Next project and its contribution to HbbTV 2.0 (support for second screen, media synchronization and user identification) are presented in [11].

III. HbbTV Browser Requirements

HbbTV standard is based on the existing standards and web technologies (Fig.1). It further profiles them and introduces few novel solutions.

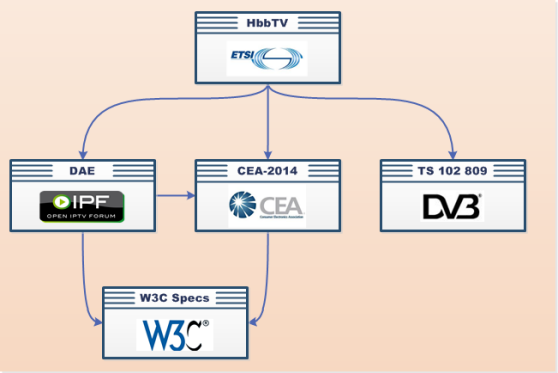


Figure 1. HbbTV and existing standards overview

CEA-2014-A (also known as Web4CE) [12] refers to the problem of displaying web pages on the DTV devices. It defines CE-HTML (Consumer Electronics HTML) language which is based on the following web languages:

- XHTML 1.0 transitional/strict,
- Ecmascript 262, 3rd edition,
- DOM Level 2.0,
- CSS TV Profile 1.0,
- XMLHttpRequest.

CE-HTML also specifies some additional features. HbbTV standard addresses on the following additional features: support for text input using remote control (Multi-Tap), definition for key codes which are sent to the applications (solves diversity codes problem on the DTV devices) and MIME type for CE-HTML applications (application/ce-html+xml).

CEA-2014-A, along with the CE-HTML, also defines embedding of non-linear audio/video content in an application (using JavaScript plugin). It specifies DOM event handling (for example key events) and still image formats (JPEG, GIF, and PNG).

DAE (Declarative Application Environment) [13] defines embedding of linear audio/video content in an application and specifies JavaScript API for applications running in a DTV environment. It specifies navigation using CSS3 directional focus navigation while the nav-up, nav-down, nav-left, nav-right properties are used by the applications.

ETSI-TS 102 809 [14] defines DTV transport of information about applications and transport of applications via DSM-CC standard. HbbTV browser has to support DVB URLs.

Along with the existing technologies HbbTV standard defines the application lifecycle and gives some recommendations. HbbTV browser should have cookie support and should not offer a history UI (user interface) for HbbTV applications.

HbbTV 2.0 specification announces the use of HTML5 features. HTML5 standard is not yet defined, but is being widely adopted technology by all major browsers. Main features which HTML5 brings are:

- abstracting the content from the hardware (audio, video and other components),
- WebSocket – an asynchronous protocol providing full-duplex communication over a single TCP connection. Server can send messages to the client at any moment,
- WebRTC – enables real-time communication, for example: voice call, video chat and P2P file sharing without plugins,
- localized storage.

Data exchange between application and DTV middleware in the existing solution is realized via JavaScript plugins (HbbTV embedded objects). Plugins are implemented using NPAPI (Netscape Plugin Application Programming Interface) [15] as a cross-platform plugin architecture. NPAPI supports initializing, creating, destroying and positioning plugin content,

scripting, printing, full-screen plugins, windowless plugins and content streaming. For HbbTV embedded objects significant value in browsers has scripting support, asynchronous callback support and object array support.

IV. Comparison of Web Browsers

Android version 4.4.x introduced Chromium-based default web browser. Android web browser in older versions has been based on the WebKit. Accordingly, HbbTV software solution in versions older than this has been based on the WebKit engine. With the new Android upgrade and emerge of Chromium-based browser the question is which web browser provides good performances and cost-effective updates. Two solutions arise: preserving WebKit-based browser and its porting of native parts on the upgraded Android version or switching HbbTV solution to the Chromium-based browser (making it HbbTV compliant).

WebKit-based browser is HbbTV compliant, as part of the existing HbbTV solution. Support for CE-HTML, DOM events handling, OIPF MIME types, CSS 3.0 navigation properties and DOM extension with HbbTV required objects (KeyEvent, KeySet and Channel) has been implemented. HbbTV JavaScript object for communication with DTV middleware are implemented using NPAPI. WebKit-based browser supports NPAPI architecture by default. It is dependent on native Android APIs. With the upgrade of Android these APIs has changed. Transition of HbbTV solution relying on WebKit based browser to Android 4.4.x requires significant effort on porting this browser's native part.

Chromium-based browser, compared to WebKit-based browser, is already integrated with the native APIs on the upgraded Android version. This default browser is not HbbTV compliant. HbbTV browser related features such as DOM object extension with new objects, CE-HTML parsing, support for OIPF MIME types, CSS 3.0 navigation properties and specific handling of DOM events need to be implemented. Cookie support exists but there is no support for DVB URLs.

In the default Chromium-based browser plugins are disabled. HbbTV requires plugins' support, so Chromium-based browser needs to be reconfigured to support plugins. Plugin support is exposed through PPAPI (Pepper Plugin API) [16]. NPAPI in this browser is deprecated and will be excluded in the future. HbbTV JavaScript plugins are implemented using NPAPI. Since PPAPI doesn't provide API for scripting objects, plugins can't be rewritten using PPAPI. Instead, NPAPI support has to be revived in the Chromium-based browser.

Main question is which solution leads to faster HbbTV release. WebKit-based browser is already HbbTV compliant and needs be ported on modified native Android API. Chromium-based browser runs on Android 4.4.x and newer versions as default browser, but needs to be modified to be HbbTV compliant. Along with development effort, web browser performances and support for HbbTV relevant and next generation features are important. Authors of the [5] conclude that Chromium's weak point is multi-tabbing. This feature is irrelevant for HbbTV browser; standard specifies only one active application at the moment.

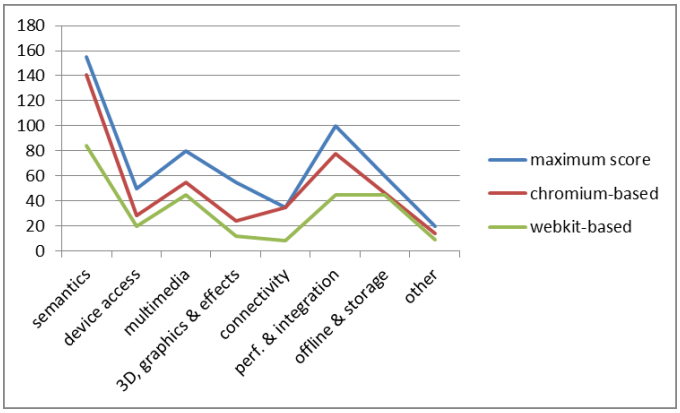


Figure 2. HTML5 support test results

HbbTV 2.0 introduces some of the next-generation features. HTML5 support is one of them. Tests for HTML5 browser support already exist. They test currently specified features by the HTML5 specification and tests change accordingly to the specification. One such test [17] is executed on both browsers. This test validates support for HTML5 semantics, device access, performance and integration, multimedia, 3D graphics and effects, connectivity, offline and storage and other. Some non-HTML5 features are also validated but are not included into final result. Results for these two browsers are depicted in the Fig. 2. They show how many scores browser gets out of maximum.

Results depicted in Fig.2 indicate that Chromium-based browser has better overall HTML5 support than the WebKit-based.

HbbTV applications communicate with the DTV middleware and achieve its functionalities via JavaScript plugins. Browser's JavaScript engine performances influence the performances of the HbbTV applications and thus on the user experience. The benchmarks measure performances of the JavaScript engine.

SunSpider [18] as JavaScript benchmark validates the core JavaScript language. It focuses on the actual JavaScript developers' use today trying to avoid microbenchmarks. Tests are executed multiple times to determine measurement error range. Results (Fig. 3) are given in milliseconds.

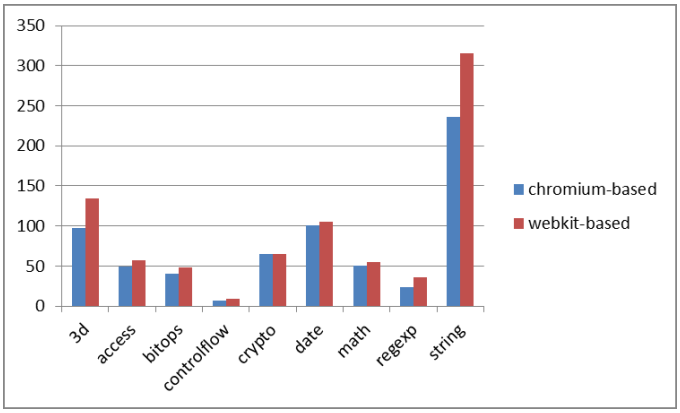


Figure 3. SunSpider benchmark test results

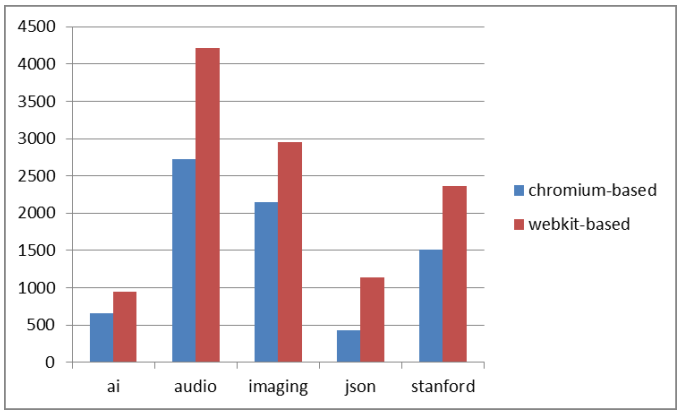


Figure 4. Kraken benchmarks test results

Kraken [19] is JavaScript performance benchmark which measures the speed of several use cases from real-world applications and libraries. It is based on SunSpider. It includes an implementation of the search algorithm, audio processing usage, image filtering routines, JSON parsing and cryptographic routines. Results (Fig. 4) are also reported in milliseconds.

V8 Benchmark Suite [20] reflects pure JavaScript performance. Results are shown in Fig. 5 and higher scores mean better performances.

All tests are executed on the reference platform and are here to give mutual performance overview of these two browsers. From the test results Chromium-based browser is considered as browser with better JavaScript performances, but in the real world it is hard to see JavaScript performance difference between these browsers. Overall user experience considering speed is satisfying for average user.

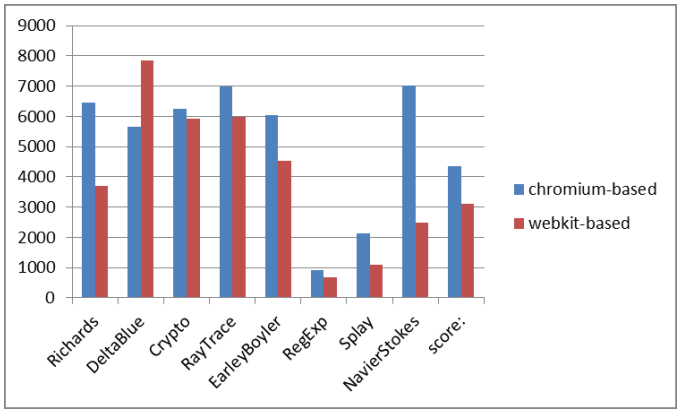


Figure 5. V8 Benchmark Suite results

V. CONCLUSIONS

This paper analyzes problems of the HbbTV browser upgrade which comes with the Android platform upgrade. Overview of the HbbTV compliant browsers based on two major Android web browsers is given. HbbTV required features and browser support for them is presented. The work focuses also on the support for next-generation features such as HTML5 and gives comparison of the browsers' JavaScript performances. To achieve faster HbbTV software releases next step would be choosing one web browser, making it as much as possible platform independent and its use in HbbTV solution across all Android versions.

REFERENCES

[1] ETSI TS 102 796: Hybrid Broadcast Broadband TV; V1.1.1 .

[2] IDATE: "TV 2010, Markets & trends, facts & figures".

[3] O. Baysal, I. Davis, and M. W. Godfrey, "A tale of two browsers," Proceeding of the 8th working conference on Mining software repositories , ser. MSR '11, 2011, pp. 238–241.

[4] F. Khomh, T. Dhaliwal, Y. Zou, and B. Adams, "Do faster releases improve software quality? An empirical case study of Mozilla Firefox," Proceedings of the international working conference on Mining software repositories, Jun 2012.

[5] V. Anand, D. Saxena, "Comparative study of modern Web browsers based on their performance and evolution ," Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on, Dec. 2013.

[6] J. Nielson, C. Williamson and M. Arlitt , "Benchmarking Modern Web Browsers ", Proceedings of IEEE HotWeb 2008, October 2008.

[7] N. Kuzmanovic, V. Mihic, T. Maruna, M. Vidakovic and N. Teslic, "Hybrid broadcast broadband TV implementation in java based applications on digital TV devices", IEEE Trans. Consumer Electronics, 2012, pp.1056-1062.

[8] M. Milosevic, S. Markovic, B. Mlikota, M.Zivanovic and B.Prtvar, "Porting of HbbTV functionality on DTV platform based on Android OS" , TELFOR 2012.

[9] ETSI TS 102 796 : Hybrid Broadcast Broadband TV; V1.2.1.

[10] M.Milosevic, B.Mlikota, M.Radonjic and T. Maruna, "HbbTV 1.5 additions: one design, implementation and integration " , ICCE Berlin 2014, in press.

[11] M.O van Deventer, J.J de Wit, M. Guelbahar, B. Cheng, F. Gomez Marmol, C. Köbel, C. Köhnen, G. Rozinaj, and B. Stockleben: "Towards next generation Hybrid broadcast broadband, results from FP7 and HbbTV 2.0", IBC2013 Conference, Amsterdam, Netherlands , 2013.

[12] CEA-2014 revision A: Web-based Protocol and Framework for Remote User Interface on UPnP™ Networks and the Internet (Web4CE) .

[13] Open IPTV Forum Release 1 specification, volume 5: Declarative Application Environment v1.1.

[14] ETSI TS 102 809: Digital Video Broadcasting (DVB); Signaling and carriage of interactive applications and services in hybrid broadcast - broadband environments; V 1.1.1.

[15] NPAPI documentation : Gecko Plugin API Reference available at https://developer.mozilla.org/en/Gecko_Plugin_API_Reference, accessed 5th Sept..2014.

[16] Pepper Plugin API documentation: available at https://code.google.com/p/ppapi/, accesse 5.9.2014.

[17] HTML5 Test : available at https://html5test.com/, accessed 5th Sept. 2014.

[18] SunSpider 1.0.2 JavaScript Benchmark: available at https://www.webkit.org/perf/sunspider-1.0.2/sunspider-1.0.2/driver.html , accessed 5th Sept. 2014.

[19] Kraken JavaScript Benchmark (version 1.1) : available at http://krakenbenchmark.mozilla.org/kraken-1.1/driver.html, accessed 5th Sept. 2014.

[20] V8 Benchmark Suite - version 7: available at http://v8.googlecode.com/svn/data/benchmarks/v7/run.html, accessed 5th Sept. 2014.

POSTER SESSION

Successive absorption and refraction in ultrathin molecular nano-films

Jovan P. Šetrajčić
Department of Physics
Faculty of Science, University of Novi Sad
Novi Sad, Vojvodina, Serbia
bora@df.uns.ac.rs

Siniša M. Vučenović
Department of Physics
Faculty of Science, University of Banja Luka
Banja Luka, Republic of Srpska, B&H
vucenovic.s@gmail.com

Blanka Škipina
Faculty of Technology, University of Banja Luka
Banja Luka, Republic of Srpska, B&H
blanka.skipina@gmail.com

Svetlana Pelemiš
Faculty of Technology, University of Srpsko Sarajevo
Zvornik, Republic of Srpska, B&H
alannica@gmail.com

***Abstract*—Based on the formed model of nano-film crystal structures, changes of optical properties due to the presence of borders for the case of symmetrical ultrathin films are theoretically investigated in this paper. Influence of five border parameters on the occurrence of localized exciton states is examined, as well as their relation with the effects of discretization and selection of resonant absorption of present electromagnetic radiation. Used combined analytical-numerical calculation to find the allowed energy states of excitons and their spatial distribution (per layers) along the axis perpendicular to surface planes. We determined permittivity for the observed models of these ultrathin dielectric films and explored the influence of boundary parameters on the occurrence of discrete and selective absorption.**

Keywords - nano-films, excitons, permittivity, absorption.

I. INTRODUCTION

*Obtaining fundamental information on different physical and chemical properties of materials and their wide practical (technical and technological) applications in nano-, bio- and optoelectronics intensified the theoretical research of low dimensional crystal system (nanostructures: ultrathin films, quantum wires and quantum dots and so on. [1]). Compared with characteristics of appropriate "big" samples, specificity of these "small" structures is reflected in the fact that the presence of close boundary surfaces leads to altered very general known properties of these materials and the occurrence of non-specific phenomenon (as a result of the effects of dimensional quantization) [2,3]. The excitons are responsible for the dielectric, optical (absorption, dispersion of light, luminescence), photoelectric and other properties of crystals [4]. In this study we observed ultrathin dielectric films (with a thickness not exceeding of ten atomic planes). Representative examples of these structures are molecular crystals, with in them occurring elementary excitations – excitons, as a result of interaction of the external electromagnetic field and electrons in crystals. Using dispersion law for the excitons and their density of states, we will theoretically define the relative permittivity, and throughout the optical properties of a system.

* This work was partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No: OI-171039 and TR-34019) and by the Provincial Secretariat for Science and Technological Development of Vojvodina (Grant No: 114–451–2048) as well as by the Ministry of Science and Technology of the Republic of Srpska (Grant No: 19/6-020/961-23/12).

II. EXCITONS IN NANO-FILMS

Standard effective excitonic Hamiltonian in harmonic approximation [4,5] has following form:

$$H = \sum_{\vec{n}} \Delta_{\vec{n}} B_{\vec{n}}^{+} B_{\vec{n}} + \sum_{\vec{n},\vec{m}} X_{\vec{n}\vec{m}} B_{\vec{n}}^{+} B_{\vec{m}} \,, \tag{1}$$

where $B_{\vec{n}}^{+}$ and $B_{\vec{n}}$ are operators of creation and annihilation of exciton on site $\vec{n}$ of crystal lattice; $\Delta_{\vec{n}}$ represents energy of exciton isolated on this site, and $X_{\vec{n}\vec{m}}$ are matrix elements of excitons transfer from site to site. In this model we assume that energy of exciton isolated on site is about 100 times bigger compared with the energy needed for transfer of excitons.

We will analyze this system with the method of two-times temperature-depended Green's functions [6-8] mainly for advantages which this method provides. Green's function has given in following form:

$$G_{\vec{n}\vec{m}}(t) = \left\langle\left\langle B_{\vec{n}}(t) \middle| B_{\vec{m}}^{+}(0) \right\rangle\right\rangle, \tag{2}$$

and satisfied following equation of motion:

$$i\hbar \frac{d}{dt} G_{\vec{n}\vec{m}}(t) = i\hbar\delta(t)\delta_{\vec{n}\vec{m}} + \Delta_{\vec{n}} G_{\vec{n}\vec{m}}(t) + \sum_{\vec{l}} X_{\vec{n}\vec{l}} G_{\vec{l}\vec{m}}(t) \,. \tag{3}$$

Systems which have structures containing boundary planes are usually called film-structures [9-11]. Dimensions in crystal nano-films are unlimited in XY directions, while in Z direction has finite width: $L=Na$. We will observe dielectric nano-film (Fig.1) which could be made by the controlled doping within bulk structures, or various methods of deposition on substrate materials [11]. Existing boundary layers change energies of excitons on site and transfer of exciton energies between boundary planes ($n_z=0$; $n_z=N$) and first inner planes ($n_z=1$; $n_z=N-1$) within ultrathin film as well, which define perturbation conditions [11-15]:

$$\begin{aligned} \Delta_{\vec{n}} &\equiv \Delta\ \left[1+d\left(\delta_{n_z,0}+\delta_{n_z,N}\right)\right]; \\ X_{\vec{n},\vec{n}+\vec{\lambda}} &\equiv X\ \left[1+x\left(\delta_{n_z,0}+\delta_{n_z,N-1}\right)\right]; \\ X_{\vec{n},\vec{n}-\vec{\lambda}} &\equiv X\ \left[1+x\left(\delta_{n_z,1}+\delta_{n_z,N}\right)\right], \end{aligned} \tag{3}$$

where parameter d defines perturbation on site of crystal nodes, and parameter x defines perturbation of energy transfer along Z – axe.

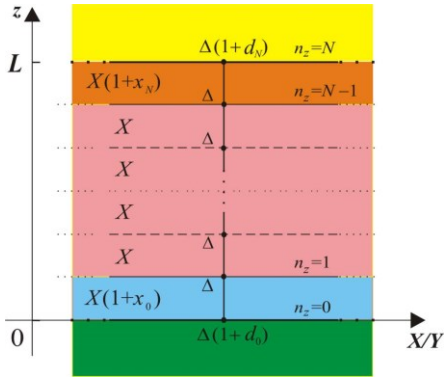


Figure 1. Model of ultrathin film.

Taking into account boundary conditions (3), Hamiltonian (1) and equation of motion (2) we could write equations for the Green's functions [11-15]. After performing of Furrier transformations, unlimited in time but limited in space (limitation is in Z – axe), we get system:

$$\begin{aligned} &G_{n_z,m_z}\left[\rho-\frac{\Delta}{|X|}d\left(\delta_{n_z,0}+\delta_{n_z,N}\right)\right]+ \\ &+G_{n_z+1,m_z}\left[1+x\left(\delta_{n_z,0}+\delta_{n_z,N-1}\right)\right]+ \\ &+G_{n_z+1,m_z}\left[1+x\left(\delta_{n_z,1}+\delta_{n_z,N}\right)\right]=\frac{i\hbar}{2\pi|X|}\delta_{n_z,m_z}\ , \end{aligned} \tag{4}$$

where:

$$\rho=\frac{\hbar\omega-\Delta}{|X|}+2\left(\cos ak_x+\cos ak_y\right).$$

System of equations (4) represent system of N+1 non-homogenuos algebraic difference equations for unknown Green's functions. We actually don't need to find out solution of the system (4), it is sufficient to calculate determinant of system (4) and equalize determinant with zero [11-15], i.e. to calculate poles of the Green's function which define dispersion law of excitons. In such a way we get N+1 solutions $\rho\equiv\rho_\nu$; $\nu=1,2,...,N+1$.

We will present dispersion law for excitons in ultrathin film in non-dimensional form, i.e. calculated reduced energies

$$E_\nu=\frac{\hbar\omega-\Delta}{|X|}\equiv\rho_\nu-R_{xy} \tag{5}$$

in dependence on function $R_{xy}\equiv 2\left(\cos ak_x+\cos ak_y\right)$. On figures 2 – 4 have shown excitons dispersion laws – on figure 2 for the unperturbed (ideal) film; on figures 3 and 4 for d-perturbed and x-perturbed 5-layer film, respectively.

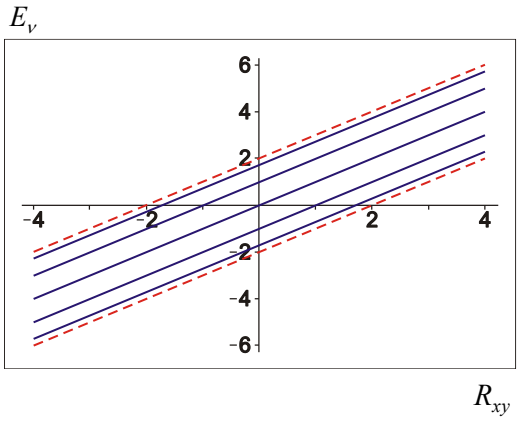


Figure 2. Exciton dispersion law for ideal film.

Comparing nonsymmetrical perturbed (lateral graphics) with symmetrical perturbed (central graphics) films from the figures 3 and 4, one can see clearly nonsymmetrical /symmetrical disposition of the localized states, taking into account that some localized levels overlap in the case of symmetrical perturbation. This overlapping happened only for the d – perturbation, which is the significant one!

In all graph the full lines represent exciton energy levels in films, while dashed lines represent exciton energy levels in bulk crystals. On the first sight one can see absence of zero energies and the presence of discrete energy levels in film structures. Number of the possible energy states is equal to the number of crystallographic planes within the film along Z-axe.

From the figure 3 one can see that increasing of d parameter, energy spectra spread on such a way that one energy level move to the higher energies and come out of the bulk energy border. Those energy states are known as localized or Tamm states [4]. Analyzing graphics on the figure 4 one can see that increasing of x parameter wide the whole spectra with two energy levels which comes out of the bulk energy borders, i.e. with Tamm states.

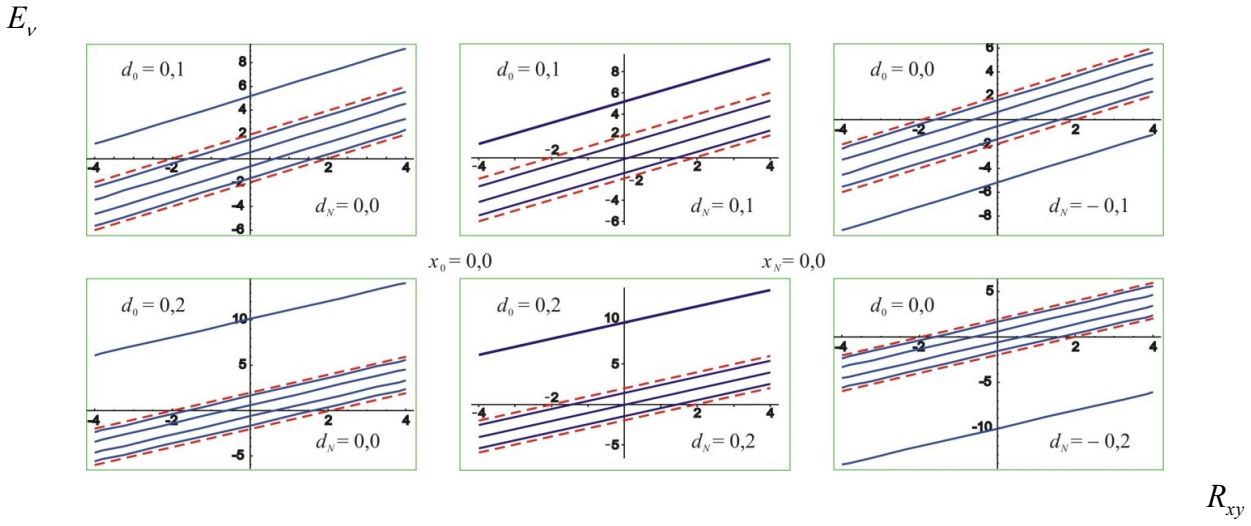


Figure 3. Exciton dispersion law for *d* – perturbed 5-layered film.

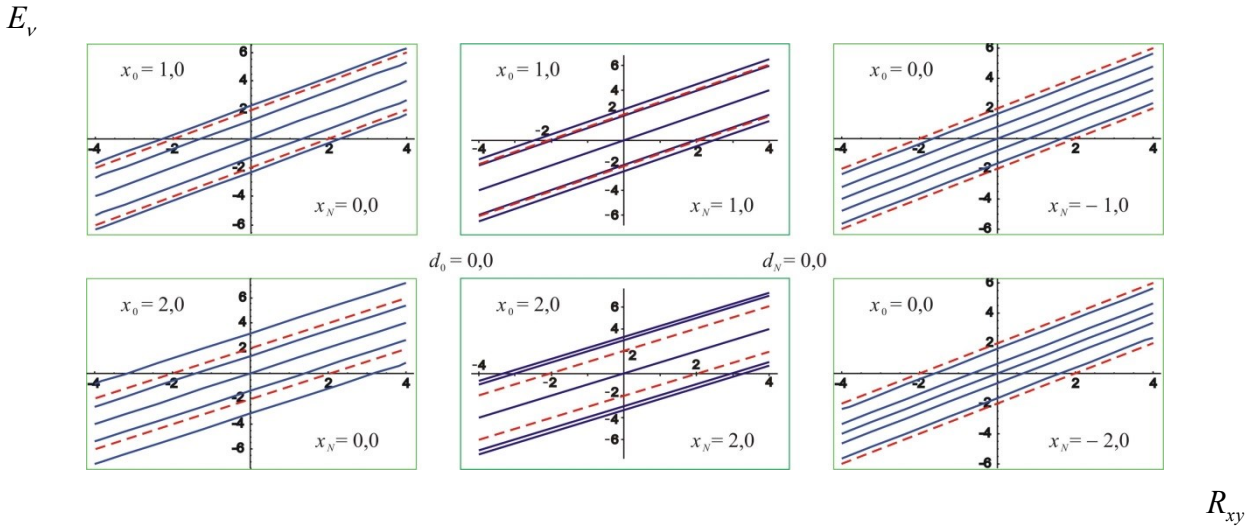


Figure 4. Exciton dispersion law for *x* – perturbed 5-layered film.

III. OCCUPANCY OF EXCITON ENERGY LEVELS AND DIELECTRIC PROPERTIES OF FILMS

To calculate the probability of finding (or creation) of excitons on site and spatial distribution of exciton energy levels within film structures, we need to find out the spectral weight of specific Green's functions. We start from the system of equations for the Green's functions (8), but in matrix form:

$$\hat{D}_{N+1}\widetilde{G}_{N+1} = \hat{K}_{N+1}, \qquad (6)$$

where $\hat{D}_{N+1}$ is matrix which represents determinant of the system, $\widetilde{G}_{N+1}$ and $\hat{K}_{N+1}$ are vectors of Green's functions and Kronecker delta function, respectively. If we imply in (6) inverse matrix $\hat{D}_{N+1}^{-1}$ from the left side, knowing that inverse matrix could be represent throughout adjugate (or adjunct) matrix whose terms D_{ik} are cofactors of elements d_{ik} of direct matrix, we could calculate Green's functions where numerator represents spectral weight $g_{n_z}(\rho_\nu)$, i.e. probability of finding excitons states ρ_ν [9,11-15]:

$$G_{n_z} = -\frac{i\hbar}{2\pi|X|}\sum_{\nu=1}^{N+1}\frac{g_{n_z}(\rho_\nu)}{\rho-\rho_\nu}. \qquad (7)$$

To calculate dynamical permittivity of film we use general expression [4-8], keeping in mind that Green's function (and related dielectric permittivity) are strongly depended of the

number (or site) of the crystalline planes of the symmetrical film:

$$\varepsilon_{n_z}^{-1}(\omega)=1-2\pi i F\ \left[G_{n_z}(\omega)+G_{n_z}(-\omega)\right], \qquad (8)$$

where F is structural factor [4,5]. Including expression for the Green's functions in (8) we obtain:

$$\varepsilon_{n_z}^{-1}(\omega)=1-\frac{\hbar F}{|X|}\sum_{\nu=1}^{N+1}\sum_{s=+,-}\frac{g_{n_z}^{\nu}}{\rho_s-\rho_\nu}, \qquad (9)$$

where $\rho_{\pm}=\dfrac{\mp\hbar\omega-\Delta}{|X|}+2\left(\cos ak_x+\cos ak_y\right)$.

Expression (9) shows how dielectric permittivity of ultrathin films depend on frequency (or energy) of external (or initial) electromagnetic excitation. This is actually dielectric response of the symmetrical molecular film on external electromagnetic field.

IV. Optical properties of the film

Dispersion of dielectric permittivity (or permeability) is function of their dependence on frequency $\varepsilon(\omega)$. With respect to the fact that $f(t)$ is real function, from the expression:

$$\varepsilon(\omega)=1+f(\omega)=1+\int_0^{\infty}f(t)\mathrm{e}^{-\omega t}dt \qquad (10)$$

arise that dielectric permittivity must be complex function, i.e.:

$$\varepsilon(\omega)=\varepsilon'(\omega)+i\varepsilon''(\omega) \qquad (11)$$

where ε' is their real and ε'' imaginary part.

Changing the sign of ω we obtain that $\varepsilon(\omega)$ turns into $\varepsilon^{*}(\omega)$, i.e. $\varepsilon'(-\omega)+i\varepsilon''(-\omega)=\varepsilon'(\omega)-i\varepsilon''(\omega)$, from where we obtain $\varepsilon'(-\omega)=\varepsilon'(\omega)$ and $\varepsilon''(-\omega)=-\varepsilon''(\omega)$.

Dispersion relation gives opportunity to calculate one of the parts of dielectric permittivity (real or imaginary, which is for example known from the experiment) if we know the other part (imaginary or real), for the whole frequency range. Of course, knowing experimentally data for the whole frequency range is practically impossible, but also not even necessary. For example, if we are interested in $\varepsilon'(\omega)$ for the frequency ω, contribution of $\varepsilon'(x)$ for the values x which are very far from the ω is not of crucial importance. Shape of the function $\varepsilon'(\omega)$ in some particular frequency ω is determined by the values $\varepsilon''(x)$ for the values x which are close to the ω. Relation between curves $\varepsilon'(\omega)$ and $\varepsilon''(x)$ for the values x in the vicinity of the ω is determined by the Velicky, who has shown that peak of the refraction index match to the absorption threshold, and peak of the absorption match to the drop of the refraction index.

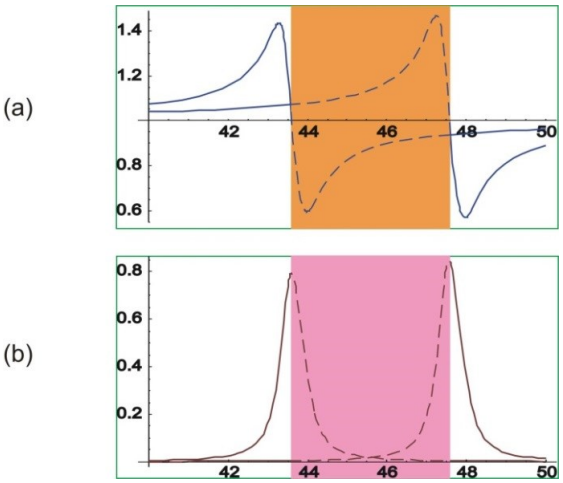


Figure 5. Refraction and absorption indicies of the bulk.

These conclusions are confirmed experimentally. Maxwell has shown that dielectric permittivity of the material is equal to the square of the refraction index. Dielectric permittivity in the case of the dispersion is complex (as we have shown before), with real and imaginary parts closely related with the optical properties of the medium. Introducing complex refraction index:

$$\eta=n+i\kappa, \qquad (12)$$

we will assume that it is related with complex dielectric permittivity with Maxwell equation:

$$\varepsilon(\omega)=\varepsilon'(\omega)+i\varepsilon''(\omega)=\eta^2=n^2-\kappa^2+2in\kappa. \qquad (13)$$

From (13) we can determine physical meaning of real and imaginary part of complex refraction index (12), i.e.:

$$\varepsilon'(\omega)=n^2-\kappa^2;\ \varepsilon''(\omega)=2n\kappa. \qquad (14)$$

Based on this we can find expressions for refraction and absorption indices in next form:

$$\kappa(\omega)=\sqrt{\frac{\varepsilon'}{2}\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}-1\right]}$$

$$n(\omega)=\sqrt{\frac{\varepsilon'}{2}\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}+1\right]}. \qquad (15)$$

On the Fig.6 are shown dependence of relative dynamic permittivity, absorption and refraction indices from the reduced energy of external e.m. field, for the 5-layered film. All graphs show this dependence for the external (boundary) crystalline planes, first inner planes and medium plane.

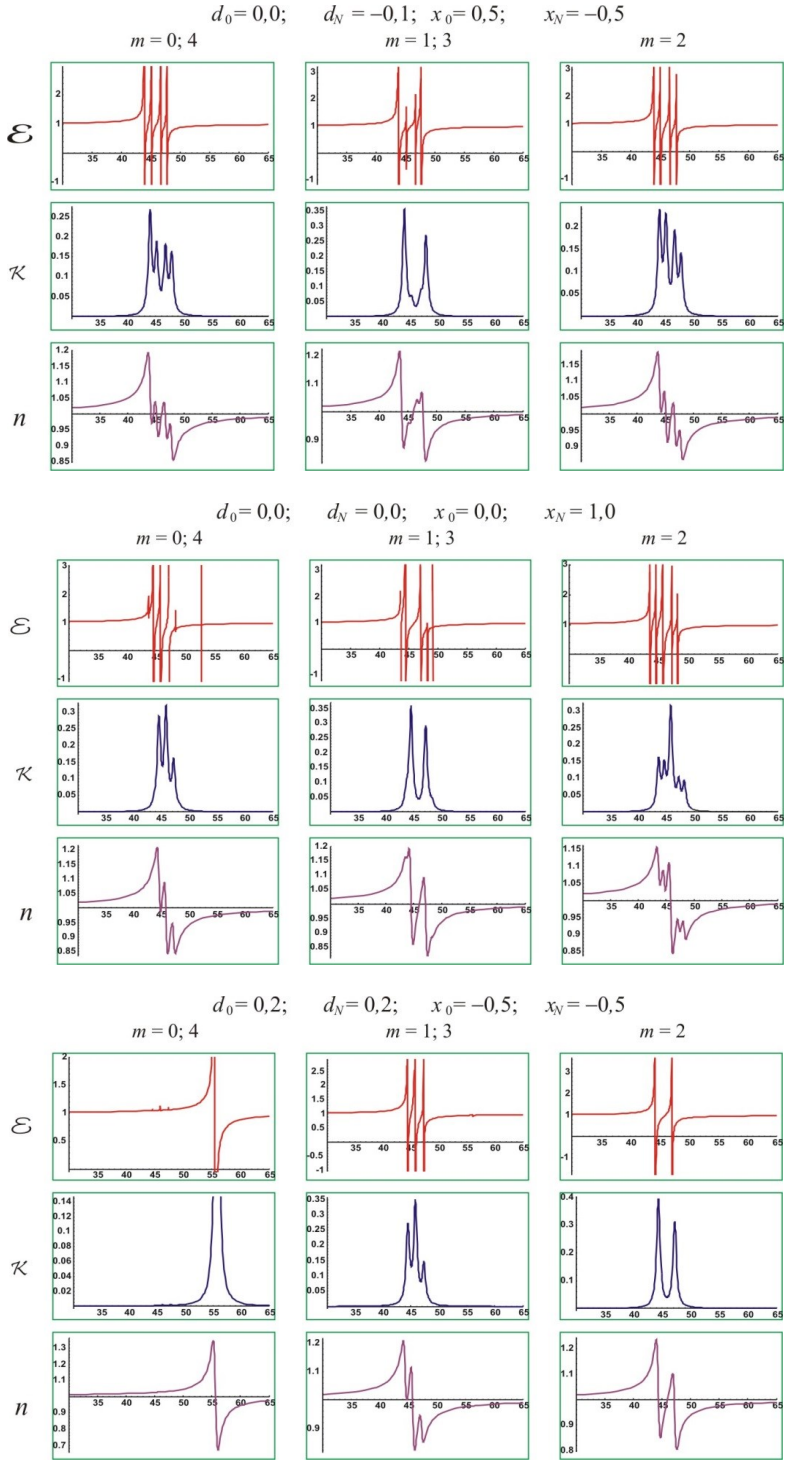


Figure 6. Relative permittivity, absorption and refraction indices for 5-layered perturbed film in dependece of relative energy .

On the first upper graphs in Fig. 6 are shown permittivity of ideal film with 4 inner atomic layers [more detail in 12]. In the second and third row are shown graphs for absorption and refraction indices in dependence of reduced frequency of external electromagnetic field for a 4-layered dielectric film when term Δ has been changed with perturbation parameter d_0 (or d_N). One can see that number of resonant peaks (frequencies where $\varepsilon\ (\omega_r) \rightarrow \pm\ \infty$) depend on number or position of atomic plane n_z for which dielectric permittivity has been calculated, but also from values of perturbation parameters $d_{0/N}$. When we increase perturbation parameter, the absorption zone spreads, with the dominant resonant peak noticed only on boundary plane of the film ($n_z = 0$ ili $n_z \equiv N = 4$). This is somewhat expected result, while spectral weight

analysis shows that the probability of finding excitons is the greatest right on those planes [12]. It is easy to notice that number of resonant permittivity peaks varies with each layer, i.e. permittivity is not only frequency depended, but also layer depended and function of perturbation parameters $x_{0/N}$. Increasing perturbations parameters $x_{0/N}$ absorption zone is widening to higher and lower energies, accompanied with simultaneously quenching of some resonant peaks, depending of layer (or plane) for which is permittivity calculated. This means that visibly greater impact on optical properties has *d*-perturbation parameter, which shifts the whole energy spectra and could make very narrow absorption, making in that way film structure some kind of a monochromatic filter. From all graphs one can see that smoothly increasing of refraction index has its breakdown on frequencies where absorption properties change (absorption index). On the frequencies where absorption index increase – refraction index decrease; where absorption indices change more rapidly (increase or decrease) are peaks of refraction indices (singularities). Those results are in good agreement with Kramer's theory, with only difference that those effects occurs on the beginning and end of absorption zone, while in film structures effects are discrete, giving quantum and very selective narrow absorption zone.

V. CONCLUSION

There are a significant differences in dispersion law (microscopic, i.e. quantum) and dielectric response (macroscopic, but dimensionally quantum) of excitons between bulk and symmetrical ultrathin film structures. This is a consequence of dimensional limitations along one axe, but also of perturbation effects on boundary planes and layers. An energy spectrum of excitons in symmetrical ultrathin film is discrete one with a number of energy levels equal to the number of atomic planes. Increasing of energy of excitons on boundary layers, i.e. crystal nodes, moves spectrum towards higher energies (frequencies), while increasing energy transfer between boundary and first inner neighbor planes symmetrically spreads spectrum towards higher and lower energies (frequencies). There is a possibility of appearance of Tamm's (or localized) states, with probability that those states occurs significantly increase with increasing of perturbation parameters. Dielectric response of ultrathin films show properties of distinct selectivity, with appearance of resonant absorption peaks on exactly determined energies (frequencies). Number and disposition of those resonant peaks depend on number of atomic layers and perturbation parameters. These properties give advantage of film structures above balk (which have continual dielectric response in certain energy zone). In this sense ultrathin film structures could be used as some kind of monochromatic filters of external radiation.

ACKNOWLEDGMENT

The results of this study are the product of a great mind of our friend and collaborator, the late Igor Vragović.

REFERENCES

[1] Wood K., Hammond P., Schmidt D., Wrightman S. and Andaya B., Thin Film Delivers Drugs, Biophotonics, Cambridge, 12.02.2008; (http://www.photonics.com)

[2] Tringides M.C., Jatochawski M. and Bauer E., Physics Today, pp.50-54, April 2007.

[3] Davison S.G. and Steslicka M., Basic Theory of Surface States, Clarendon Press, Oxford, 1996.

[4] Agranovich V.M. and Ginzburg V.L., Crystaloptics with Space Dispersion and Theory of Excitons, Nauka, Moskwa, 1979.

[5] Mirjanić D.Lj., Kozmidis-Luburić U.F., Marinković M.M., Tošić B.S., Can.J.Phys. 60, p.1838, 1982.

[6] Rickayzen G., Green's Functions and Condensed Matter, Academic Press, London, 1980.

[7] Mahan G., Many Particle Physics, Plenum Press, New York, 1990.

[8] Dzialoshinski I.E. and.Pitaevski L.P., Zh.eksper.teor.Fiz**.** 36, p. 1977, 1959.

[9] Pelemiš S., Škipina B., Vučenović S.M., Mirjanić D.Lj. and Šetrajčić J.P., Tehnika – Novi materijali 2, p. 13, 2008.

[10] Cottam M.G., Tilley D.R., Introduction to Surface and Superlattice Excitations, University, Cambridge, 1989.

[11] B.Škipina, D.Lj.Mirjanić, S.M.Vučenović, J.P.Šetrajčić, I.J.Šetrajčić, A.J.Šetrajčić-Tomić, S.S.Pelemiš and B.Markoski, Selective IR Absorption in Molecular Nanofilms, *Optical Materials* **33**, 1578-1584, 2011.

[12] S.M.Vučenović, B.V.Škipina, S.S.Pelemiš, A.J.Šetrajčić-Tomić, S.Armaković, J.P.Šetrajčić, I.J.Šetrajčić, B.Markoski, Optical Properties Analysis of Ultrathin Crystalline Structures with Broken Symmetry, *4th MolMat*, Montpellier, France, 2010.

[13] J.P.Šetrajčić, B.Markoski, D.Rodić, S.S.Pelemiš, S.M.Vučenović, B.Škipina and D.Lj.Mirjanić, Absorption Features of Symmetric Molecular Nanofilms,*8th ICNN*, Thessaloniki (Greece), 2011.

[14] J.P.Šetrajčić, S.S.Pelemiš, S.M.Vučenović, V.M.Zorić, S.Armaković B.Škipina and A.J.Šetrajčić, Absorption Features of Symmetric Molecular Nanofilms, Proceedings 27th MIEL, 127-130, 2010.

[15] Pelemiš S.S., Šetrajčić J.P., Markoski B., Delić N.V. and Vučenović S.M., Selective Absorption in Twolayered Optic Films, J.Comput.Theor.Nanosci. 6(7), 1474-1477, 2009.

Low-Cost Multi-Phase DC/DC Buck Converter Test Circuit with Simple Control for Testing Multi-Phase Inductors

Nikola Lečić, Akhil Chandran Mukkattu Kuniyil, Goran Stojanović and Aleksandar Pajkanović
Faculty of Technical Sciences, University of Novi Sad, Serbia
niklecic@gmail.com, akhil@uns.ac.rs, sgoran@uns.ac.rs, apaj@uns.ac.rs

***Abstract* — In this work six-phase DC/DC converter with simplified low-cost topology will be represented. Circuit control block has been developed as direct control concept, without control loop for output signal. Different operating modes are easily adjustable with DIP-4 switch. Duty cycle in all operating modes is 50%. Small resistance power resistor with 50 mΩ resistance has been positioned between ground point and load to provide possibility to circuit user to analyse transient response of test circuit. Operating frequency of circuit is externally adjustable with laboratory frequency counter. The circuit is simple for use and variable inductor topologies, single or multiphase inductors can be tested with this circuit.**

Keywords - Circuit topology, DC-DC power converters, Inductors, Testing

I. INTRODUCTION

In order to power the next generation microprocessors which require about 1 V voltage and up to 100 A current, the number of phases in the interleaved multi-phase synchronous buck converter has been increasing. Some of today's designs require as many as 8 phases. Selecting the optimum number of phases is determined by many factors, e.g. output current, system efficiency, the transient requirement, thermal management, costs of capacitors, MOSFET performance, size restriction, and overall system cost. Phase selection is further complicated due to continually changing current requirements making a scalable multiphase converter necessary [1]. A tendency to miniaturize electronic components began in the 1990s. Progress also occurred in surface-mounting technology, and attempts have been made to accomplish high density, incorporation of ferrite inductors into a printed circuit board [2]. Recent technological developments concerning electrical devices demand higher performance of the inductor i.e. higher frequency application and further miniaturization [3]. One of the major advantages of the multiphase buck converter is the ripple cancellation effect, which enables the use of a small inductance to improve the transient response and to minimize the output capacitance [4]. Thus, there are many designers working on development of new inductor topologies in order to improve performances of modern DC/DC converters.

One of the main steps in process of inductor development are operating tests which provide good feedback to designers about performances of developed inductor topologies. While working on development of new multi phase inductor topology, authors developed six-phase test circuit presented in this work.

II. TEST CIRCUIT TOPOLOGY

Six-phase DC/DC converter is built as a test circuit in order to test inductors in operating conditions. Test circuit was divided in two functional blocks: Control block and Power block.

Main part of simple control block represented in Fig. 1. is Atmel's AT TINY26 controller. This micro-controller sending control signals via circuit (M74HCT541B1R) to the power block of test circuit, actually to the gates of power MOSFETs. Control circuit was designed in such way that can be synchronized by external frequency counter, instead with crystal oscillator. This kind of realization make possible for user to change frequencies of control signal very easy during tests with simple change of synchronizing signal frequency generated at frequency counter and connected to XTAL input of controller. This control block is able to create test signal at maximal driving frequencies about 5 MHz depending on selected operation mode. DIP SW in control circuit is used to select operation modes of test circuit. Operating modes of this circuit are as follows: 1x1, 2x1, 3x1, 4x1, 5x1, 6x1, 3x2 and 2x3. In order to test on e.g. six-phase inductor performance, test circuit operating modes 6x1, 3x2 and 2x3 have been used.

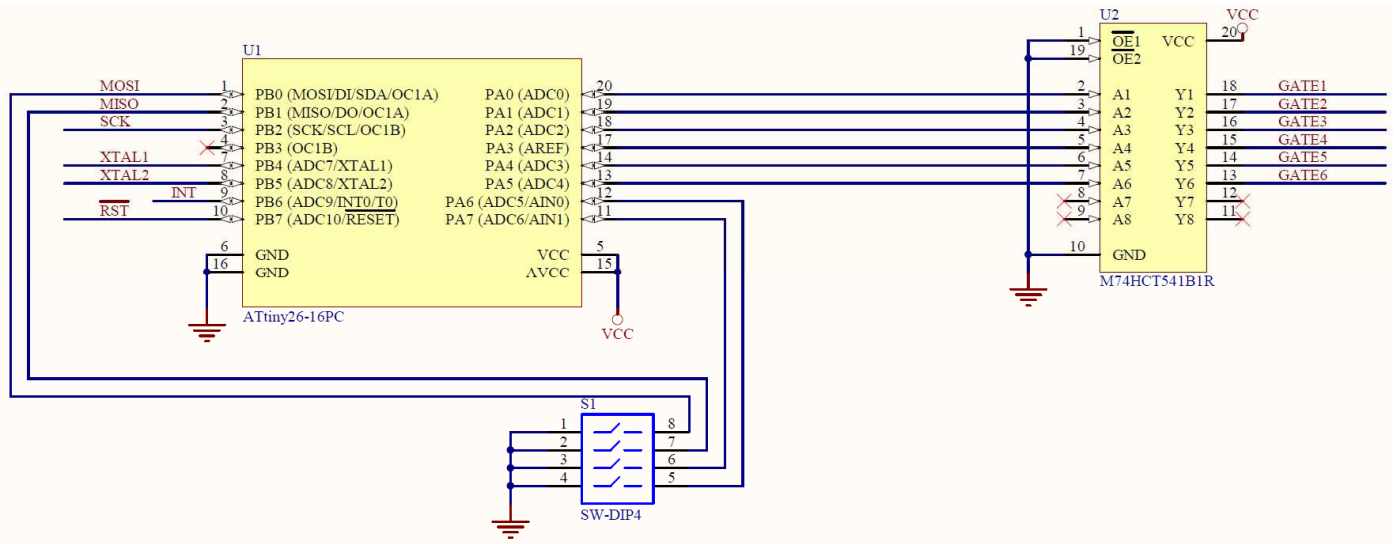


Fig. 1. Control block of test circuit

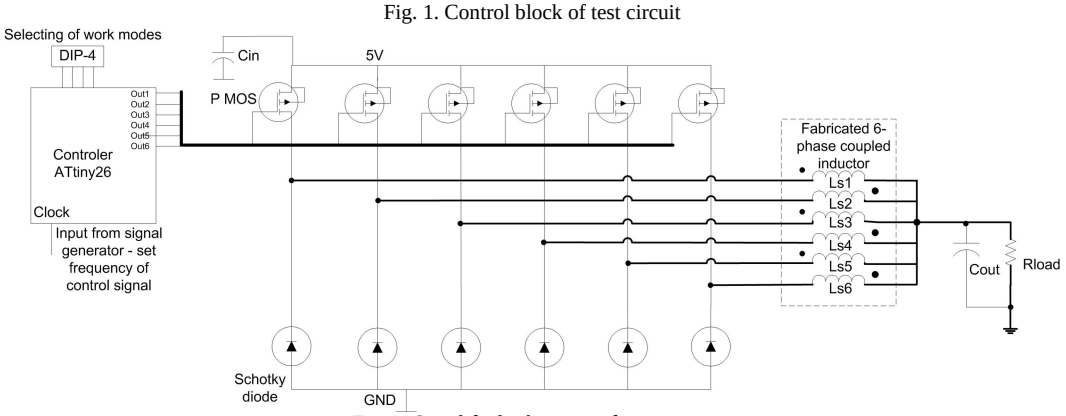


Fig. 2. Simplified schematic of test circuit

Schematic of test circuit power block was developed as simplified Buck converter concept (MOSFET-diode combination) realized with six phases. SMD P-MOS transistors IRF5305 are selected as power switch component in combination with fast Schotky diode for inductor relaxation cycle. All phases are interconnected at circuit output and provide controlled power supply signal to the load. Schematic with simplified control block and complete power block is shown in Fig. 2. Test circuit hardware with tested inductor sample is presented in Fig. 3.

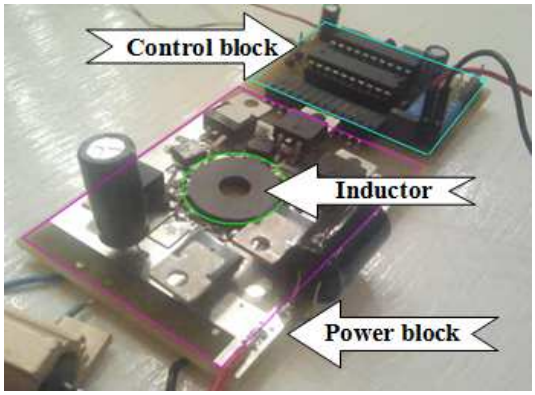
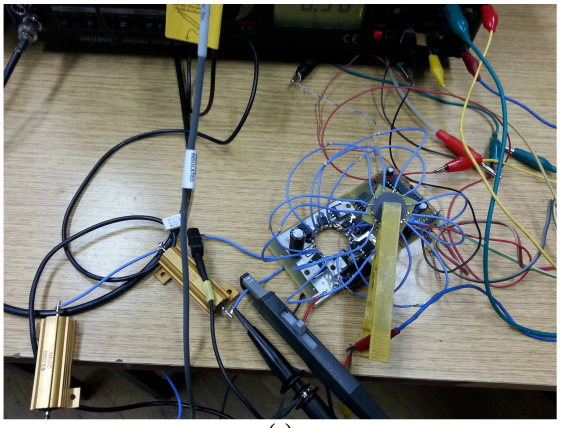


Fig. 3. Test circuit hardware

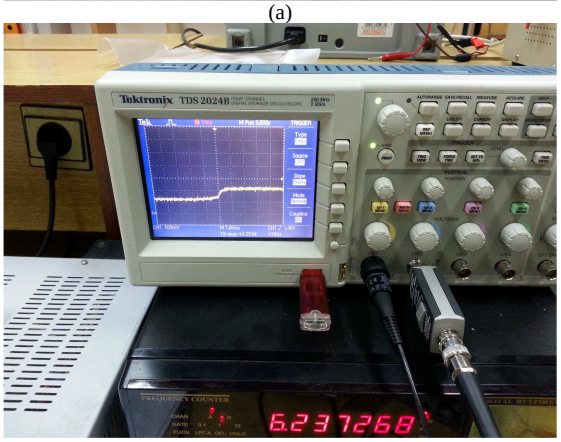
III. HARDWARE IN OPERATION MODE

Test circuit during operating tests is presented in Fig. 4a. Test equipment and frequency counter are shown in Fig. 4b. Six-phase inductor sample has been tested in the test circuit in operating mode 6x1 (Fig. 5a). Circuit output Voltage and AC current with 1 Ω power load were measured at Tektronix TDS 2024B oscilloscope and have been presented in Fig. 5b. Operating frequency was 1.1 MHz and phase inductance of inductor was 250nH.

Efficiency of the circuit is around 73% depending on selected operation mode and properties of tested inductor. In the future this efficiency can be increased by implementing additional input and output condensers and optimizing control signal operating frequency. For the purpose of initial operating analysis of inductor samples efficiency of the circuit is not important, because most of the comparative tests between commercially available and developed inductors in operation mode can be easily conducted with this circuit.

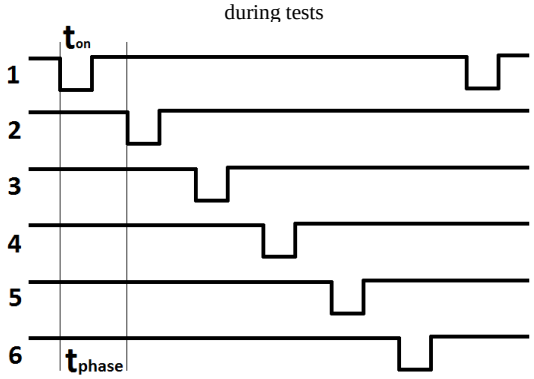


(a)

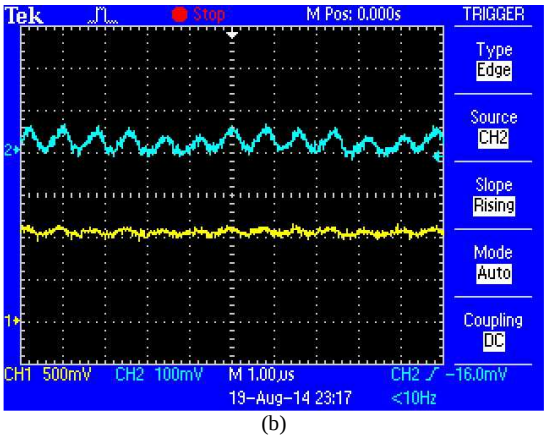


(b)

Fig. 4. Operating tests of inductor sample in the test circuit: a) active hardware with attached test probes and b) frequency counter and oscilloscope used during tests



(a)



(b)

Fig. 5. 6x1 operating mode: a) control signal and b) output voltage and AC current signal of test circuit in 6x1 operating mode

IV. CONCLUSION

Six-phase DC/DC buck converter circuit was developed with intention to test different prototypes of inductors. This test circuit was not designed to be high efficiency power supply. The main purpose of such circuit is to provide easy testing procedure to user. Operating modes can be selected easily by change of active phases in DIP4 switch combination. Adjusting of operating frequency by the external frequency counter provide comfort work with the test circuit. Different types of single inductors or multi-phase inductors up to six phases can be tested with this circuit.

ACKNOWLEDGMENT

This work is partly supported by the project SENSEIVER funded from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 289481.

REFERENCES

[1] W. Huang, G. Schuellein, and D. Clavette, "A Scalable Multiphase Buck Converter with Average Current Share Bus", Applied Power Electronics Conf. and Expo. APEC '03, Vol. 1., pp. 438-443, Feb. 2003.

[2] N. Varalaxamil, K. Sivakumar, "Sintering applications", Chapter 10: "Development of a Stress Insensitive MgCuZn-NiCuZn Composite Ferrite Useful for Microinductors Applications", published 2013.

[3] N. Varalaxmi1, K. Sivakumar, "Development of Non-Shrinking Soft Ferrite Composition Useful for Microinductors Applications", World Journal of Condensed Matter Physics, vol. 1, pp. 105-119, 2011.

[4] P. Xu, J. Wei, and F. C. Lee, "Multiphase Coupled-Buck Converter—A Novel High Efficient 12 V Voltage Regulator Module", IEEE Trans. on Power Electronics, VOL. 18, pp. 74-82 , Jan. 200

Smart meter privacy by suppression of low power frequency components

Srđan Đorđević and Marko Dimitrijević
Faculty of Electronic Engineering
University of Niš
Niš, Serbia
{srdjan.djordjevic, marko.dimitrijevic}@elfak.ni.ac.rs

Slobodan Bojanić
Escuela Tecnica Superior de Ing. de Telecomunicacion
Universidad Politecnica de Madrid
Madrid, Spain
slobodan@die.upm.es

Abstract—**This paper focuses on the problems associated with privacy protection in smart grid. We will give an overview of a possible realization of a privacy-preserving approach that encompasses privacy-utility tradeoff into a single model. This approach proposes suppression of low power frequency components as a solution to reduce the amount of information leakage from smart meter readings. We will consider the applicability of the procedure to hide the appliance usage with respect to the type of home devices.**

Keywords- load monitoring, privacy protection, smart metering

I. INTRODUCTION

The increased interest in modern power systems or smart grid is a consequence of the fact that the traditional power system is not able to meet growing energy demands. Smart grid characterizes the two-way flow of electricity and information, which allows more effectively monitoring and control of energy use. Smart meter periodically and automatically transmits readings to the utility which enables efficient load balancing. However, collecting and transmitting measuring data also poses a serious privacy threat since an unauthorized third party is able to intercept data in transmission. They could process power consumption data to extract appliance usage and observe some events. This information together with demographic data provide an opportunity to model a typical human behavior [1].

Any privacy preserving technique makes a tradeoff between privacy and data collection requirements. The first attempt to develop a general theoretical framework for privacy-utility tradeoff was proposed by Sankar et al. [2]. Such theoretical framework enables us to make a compromise between the lost of privacy and the precision of aggregated measurements. The second benefit of a theoretical abstraction is a possibility to create procedures which are technology-independent. It is shown that an optimal compromise between privacy preserving and utility can be achieved by filtering out low power frequency components.

This paper focuses on the technological defenses against unwanted and unauthorized monitoring. The following section is related to the Non-Intrusive Load Monitoring (NILM). In the third section we investigate realization of a privacy-preserving approach that encompasses privacy-utility tradeoff into a single model. The fourth section considers the applicability of the proposed technique to hide the appliance usage with respect to the type of home devices.

II. NON-INTRUSIVE LOAD MONITORING

The metering data which smart meter transmits to the utility provide an opportunity to observe the daily activities of a person remotely. If an attacker intercepts data in transmission he is able to determine individual appliance operation schedules by using NILM algorithm. Therefore the smart grid privacy solutions must carefully examine how NILM algorithm can be used to extract personal information and develop more effective means of protection from it.

In order to save energy and use energy resource efficiently it is necessary to implement the load identification of appliances in individual households. The traditional load monitoring system which is based on Supervisory Control and Data Acquisition (SCADA) requires sensors attached to each appliance in the house and a home area network. The implementation of this system is very complex. In addition, it is unreliable and not scalable as a consequence of the large number of sensors.

The second method used to detect the operation state of individual electric appliances is NILM, proposed by Schweppe and Hart [3]. This method uses the nonintrusive monitoring concept by which individual loads need not be instrumented. An NILM system uses only aggregate power consumption signals from a sensor at the power service entrance. This load monitoring method is widely accepted because of the simple hardware installation.

The appliance signature represents a specific characteristic that makes a load unique. There are two main classes of nonintrusive signatures:

- Steady-state
- Transient

The steady-state signature (SS) represents the set of parameters that are derived under steady-state operation of the appliances. The NILM methods which use this kind of signature identify devices based on the steady state change of real and reactive power. The behavior of some appliances can be tracked from positive and negative real power variations. According to the method proposed by Hart step changes of the power consumption are grouped into clusters in order to extract

individual appliance usage. The important advantages of the steady state approach are: minimal hardware requirements, low sampling frequency and additivity of the SS signatures.

Any load identification algorithm needs to employ a library of load models during pattern recognition stage. In order to build a signature library, it is necessary to make a classification of electric home appliances.

A. Electric appliance classification

The electric equipments can be classified into following six categories depending on the electrical characteristics [4]: resistive appliances, motor driven appliances, pump operated appliances, electronically-fed appliances, electronic power control appliances and fluorescent lighting.

The classification is performed based on the: duration and shape of the current transients, current harmonics and the other parameters of the voltage or current signal. Resistive appliances draw current that is in phase with the voltage. Therefore, there is the only real power present on the electrical network. The current signal drawn by resistive appliances does not contain higher harmonics.

The second category of electric equipments contains an electric motor. The current signal drawn by motor driven appliances characterize a long transient and presence of odd-numbered harmonics. The appliances which contain a pump operated by an electric motor are often classified separately from other motor driven appliances, because the differences in switching-on transients. Common household appliances that belong to this category include refrigerators, freezers, washing machines etc.

Electronically-fed appliances are low power loads which use Switch Mode Power Supply (SMPS). The fundamental characteristic of any SMPS is high power efficiency and controllability. The SMPS current has a significant amount of triplen harmonics (3^{rd}, 9^{th}, 15^{th}, etc.) and high THD.

Electronic power control appliances are loads which characteristic significantly depend on power level at which they operate. Appliances such as halogen lights, some vacuum cleaners and cookers belong to this group.

Fluorescent light sources belong to the inductive loads which characterize a substantial phase shift between current and voltage and a long two-step switching on transient. They posses current spectrum dominated by third harmonic.

III. PRIVACY-PRESERVING TECHNIQUE BASED ON PRIVACY-UTILITY TRADEOFF

There are two approaches in the privacy preserving techniques proposed so far, non-cryptographic approaches and cryptographic approaches. Cryptographic techniques allow operation of the utility without access to meter readings. The main drawback of this approach is demand to implement protocols, and software on each smart meter.

The privacy-preserving techniques can be centralized or distributed. The more common approach uses data aggregation, which performs gathering and computing data in a gateway. The second approach is to use security techniques on the side of the smart meter which communicate directly to the utility.

The technique which ensures household privacy without utility cooperation is masking and obfuscation of metering data. One such solution named, Battery-based Load Hiding (BLH) [5, 6], uses controllable batteries which are charged and discharged at strategic times to hide the load demand. The BLH algorithm tries to prevent NILM by keeping metered load constant.

The theoretical framework proposed by Sankar et al. [2] shows that optimal privacy preserving solution requires suppression of low power frequency components. Low power frequency components are typically caused by short-lived fluctuations in energy consumption and reveal a great amount of information about human behavior. In contrast, frequency components that have high power are caused by continuously running appliances and contain much less information. A privacy-preserving technique which is based on utility-privacy preserving model must realize two functions, which are as follows [7]:

- Estimation of harmonic components in power system
- Removing certain frequency components from measuring data signal

Traditionally, harmonic analysis of current or voltage signal is performed by using FFT algorithm or bandpass filtering. The FFT algorithm can be implemented only if the number of samples per period is an exponentiation of 2. Therefore, one must determine the period of the fundamental frequency and adjust sampling frequency.

The key part of the proposed privacy preserving technique is to determine active power harmonics that can be suppressed. Removing the frequency components must be done adaptively from the power consumption spectrum.

The complex power can be expressed in terms of the active and reactive power harmonics in the following way:

$$S = P + jQ = \sum_{k=1}^{N} S_k = \sum_{k=1}^{N} P_k + j\sum_{k=1}^{N} Q_k \tag{1}$$

The proposed procedure of removing low power frequency components from measuring data signal results in the approximated active and reactive power, as follows:

$$P' = P - \sum_{k \in M} P_k \qquad Q' = Q - \sum_{k \in M} Q_k \tag{2}$$

where: **M** is the set of harmonics which are suppressed.

The approximation is acceptable if the active power relative error does not exceed the upper limit, δ_p.

$$\frac{P'-P}{P} < \delta_P \tag{3}$$

All other power quality measurements are performed without any approximation since they do not reveal privacy information.

IV. MEASUREMENTS

A. Measurement conditions and setup

Measurements and data analysis are performed using a real-time system for polyphase nonlinear loads analysis, described in [8]. The system is based on virtual instrument paradigm, using National Instruments NI9225 and NI9227 acquisition modules and PXI controller running real-time operating system (RTOS). In these measurements, we used 50 kSa/s sampling rate and 24bit accuracy. The current and voltage ranges are $\pm 5A_{RMS}$ and $\pm 300V_{RMS}$, respectively. All measurements are conducted using one acquisition channel, i.e. one phase.

The current spectra and waveforms are shown on Fig. 1 to 5. Fig.1 presents a current spectrum and waveform in case of pure linear resistive load, 100 W nominal power incandescent lamp. This measurement is given as reference. The waveform is almost sinusoidal, and one can observe small magnitude of 3rd, 5th and 7th harmonics. These harmonics exist due to power grid supply voltage, which is not pure sinusoide, having some harmonic pollution.

B. Results

Fig. 2 and 3 depicts measured results for Compact Fluorescent Lights (CFL) lamps with 15W and 20W nominal power, respectively. The waveforms are almost identical, and both spectra have the same harmonic structure, with odd harmonics. The envelope of spectra is approximately exponentially decaying function, with a small drop in 5th harmonic. The only observable difference between two spectra is different magnitudes of same order harmonics.

The measured data related to the CRT monitor are shown in Fig. 4. The waveform resembles the sine function. The spectrum contains low frequency components, i.e. 5th–9th harmonics, with small magnitude. This is an example of well compensated SMPS.

Finally, Fig. 5 represents spectrum and waveform obtained by measuring portable PC supply current. This is an example of time non-invariant electronic appliance. The power consumption, as well as spectrum structure of such nonlinear load varies in time, depending on working conditions (i.e. battery status) and activities within a PC, causing different CPU, GPU and I/O load [9]. This information can be also used for eavesdropping the computer activity, as shown in [10].

The measurements shown in Fig. 5 are performed with fully charged battery, when PC is in idle state – only OS and core services running. The AC/DC power converter draws current from the grid in bursts, causing heavily distorted waveform. The corresponding spectrum contains odd harmonics, with sinc function as envelope.

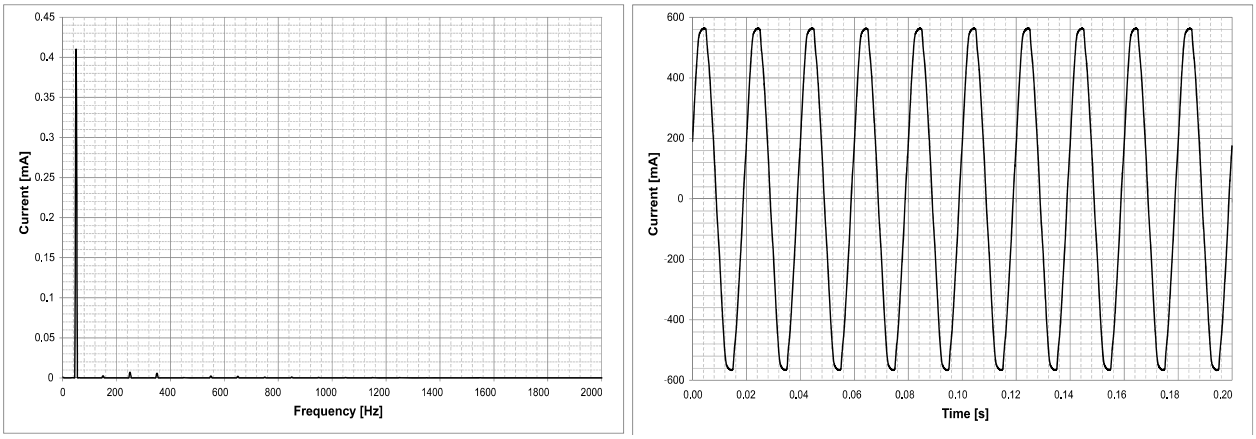


Figure 1. Incandescent lamp, 100W nominal power – spectrum (left) and waveform (right)

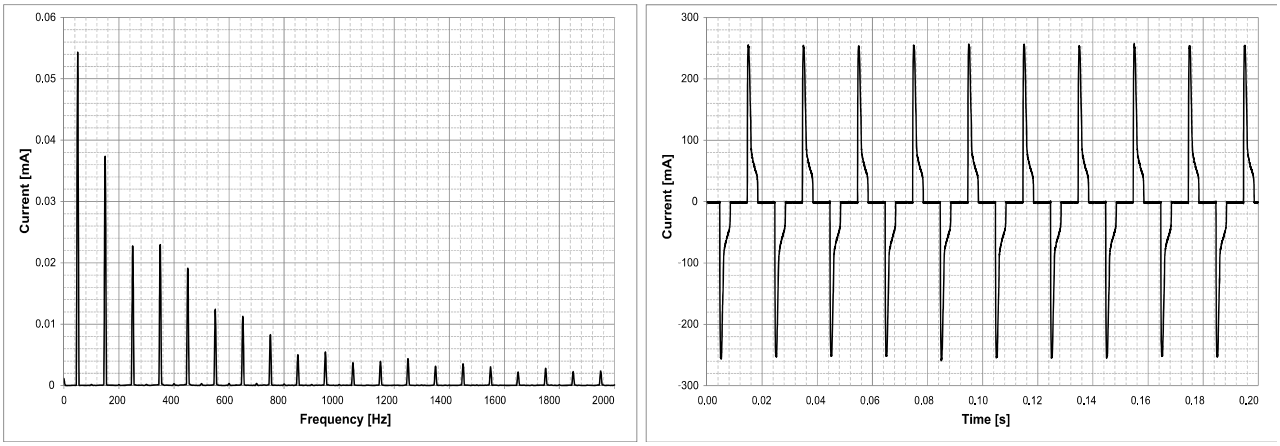


Figure 2. 15W nominal power CFL – spectrum (left) and waveform (right)

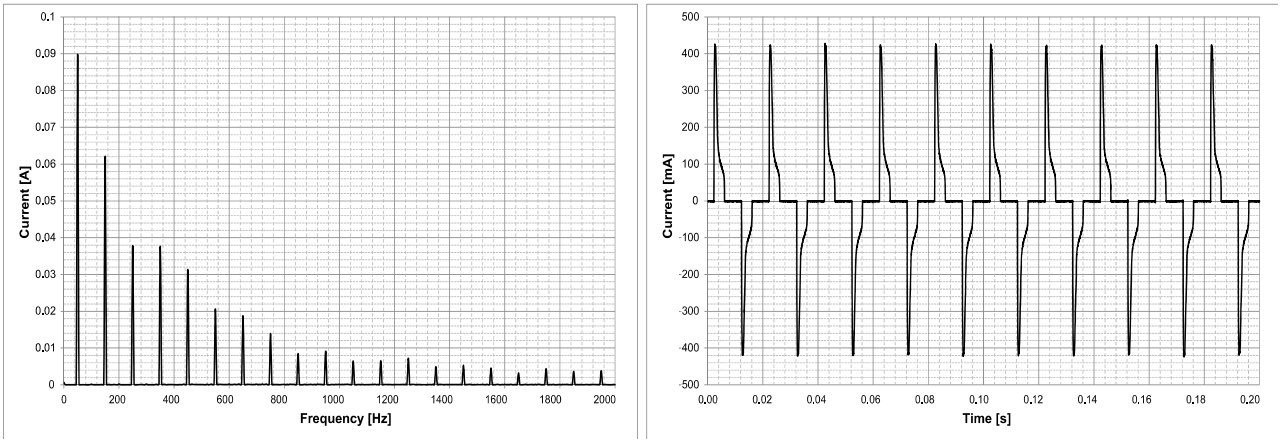


Figure 3. 20W nominal power CFL – spectrum (left) and waveform (right)

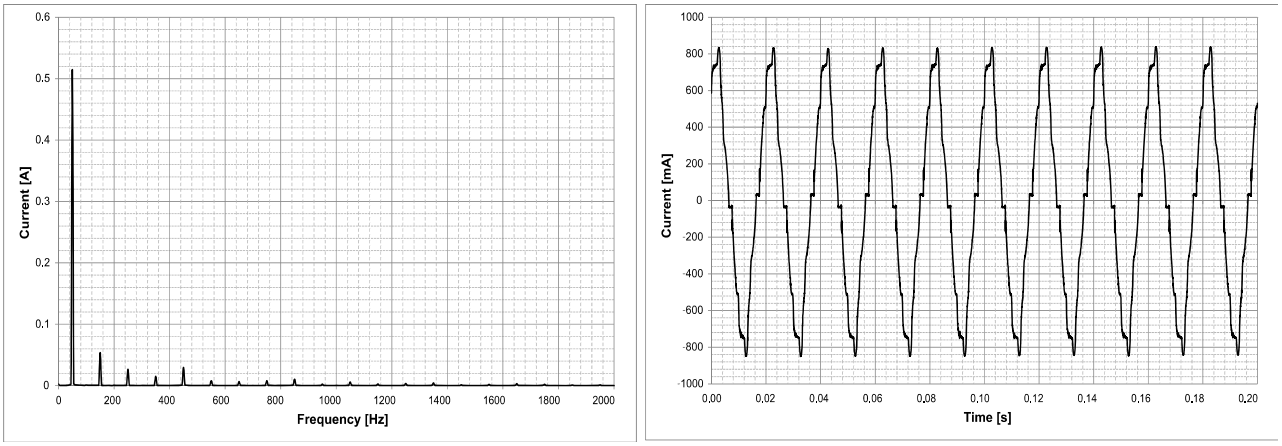


Figure 4. CRT monitor – spectrum (left) and waveform (right)

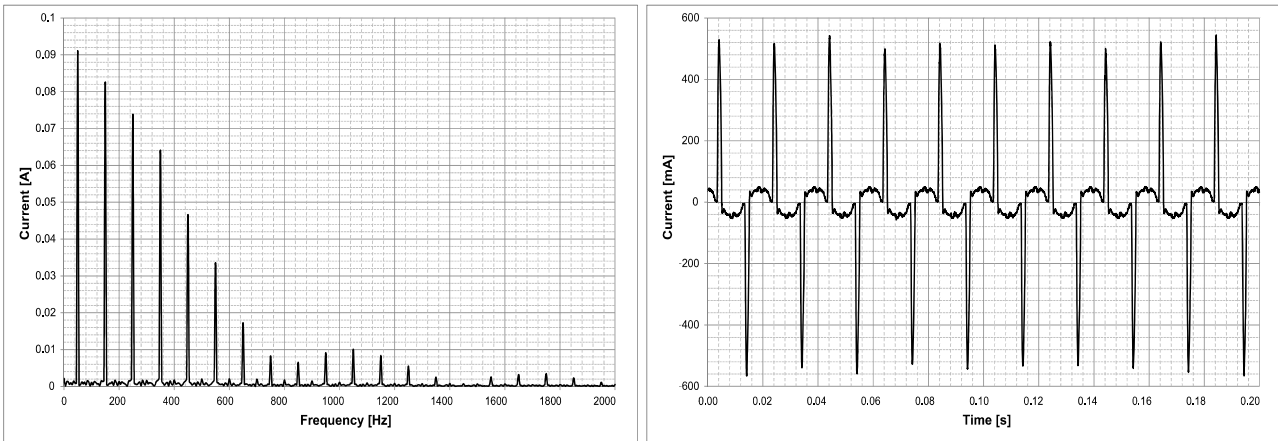


Figure 5. Portable PC – spectrum (left) and waveform (right)

V. CONCLUSION

By observing measured results, especially related spectra, one can conclude that each electronic appliance has unique harmonic signature, which can be used for identification. Identifying single appliance connected into power grid requires complex pattern based algorithms.

In this proceeding, we only take in consideration harmonic magnitudes. Harmonic phases can also be measured with described system, providing more information needed for appliance identification.

The proposed privacy-preserving technique is applicable to appliances with a current spectrum having relatively strong higher harmonics. Hence, the suppression of low-power frequency components is not effective on resistive appliances. According to the measurement results, the usage of electronically-fed and electronic power control appliances can be effectively hidden. This group of appliances encompasses

household appliances such as IT equipment (mainly computers), television sets, energy-efficient lighting (LED and CFL) which are becoming nowadays more prevalent in households.

ACKNOWLEDGMENT

This research was partly funded by The Ministry of Education and Science of Republic of Serbia under contract No.TR32004.

REFERENCES

[1] M. A. Lisovich, D. K. Mulligan, S. B. Wicker, "Inferring Personal Information from Demand-Response Systems," Security & Privacy, IEEE, vol. 8, no. 1, pp. 11-20, 2010.

[2] L. Sankar, S. Rajagopalan, S. Mohajer, H. Vincent Poor, "Smart Meter Privacy: A Theoretical Framework," *smart grid*, *IEEE transactions on*, vol. 4, no. 2, pp. 837-846, Jun 2013.

[3] G. Hart, "Nonintrusive appliance load monitoring," Proceedings of the IEEE, vol. 80, no. 12, pp. 1870–1891, 1992.

[4] F. Sultanem, "Using appliance signatures for monitoring residential loads at meter panel level," IEEE Transactions on Power Delivery, vol. 6, no. 4, pp. 1380–1385, 1991.

[5] S. McLaughlin, P. McDaniel, W. Aiello, "Protecting consumer privacy from electric load monitoring," In Proceedings of the 2011 ACM CCS: 87-98.

[6] W. Yang, N. Li, Y. Qi, et al. "Minimizing private data disclosures in the smart grid," In Proceedings of the 2012 ACM CCS: 415-427.

[7] S. Đorđević, S. Bojanić, "Tehničko rešenje za zaštitu privatnosti u novim elektroenergetskim mrežama zasnovano a optimalnom balansu između privatnosti i funkcionalnosti," Proc. of the LVIII Conf. of ETRAN, 2013, Vrnjačka Banja pp. EL1.1.

[8] M. Dimitrijević: "Electronic System for Polyphase Nonlinear Load Analysis," in serbian, PhD thesis, December 2012.

[9] O. Nieto-Tialdriz, M. Dimitrijević, D. Stevanović, D. Mirković "Energy Profile of a Personal Computer," Proc. of the LVI Conf. of ETRAN, 2012, Zlatibor pp. EL3.3-1-4.

[10] M. Andrejević Stošović, M. Dimitrijević, V. Litovski, "Computer Security Vulnerability Seen From the Electricity Distribution Grid Side," Applied Artificial Intelligence, Taylor & Francis Ltd., vol. 4, no. 28, pp. 323-336, London, 2014.

Computer-Assisted Performance Assessment of Outdoor Substation Grounding Systems

Boško Mijatović
Department for Energy and Telecommunications
Civil Engineering Institute "IG"
Banja Luka, Bosnia and Herzegovina
bosko8@gmail.com

Čedomir Zeljković
Faculty of Electrical Engineering
University of Banja Luka
Banja Luka, Bosnia and Herzegovina
cedomir@etfbl.net

***Abstract*—The performance of grid grounding system is assessed by using a commercial software package. The most influential input variables are systematically varied and their impact on the system is observed and discussed. The corrective measures are suggested in order to bring the design parameters of the grounding system within their permissible limits.**

Keywords-computer aided design; grounding system; outdoor substation; performance assessment

I. INTRODUCTION

The ultimate purposes of substation grounding systems are to keep the people safe from dangerous electrical shock inside or near the substation and to provide the drainage of fault currents while maintaining the reliable operation [1]. Successful design of a grounding system comes down to bringing a touch voltage and step voltage into limits defined by regulation [2]. In order to meet regulation requirements, the grounding system design may often lead to a complex geometric shapes which stretch over a large area. On the contrary, computation of the ground potential rise in the closed form is possible only for special cases having simple geometry. In realistic complex grounding systems, it is necessary to use some numerical methods, such as finite element method [3], boundary element method [4] or similar, for a precise computation.

In recent years, the developed mathematical methods are incorporated into commercial software packages which distinctly facilitated the process of the system design and visualization of results. Using a software package, it is possible to systematically vary the input variables and to analyze their effect on the grounding system within a reasonable period of time. In this paper we use the program CYMGRD provided by the CYME International T&D company for the design and analysis of a real grounding system for an outdoor high-voltage substation [5].

II. THEORETICAL BACKGROUND

The elements of electrical power system are grounded in order to maintain their potential at approximately the potential of earth. To provide a low-impedance contact, the grounding system inevitably contains a set of metal components buried underground. During the faults or highly unbalanced power system operation the grounding system conducts some current. Since the grounding impedance is never as low as zero, this current always produces some voltage drop. Therefore, the potential rise is occurred on the grounded masses. Due to the current flow through the ground, the potential of the soil surface around the grounding system is also increased in comparison with the potential of remote earth. A typical chart of the soil surface potential above the grounding grid is shown in Fig. 1.

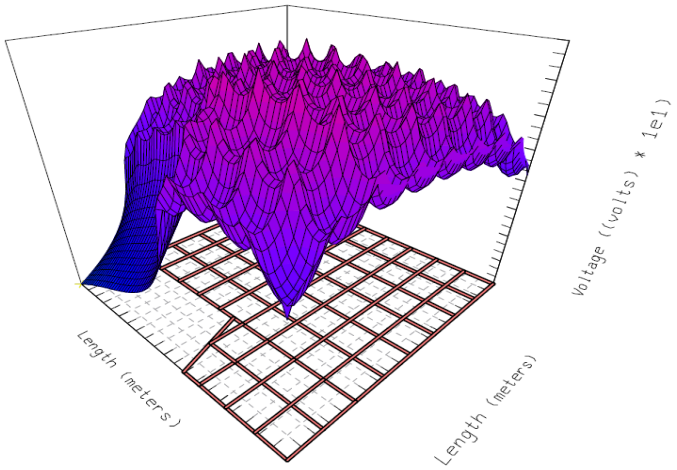


Figure 1. Illustrative chart for the soil voltage above the grounding system

As a result of the potential rise, a plenty of hazardous situations are possible to occur inside or outside of a substation. Basic electrical shock situations are shown in Fig. 2. The following labels are used in the figure: E_{mm} is the metal-to-metal touch voltage, E_k is the step voltage, E_d is the touch voltage, and E_{ip} is the transferred voltage.

Dangerous metal-to-metal touch voltage can be avoided with appropriate equipment positioning. It is not a part of grounding system design. The most common way to avoid transferred voltage endangerment is to isolate cable armoring and electrical protection from main substation grounding. After all, main challenges in grounding system design are to provide compliance of maximum touch and step voltages with their allowable values. Selected grounding grid design should provide safety in substation exploitation but also taking into account a financial aspect of grounding construction and maintenance.

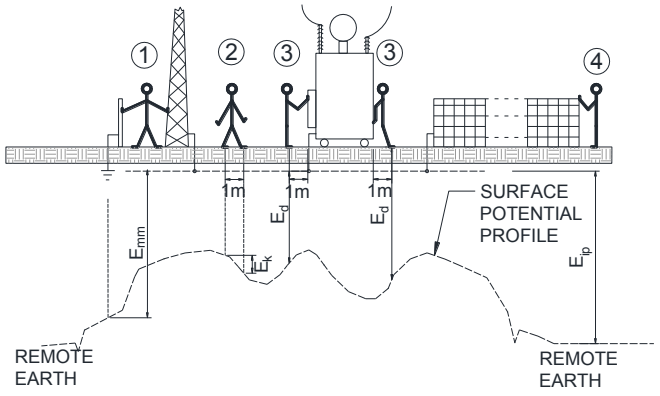


Figure 2. Basic shock situations (modified from [2])

Exposure to touch voltage is illustrated at Fig. 3. U stands for the phase-to-ground potential, Z is the system impedance, I_f is the fault current, I_g is the grounding current, I_b is the body current, R_g is the resistance of grounding and equipment to the point H, R_b is the body resistance, R_{sl} is the surface layer resistance, H is the hand contact point, and F is the foot contact point.

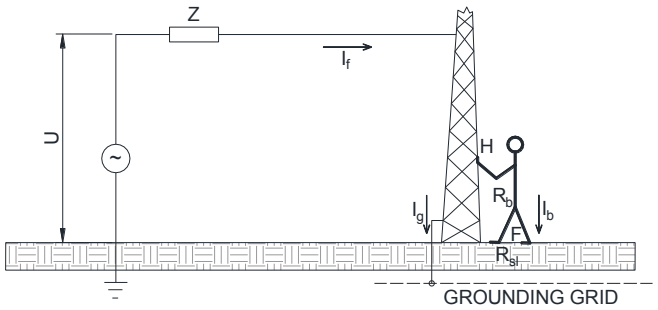


Figure 3. Exposure to touch voltage (modified from [2])

Equivalent touch voltage circuit, corresponding to Fig. 3 is shown in Fig. 4.

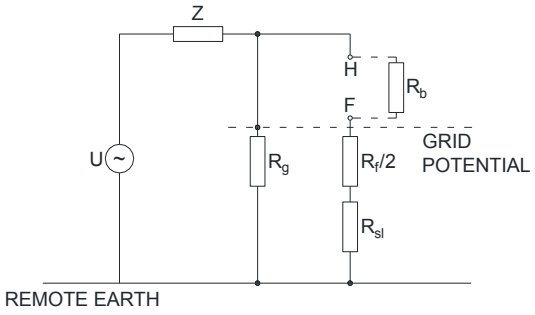


Figure 4. Equivalent touch voltage circuit

The potential of the contact point H corresponds to grounding potential:

$$U_H = I_g \cdot R_g = \frac{U}{Z + R_g} \cdot R_g = \frac{U}{\frac{Z}{R_g} + 1} \tag{1}$$

In Eq. 1, Z and U are system defined and they cannot be affected by grounding design. U_H can be controlled by changing grounding resistance R_g. The potential of the contact point F depends on the surface potential distribution. Distribution of potential is a function of grounding system arrangement and grounding potential, which is a key task of grounding system design.

Current through the body I_b after contact between points H and F will be:

$$I_b = \frac{U_H - U_F}{R_{f/2} + R_{sl} + R_b} \tag{2}$$

In Eq. 2, body resistance R_b cannot be affected by grounding design. Body current I_b can be controlled by the foot contact resistance R_f, the surface layer resistance R_{sl} and the touch potential (potential difference U_H-U_F). R_f is a subject of safety at work and can be increased by using protective shoes. The other two variables are subject of grounding system design.

Exposure to touch voltage is represented in Fig. 5, where U_{F1F2} is the potential between point points F_1 and F_2, I_b is the body current on the leg-to-leg path, R'_b is the body resistance leg-to-leg, and F_1, F_2 are the foot contact points.

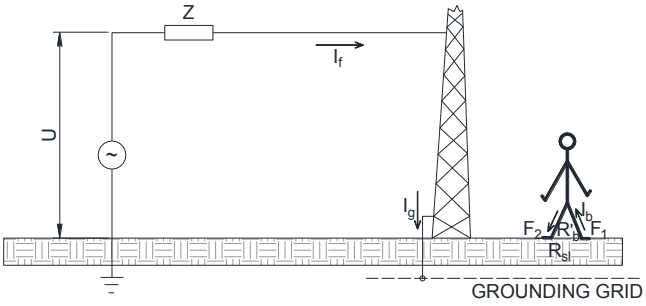


Figure 5. Exposure to step voltage (modified from [2])

Equivalent step voltage circuit, corresponding to Fig. 5 is represented in Fig. 6.

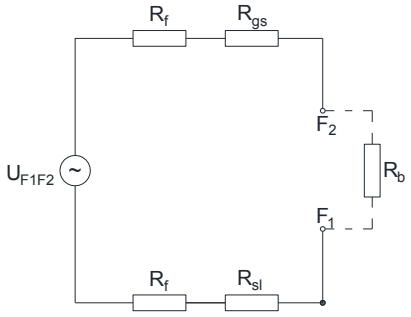


Figure 6. Equivalent step voltage circuit

The body current under potential difference between the foot contact points F_1 and F_2 (step potential) will be:

$$I_b = \frac{U_{F1F2}}{2 \cdot R_f + 2 \cdot R_{sl} + R'_b} \tag{3}$$

Similar as at touch potential calculation, fixed part of Eq. 3 is R'_b, and R_f is a subject of safety at work. Subject of grounding system design are U_{F1F2} and R_{sl}.

Although the body current is only relevant for electrical hazard assessment, touch and step potentials are more favorable for practical use in design, testing and maintenance. Relations between body current and touch (step) potential are fixed for specific grounding system, fault current and body resistance, as it is represented in Eq. 2 and Eq. 3. Maximum allowable body current corresponding to maximum allowable touch and step voltage. Therefore, in the reminder of the paper, the touch and step potentials will be the outputs of the primary importance.

III. TEST CASE

A. Substation Description

A test case in this paper is a real outdoor substation for connection of a hydro power plant on a 220kV overhead line. Substation is located on a double layer soil (causeway) with lower layer resistivity of 100Ωm and upper 8m thick layer with approximate resistivity of 800Ωm. Live-to-ground fault current intensity of 9.268kA under -83.2° angle is obtained from a separate study. All calculations are performed for a fault duration of 0.1 seconds, which is a typical time of fault detection and disconnection without automatic reclosing system.

Switchgear disposition within the substation is shown in Fig. 7.

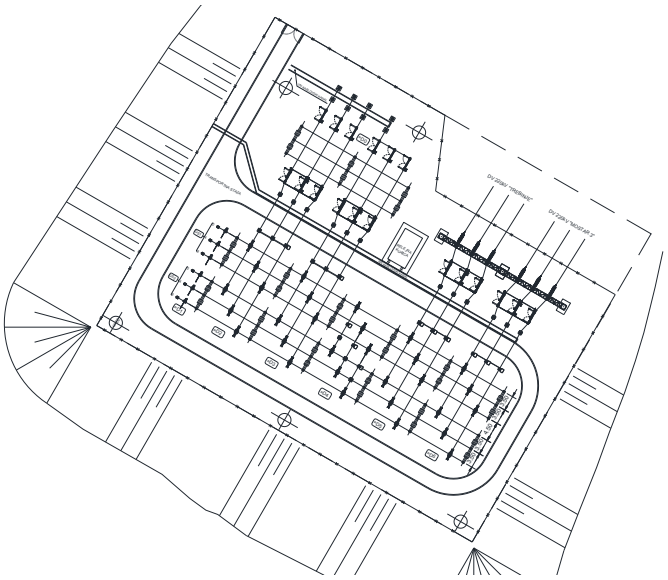


Figure 7. Disposition of the switchgear

B. Grounding System

Substation grounding layout is chosen to ensure connection of metal equipment, low grounding resistance, as well as protection from dangerous step and touch voltage. Minimum size of grounding grid is determined by size of area under switchgear equipment, maximum size by substation area (inside the fence).

There are two approaches to substation fence grounding. First and widely used is separation from primary grounding in order to prevent danger voltage transfer outside of protected area. Second approach is metal connection of fence grounding and primary grounding. Both approaches will be considered in simulations presented in this paper.

Configuration of grounding system designed for substation from Fig. 7 is represented in Fig. 8. Grounding grid layout from Fig. 8 is with separated grounding fence widely used for grounding systems ground in low resistivity soil.

In most of the cases, especially in high resistivity soil, use of grounding system from Fig. 8 will not ensure protection from dangerous step and touch potential. In this case it will be necessary to provide additional steps to upgrade the performance of the grounding system.

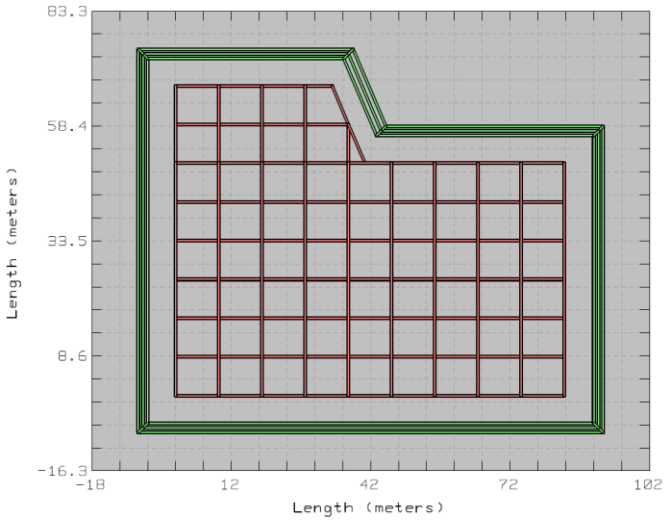


Figure 8. The configuration of the grounding system

The key variables affecting the performance of outdoor substation grounding systems limited to area in a surrounding fence are:

- use of high resistivity soil as surface layer,
- dimension of area used for grounding grid,
- use of grounding rods,
- use of low resistivity soil as laying bed for grounding grid system.

Influence of above listed variables to grounding system will be reviewed in this paper and represented in figures and table below.

IV. SIMULATION RESULTS AND DISCUSSION

A. Software

Considering the complexity of electromagnetic processes around grounding electrodes, as well as extensive dimension of substation grid, computer assistance is important factor in modern outdoor substation grounding system design. Simulations presented in this paper were performed using dedicated software CYMGrd, provided by CYME International T&D Company.

B. Simulation Results

1) Reference Case

Simulations of the grounding system in this paper are provided for low resistivity ground with resistivity of 100Ωm, as well as for double layer soil with high resistivity upper layer, as described for subject substation in section III A. The basic results are shown in Table I (for low resistivity) and Table II (for high resistivity).

TABLE I. GROUNDING SYSTEM GROUNDED IN LOW RESISTIVITY SOIL

Maximum step potential (V)	Maximum permissible step potential (V)	Maximum touch potential (V)	Maximum permissible touch potential (V)
142,51	2879,57	2646	1092,25

TABLE II. GROUNDING SYSTEM GROUNDED IN HIGH RESISTIVITY SOIL

Maximum step potential (V)	Maximum permissible step potential (V)	Maximum touch potential (V)	Maximum permissible touch potential (V)
774,64	2879,57	9567	1092,25

The simulation results for step and touch potential represented in Table 1 and Table 2 exceed their maximum permissible values. It is thus necessary to take further action to get them reduced bellow the maximum permissible thresholds.

2) Grounding System with High Resistivity Surface Layer

Adding a thin surface layer of a high resistivity material such as gravel may be a useful corrective measure for bringing the step and touch potentials within the permissible limits. The impact of a surface layer thickness is simulated in CYMGRD software and the results are summarized in Fig. 9 and Fig. 10.

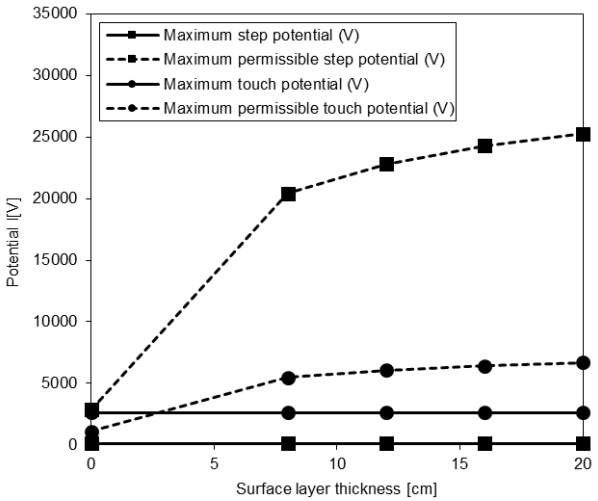


Figure 9. Grounding in low resistivity soil with surface layer

It can be noticed that using of the surface high resistivity layer will affect the touch and step maximum permissible potential, but not the maximum step and touch potential. By using the obtained results, it is possible to assess the effectiveness of this corrective method. For the system grounded in low resistivity soil (Fig. 9), a surface layer of 3 cm is sufficient to set the maximum step and touch potentials within their permissible limits. Nevertheless, a minimum thickness of 10 cm is used in practice, in order to ensure compact layer with proper granulation of material. For grounding system in high resistivity double layer soil (Fig. 10), using of the surface layer will not ensure meeting of maximum permissible touch potential and additional steps should be performed. As the simulations show, the influence of surface layer is limited for layer thickness greater than 10 cm at low resistivity and 15 cm at high resistivity soil.

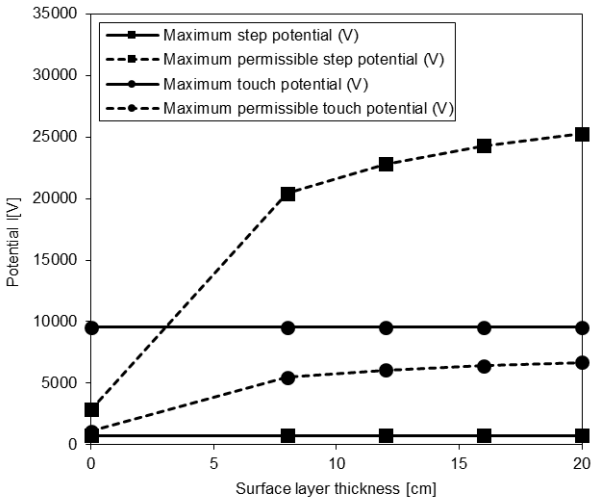


Figure 10. Grounding in high resistivity double layer soil with surface layer

3) Use of Grounding Rods

Fig. 11 and Fig. 12 represent simulation results of primary grounding system from Fig. 8, with additional grounding rods. Grounding system is grounded in low resistivity soil and double layer soil with high resistivity upper layer, same as the reference case. Figures are drawn for variable length of the grounding rods.

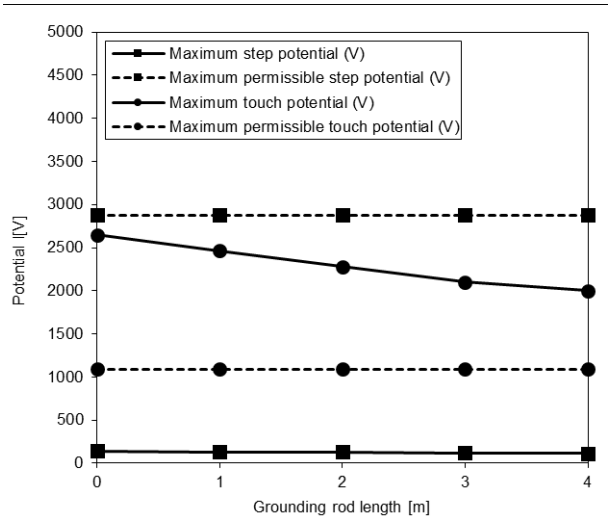


Figure 11. Grounding in low resistivity soil with grounding rods

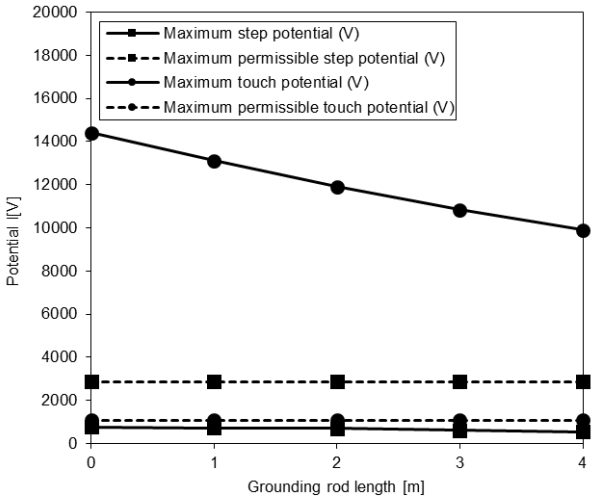


Figure 12. Grounding in high resistivity double layer soil with grounding rods

The simulation results show a greater efficiency of grounding rods in case of grounding system grounded in high resistivity soil. For grounding system in low resistivity soil, the influence of grounding rods on touch and step potential becomes less dominant as the length of the rods increases. Regardless of its positive effects, the use of grounding rods will never decrease the maximum touch potential below the permissible value, for any type of soil and any grounding rod length.

4) Use of Low Resistivity Soil as Encasement for Grounding Rods

Fig. 13 and Fig. 14 illustrate how the grounding rods encased in low resistivity soil layer affect the characteristics of the grounding system. Primary grounding system is grounded in low resistivity soil or double layer soil with high resistivity upper layer, without low resistivity laying bed. Grounding rods used in calculation are 4 m long, with a diameter of 63 mm, embedded in primary grounding grid junction points. Figures are drawn for fixed length of grounding rod and variable thickness of encasement layer.

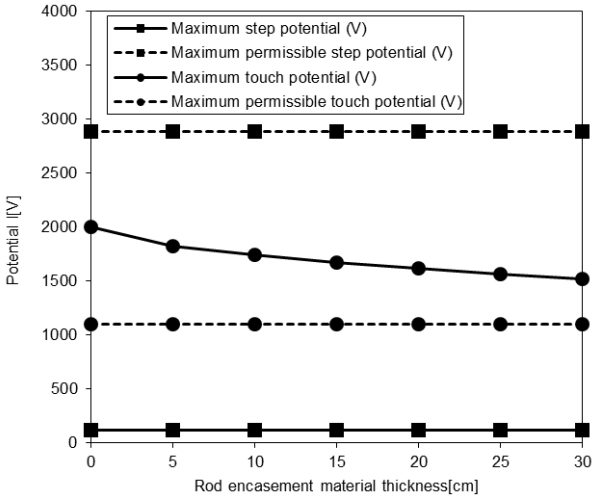


Figure 13. Grounding in low resistivity soil with encased grounding rods

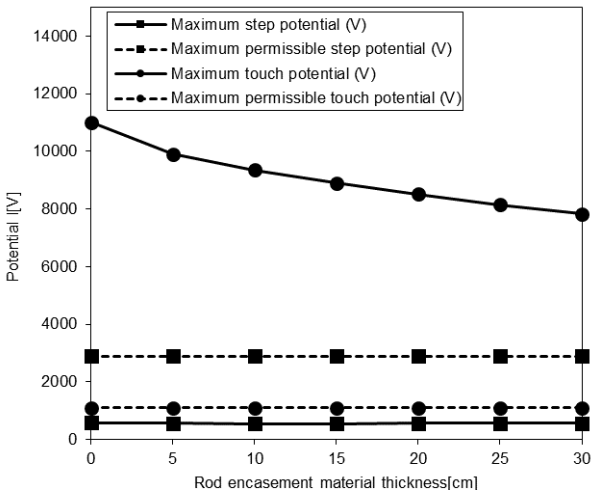


Figure 14. Grounding in high resistivity double layer soil with encased grounding rods

The simulation results show a greater efficiency of grounding rods encasement in case of grounding rods grounded in high resistivity soil. Positive influence to step and touch potential will drop down after encasement material thickness greater than 5 to 10 cm. Influence to step potential will even be negligible because of increasing fault current density at the zone of grounding rods. Increasing of current density will increase gradient of potential at the soil surface, and consequently step potential in the proximity of grounding rod.

5) Widening of Area Used for Grounding Grid

All previous simulations of grounding system grounded in double layer soil with high resistivity upper layer shown significantly greater touch potential from its permissible limits. In order to improve performance of grounding system in high resistivity soil, it is necessary to simultaneously use two or more elaborated additional measures.

The final part of simulation includes a selection of three variant solutions for grounding system grounded in double layer soil with high resistivity upper layer. Each solution involves widening of primary substation grounding grid to total available area, with connection of fence grounding to substation primary grounding system. The characteristics of selected grounding system variant solutions are represented in Table III. Variant solution 1 from Table III includes a wider grounding system with surface high resistivity layer, without grounding rods. Variant solutions 2 and 3 are based on variant 1, modified with grounding rods length of 2 m and 3 m, respectively.

Although all variant solutions presented in Table III fulfill criteria for maximum permissible touch and step potential, they are different from techno-economical aspect. Variant solution 1 is most economically preferred, but solution 2 is less sensitive to drainage conditions, which make it technically preferred. Variant 3 is a variation of variant 2, a bit expensive, but with a greater safety margin in maximum touch potential. Considering benefits of all variant solutions, solution 3 is found as the most favorable for high-resistivity dry soils.

TABLE III. VARIANT SOLUTIONS OF GROUNDING SYSTEM GROUNDED IN HIGH RESISTIVITY SOIL

Description	Variant solution 1	Variant solution 2	Variant solution 3
Grounding resistance (Ω)	1,45	1,42	1,39
Primary grounding conductor length (m)	2222	2222	2222
Grounding rod length (m)	0	2	3
Total length of grounding rods (m)	0	196	294
Rod encasement material thickness (cm)	0	0	0
Surface layer thickness (cm)	15	12	12
Maximum step potential (V)	575	559	611
Maximum permissible step potential (V)	23960	22810	22811
Maximum touch potential (V)	6223	6011	4831
Maximum permissible touch potential (V)	6363	6075	6075

V. CONCLUSIONS

In this paper the application of CYMGRD software to the performance assessment of grid grounding systems is demonstrated. The design process is illustrated on a real outdoor high-voltage substation. The first step was the identification of the most influential input variables. The inputs are then systematically varied and their impact to the performances of the grounding system are computed and discussed. Some useful corrective measures are suggested.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of CYME International T&D company which provided a time limited fully functional version of the CYMGRD software.

REFERENCES

[1] A. P. Sakis Meliopoulos, Power System Grounding and Transients, New York*:* Marcel Dekker*,* 1988.

[2] IEEE Std. 80-2000, IEEE Guide for Safety in AC Substation Grounding, IEEE: Institute of Electrical and Electronic Engineers, Inc., New York, 2000.

[3] J. He, R. Zeng, and B. Zhang, Methodology and Technology for Power System Grounding, Singapore:Wiley, 2012.

[4] I. Colominas, J. Gómez-Calviño, F. Navarrina, and M. Casteleiro, "Computer analysis of earthing systems in horizontally or vertically layered soils", Elect. Pow. Syst. Res., vol 59, no. 3, pp. 149-156, October 2001.

[5] CYMGRD, Substation Grounding Program, [Online]. Available: http://www.cyme.com/software/cymgrd/

Cost Effectiveness of a Control Strategy for Grid-Connected Photovoltaic Systems

Predrag Mršić, Čedomir Zeljković
Faculty of Electrical Engineering
University of Banja Luka
Banja Luka, Republic of Srpska, BiH
predrag.mrsic@etfbl.net

Nikola Rajaković
School of Electrical Engineering
University of Belgrade
Belgrade, Serbia
rajakovic@etf.rs

Abstract—**The electricity customers may use photovoltaic systems supported by batteries in order to fulfill a fraction of their energy requirements and to decrease the peak demand. The achievable savings primarily depend on a system control strategy. In this paper, one algorithm based on a threshold control is described and tested. The sensitivity on the most important input variables is analyzed by extensive set of numerical simulations.**

Keywords - battery energy storage system, distribution network, peak shaving, photovoltaic system

I. INTRODUCTION

In recent years, the photovoltaic (PV) systems remain the most rapidly growing energy technology worldwide. The installed PV capacity surpassed 100 GW in 2012 [1]. Nearly 38 GW was added in 2013, while new 40 GW are expected to be installed until the end of 2014 [2]. Electricity customers are encouraged to invest in such systems and to use them in a combination with grid supply in order to minimize their energy bills. The PV systems are often equipped with the battery energy storage systems (BESS) to increase the control flexibility through storing surpluses and compensating deficits of solar energy.

Different control strategies for minimization of the customer electricity bill have been proposed in the literature [3,4]. However, the studies were mainly focused on the minimization of the customer volumetric energy costs. On the other hand, the tariffs for industrial and commercial customers often include the monthly peak demand component as well. In this paper, we propose a control strategy for the PV systems that balance between the peak shaving and reducing the cost of energy.

The first steps in this study are based on simplified deterministic models. The control strategy is tested and the maximum achievable savings are calculated. As the realistic inputs such as the solar radiation and the customer load have the stochastic nature, their values may deviate from the expected values. In the latter part of the study, the inputs of importance are altered in order to quantify their influence on the final results.

II. DESCRIPTION AND MODELING OF THE SYSTEM

The system under consideration is depicted in Fig. 1. The customer load is supplied either by the utility grid or by the photovoltaic system. If the power available from the PV system exceeds the load requirements, the surplus may be transferred either to the grid or be stored in the battery for the later use. Possibility of charging the battery from the grid is not covered in this paper. In the reminder of the section, the individual models of the system components are described.

A. Photovoltaic System and Inverter

The photovoltaic system is one of the three sources for supplying the customer load. PV system and inverter are integrated into a joint model. Our model rely on the National Renewable Energy Laboratory modeling methodology documented in [5], which is implemented in a web calculator PVWatts. PVWatts combines a number of sub-models to predict overall system performance and includes several built-in parameters. The inputs for this model are irradiation, DC rating of the PV system, tilt, azimuth, efficiency of the modules and efficiency of the inverter. By entering data for the DC rating of the PV system and location of the system, tilt and azimuth in PVWatts calculator the generation of the PV system is computed. The results of the calculation should be interpreted as being a representative estimate for a similar actual system operating in a year with typical weather. The errors may be as high as 10 percent of annual energy for weather data representing long-time typical conditions.

B. Battery and Converter

The battery energy storage system capacity is expressed by the nominal capacity and number of autonomy hours. We use BESS for storing energy when PV system produces more energy than demanded by the load. We also use BEES for load supply when the PV system produces less energy than load demand.

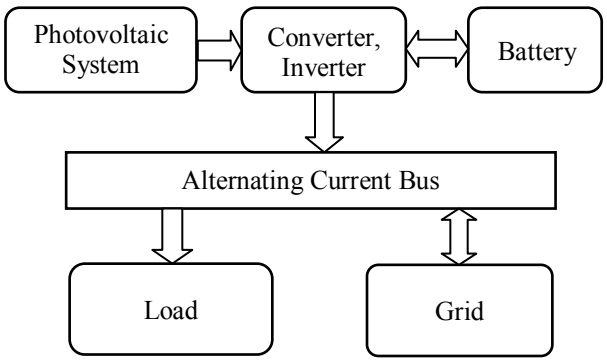


Figure 1. The grid-connected PV system

Nomenclature

Abbreviations

BESS Battery energy storage system
PV Photovoltaic

Indices

h Hour

Parameters

η_{ch} BESS charging efficiency (%)
η_{dch} BESS discharging efficiency (%)
η_{inv} DC/AC electric conversion efficiency (%)
δ BESS hourly self-discharge rate (%)
Δt Sampling interval [h]
W_{ref} The nominal capacity of BESS [kWh]
W_{Bmin} Minimum stored energy in the BESS [kWh]
W_{Bmax} Maximum stored energy in the BESS [kWh]
P_{Bmax} Maximum charge/discharge rate of the BESS [kW]
P_{Gthr} Threshold for the power supplied by the grid [kW]
P_{pvdc} DC rated power [kW]
C_{off} Electricity prices (Off-peak) [BAM/kWh]
C_{on} Electricity prices (On-peak) [BAM/kWh]
C_f Feed-in tariff [BAM/kWh]
C_{pd} Peak demand price [BAM/kW]
T_{on} Time of occurrence of the on-peak daily tariff [h]
T_{off} Time of the end of the on-peak daily tariff [h]

Variables

P_{Gmax} The monthly grid peak demand [kW]
P_B Charge/discharge rate of the BESS [kW]
E_B Stored energy in the BESS [kWh]
P_L The electrical load of the customer [kW]
P_{pv} The electricity generated by the PV system (AC) [kW]
P_G Power supplied by the grid [kW]
P_N Net power of the grid [kW]
W_{Goff} Energy delivered in the grid (Off-peak) [kWh]
W_{Gon} Energy delivered in the grid (On-peak) [kWh]
W_G Total energy delivered in the grid ($W_{Gon}+W_{Goff}$) [kWh]
W_{Loff} Energy supplied by the grid (Off-peak) [kWh]
W_{Lon} Energy supplied by the grid (On-peak) [kWh]

BEES is modeled by two equations for three working states. If the power of the battery P_B is less than zero, the battery is in charge process. This process we can describe with (1):

$$W_B(h) = \mathrm{W}_B(h-1)\left(1-\frac{\delta\Delta t}{24}\right)+\eta_{ch}\cdot P_B(h)\cdot\Delta t. \qquad (1)$$

In (1) the first summand models the self-discharge while the second describes the battery charging from the PV system. The charging process ends when the BESS reaches the maximum capacity. In the second case when the power of the battery is higher than zero, the battery is in discharging process. The discharging process ends when the BESS is on minimum stored energy. This state we can also describe with (1). Finally, the second equation models the battery at no operation. Equation (2) takes into account only self-discharge process in the battery:

$$W_B(h) = \mathrm{W}_B(h-1)\left(1-\frac{\delta\Delta t}{24}\right). \qquad (2)$$

C. Load

The load is considered known for the month under consideration. It is modeled by an array of the hourly values P_L. An example of the monthly load profile of an industrial customer from Banja Luka is shown in Fig. 2.

D. Grid

In this paper, we consider the grid as uninterrupted source of infinite power. The grid is used for supplying the base load. The grid is also used when the battery is empty and the PV system is not capable of producing enough energy. Electricity tariffs for industrial and commercial customers in many utilities worldwide include both the volumetric and peak demand component. Therefore, one of our primary targets is to reduce the grid peak demand and to decrease its negative effect on the customer energy bill. Tariff information used in this study corresponds to actual situation in the Republic of Srpska [7,8].

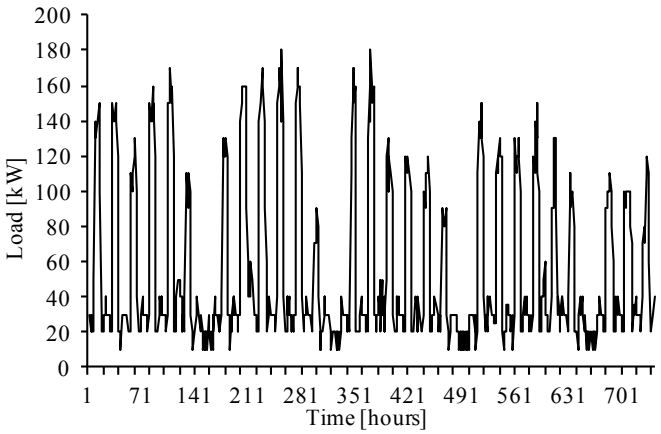


Figure 2. The monthly load profile

E. Control Strategy

We use an automatic controller to adjust the load flow in the customer AC bus for the minimization of the overall energy cost. If the net metering is used for calculating energy exchange between the customer and distribution network, the following cost function is employed:

$$C = (W_{Lon} - W_{Gon})\cdot C_{on} + (W_{Loff} - W_{Goff})\cdot C_{off} + P_{Gmax}\cdot C_{pd}. \qquad (3)$$

The photovoltaic sources, as the source of renewable energy, may be incentivized by a feed-in tariff for the energy supplied to a distribution network. In such a case, the customer cost function is written in the following manner:

$$C(P_{Gthr}) = W_{Lon}\cdot C_{on} + W_{Loff}\cdot C_{off} + P_{Gmax}\cdot C_{pd} - W_G\cdot C_f. \qquad (4)$$

Both for cases (3) and (4) the goal for controller is to minimize the cost function C. The input data for the controller are the production of PV system, load, stored energy in BESS and threshold power. The output data are charge or discharge rate of the BESS and power supplied by the grid. The proposed control algorithm is shown in Fig. 3. The controller first com-

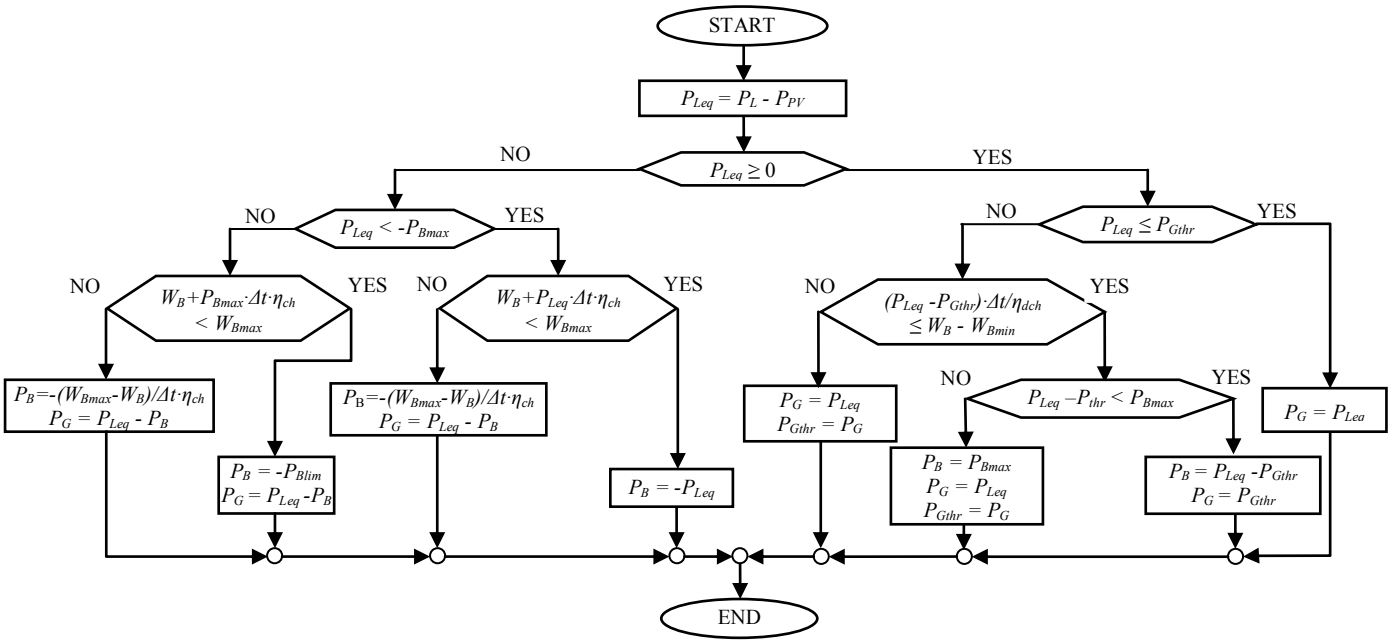


Figure 3. The control algorithm flow chart

pares the load and generation of the PV system. If the difference is negative and less than negative value of the maximum charge rate of BESS and there is enough empty capacity in BESS, the BEES is charged. In the second case, if there is not enough capacity in BEES, the BESS is charged up to its maximum capacity and the rest is delivered to the distribution network. In the third case, if the difference is negative and higher than negative value of the maximum charge rate of the BESS, the BESS is charged and distribution network is supplied. The rate of charging of the BESS and grid supply depend on the state of the BESS charge. If the difference between the load and generated energy in the PV system is positive and less than the user defined power threshold P_{Gthr}, the load is supplied from the grid. However, if difference is positive and higher than the limit of power supplied by the grid and there is enough energy in the BESS, the load is supplied by distribution network with power limit and the BESS is discharged. But if there isn't enough energy inside BESS, the load is supplied from the distribution network and additionally the power threshold must be appropriately increased.

Obviously, very important issue in this control strategy is to choose the appropriate threshold for the power drawn from the grid. In a deterministic case with all the inputs being certainly known, the optimal threshold may be found by a simulation based search process. The customer cost would then be minimized. On the other hand, if the inputs deviate from their expected values the cost effectiveness of the algorithm will be more or less reduced below the theoretical maximum.

III. SENSITIVITY ANALYSIS

In this section we examine how some important input variables influence the overall cost effectiveness of the control algorithm. We analyze the selection of the grid power threshold, the sequence of the sunny and overcast days as well as the changes in the customer load.

A. Test Case Description

Test customer is a medium industrial customer from Banja Luka. As an example, we observe the month of July. The load profile is given in Fig. 2. The production of the PV system, as calculated in PVWatts, is shown in Fig. 4. The other system parameters used in all the simulations are listed in Table I.

TABLE I. SYSTEM INFORMATION

Parameters	Value	Unit	Reference
P_{pv}	200	kW	
P_{Bmax}	100	kW	
$W_{ref,}\ W_{Bmax}$	*750*	kWh	
W_{Bmin}	*75*	kWh	
δ	0.140	% per day	[3]
η_{ch}	89	%	[6]
η_{dch}	89	%	[6]
η_{inv}	96	%	[6]
C_{off}	0.054	BAM/kWh	[7] (tariff group I)
C_{on}	0.108	BAM/kWh	[7] (tariff group I)
C_{pd}	15.867	BAM/kWh	[7] (tariff group I)
C_f	0.3178	BAM/kWh	[8]
T_{on}	7	h	[7] (tariff group I)
T_{off}	23	h	[7] (tariff group I)

B. Selection of the Grid Power Threshold

By changing the grid power threshold P_{Gthr}, the customer electricity bill is affected to a certain extent. Simulation results are shown in Fig. 5. The cost of energy are calculated by using expression (4). It can be noticed that the optimal threshold is 82 kW and the minimal energy bill for July is 2338 BAM. If the threshold is larger than 82 kW, the cost of energy increases, as we allow a greater peak demand drawn from the grid. For $P_{Gthr} > \max(P_{Leq})$ the energy costs don't any further depend on

the chosen threshold. In the second case, if the threshold is less than 82 kW, the cost of energy changes in dependent when the threshold is exceeded.

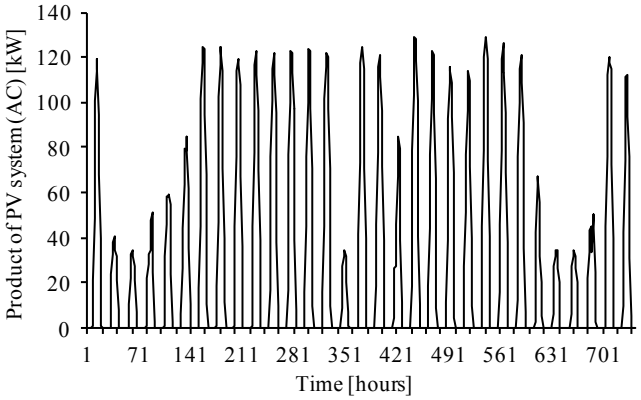


Figure 4. Production of the PV system

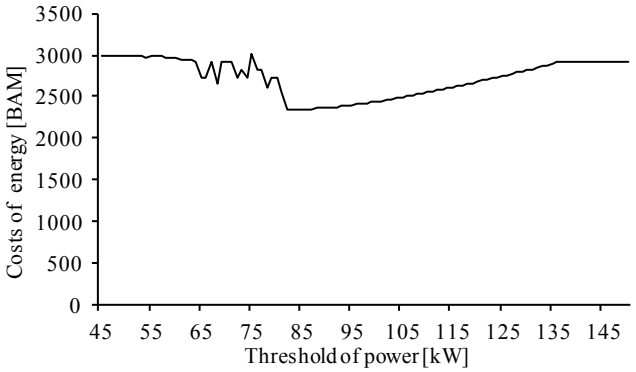


Figure 5. Cost of energy in function of power threshold

C. *Sequence of Sunny and Overcast Days*

The next experiment is a change of the sequence of sunny and overcast days, while the total energy produced by the PV system is kept constant. The impact on the grid power threshold and the cost of energy is analyzed in several cases. In the first case, the sunny and overcast days are uniformly distributed. In other cases we have several overcast days in a block. Table II presents a number of overcast days in each case, the optimal threshold power and the cost for this power. Fig. 6 presents the energy cost as the function of the chosen threshold for different cases. Fig. 7 presents threshold deviation from the one we have for the reference case.

TABLE II. NUMBER OF OVERCAST DAYS

Event	Number of overcast days in block	Optimal threshold [kW]	Cost for optimal threshold [BAM]
Case A1	1	71	2497
Case B1	3	77	2450
Case C1	4	80	2393
Reference case (R)	5	82	2338
Case D1	6	89	2447
Case E1	8	95	2519
Case F1	9	99	2539

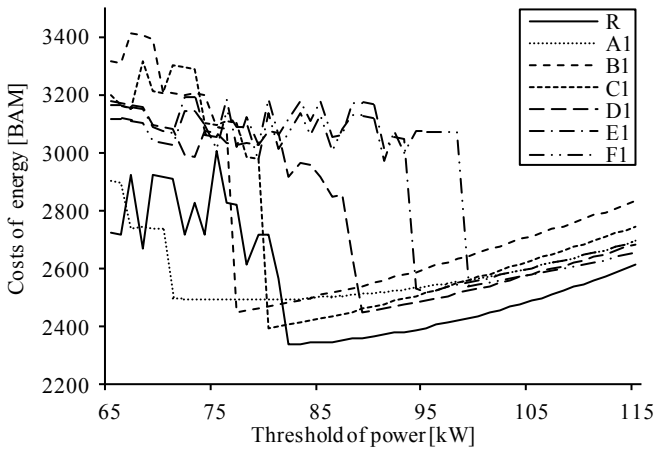


Figure 6. Cost of energy as a function of power threshold for several cases

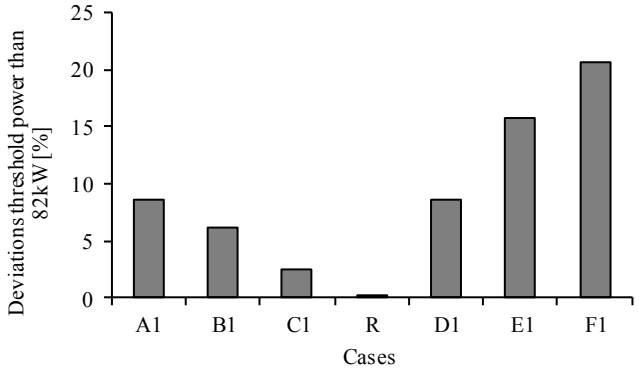


Figure 7. Deviation threshold from the reference case

Moving away from the reference state we have a higher deviation. If we mistakenly set the controller for the reference state but the production of the PV system belongs to one of the other cases, we will obtain deviation in costs. These deviations are presented in Fig. 8. It can be noticed that a block of overcast days wider than in the reference case leads to errors in the cost of more than 20 %, due to an increase in the power threshold. On the opposite, the number of overcast days less than it is in the reference case leads to an error less than 2%, because the customer can deliver surplus of their energy in the distribution network.

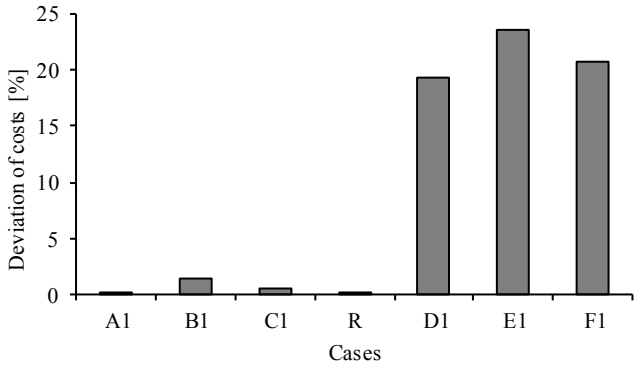


Figure 8. Deviation of energy cost from the minimum value

D. Changes in the Customer Load

Finally, the impact of load change on the cost of energy is analyzed. The load is linearly scaled by multiplying the load profile by a constant *k*. Table III presents the values of coefficient *k,* optimal threshold power and the cost related to the optimal threshold.

TABLE III. LOAD SCALING

Event	*k*	Optimal threshold [kW]	Cost for optimal threshold [BAM]
Case A2	0.85	60	1216
Case B2	0.90	68	1601
Case C2	0.95	75	1978
Reference case (R)	1.00	82	2338
Case D2	1.05	90	2689
Case E2	1.10	97	3032
Case F2	1.15	104	3372

Fig. 9 presents the cost of energy as the function of the coefficient *k*. In this figure we can notice that the optimal threshold is growing when the coefficient *k* tends towards 1.15.

If the controller is set for the reference case and the load belongs to one of cases defined here, we have deviations in the cost. These deviations are presented in Fig. 10. In this figure we can see that the coefficient *k* less than one causes the error less than 10%. However, for the coefficient *k* higher than one error is higher than 20%, because it leads to an increase in the grid power threshold.

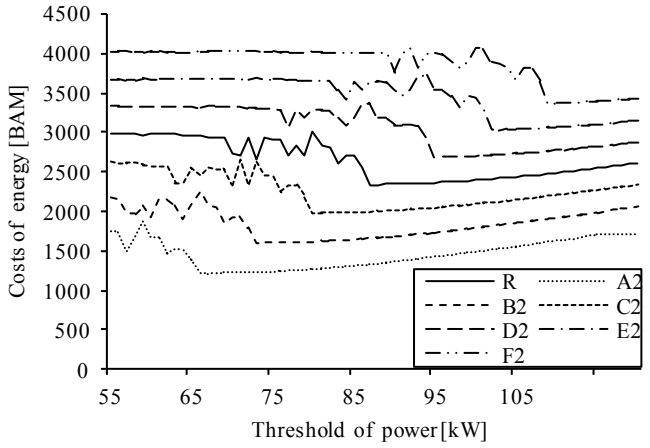


Figure 9. Cost of energy as the function of grid power threshold for considered simulation cases

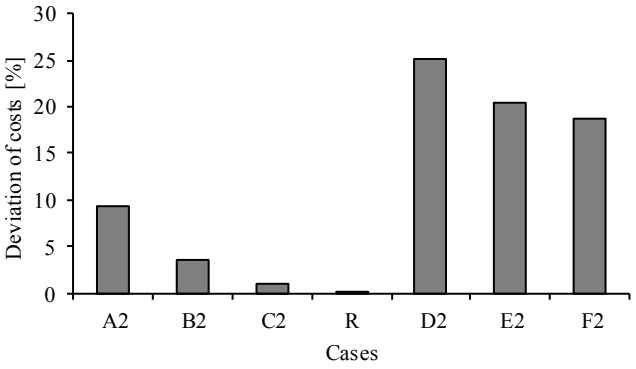


Figure 10. Deviation of energy cost from the minimum value

IV. CONCLUSIONS

The customer cost of energy is significantly sensitive on the change of the load and the number of consecutive overcast days. Our simulations show that the underestimation of number of overcast days makes the proposed control strategy less cost effective. Similarly, the control strategy performance is getting worse if the customer load is underestimated. Generally spoken, a deterministic approach is applicable to the systems where the input variables can be predicted with a solid certainty. Otherwise, if the inputs exhibit remarkable unpredictable deviations, more complex stochastic methods should be employed.

REFERENCES

[1] J.M.Roney, " World Power topped 100.000 Megawatts in 2012," Earth Policy Institute, July 2013. Available: http://www.earrh-policy.org/indicators/C47/solar_pover_2013 (accessed 20.09.2014).

[2] J.M.Roney, "China leads World to Solar Record in 2013, " Earth Policy Institute, July 2013. Available: http://www.earrh-policy.org /indicators/C47 (accessed 20.09.2014).

[3] M. Bortolini, M Gamberi, and A. Graziani, "Technical and economic design of photovoltaic and battery energy storage system," Energy Conversion ans Management, vol. 86 pp. 81-92, October 2014.

[4] J. Li and M.A.Danzer, "Optimal charge control strategies for stationary photovoltaic battery systems," Journal of Power Sources, vol. 258 pp. 365-373, July 2014.

[5] A.P.Dobos, "PVWatts version 5 manual," National Renewable Energy Laboratory, September 2014. Available: http://pvwatts.nrel.gov /downloads/pvwattsv5.pdf

[6] C.J. Rydh, B.A. Sanden, "Energy analysis of batteries in photovoltaic systems. Part I Performance and energy requirements," Energy Conversion and Management, vol. 46 pp. 1957-1979, July 2005.

[7] REERS [Regulatory Commission for Energy of the Republic of Srpska], "Tariffs" [Online]. Available: http://www.reers.ba/en/node/296.

[8] REERS [Regulatory Commission for Energy of the Republic of Srpska], "Tariffs" [Online]. Available: http://www.reers.ba/en/node/1488.

Ultracapacitors as auxiliary energy source in electric vehicles

Ivan M. Todorović, Petar Gartner, Vladimir Katić, Stevan Grabić
Faculty of Technical Science
Novi Sad, Serbia

Abstract— **Recent and promising anticipated development in energy storage technologies demands adequate energy flow control strategies that will actuate opportunities spawned by this development. In this paper, energy management strategy that relies on fuzzy logic theory is proposed, i.e. controllers that govern functioning of the converters are designed using human reasoning and interrelated fuzzy logic rather than conventional PID controllers. It will be shown how the proposed energy management allows ultracapacitors to be used as an energy source that not only eliminates shortcomings of other sources but can also be used as a tool for optimization of system functioning as a whole.**

Keywords- energy storage; energy management; ultracapacitors

I. INTRODUCTION

It is evident that many factors influence the pace of development of electric vehicles (EV) and hybrid electric vehicles (HEV) industry but the fact that, for the first time, electric cars are the best-selling new cars, in Norway for instance, shows that this industry is on the rise and that appropriate combination of ecological awareness and financial subventions will cause the market to eventually shift towards these forms of transportation rather than transportation powered by fossil fuels. With an increased demand for vehicles that are partially (HEV) or completely (EV) powered by electrical energy, fuel cells, ultracapacitors (UCs) and battery-based electrical sources will find ever increasing utilization. It is important to point out that problems and opportunities of energy storage systems and energy management in EV are very similar to the problems of energy storage found in high power systems such as wind and solar farms. That makes this topic even more important for researchers[1].

In order to use these technologies, control strategies are being developed. All of them take into account the dynamics of chemical processes that occurs in them during energy exchange with, in case of (H)EV, the motor drive. Generally, traction depends on the energy supplied by chemical sources of energy-batteries, UCs, fuel cells and mechanical sources of energy-flywheels, but how that energy is exchanged is of crucial importance for the robustness, longevity and performance of the system. In most EV applications, the battery is used as the primary source of energy, while UCs are used to complement the battery's energy delivery characteristics. In HEV applications, fuel cell is used as the primary source of energy, while batteries and/or UCs are used as secondary energy sources[2][3].

Depending on the technology used in the production of different batteries, UCs and fuel cells, their characteristics differ, but underlying similarities and general attributes can be observed as well. For the sake of clarity of this paper's goals, general properties of the aforementioned energy sources are presented below.

Contrary to other papers that tackle the problems of energy flow in (H)EV, in which each energy source is evaluated as a distinctive source with separate functions[2][3], and little or no direct association with other sources in the system, in this paper UCs are seen exclusively as the *slave* energy source to the other energy sources that are seen as the *master* energy sources. It can be argued that the differences in these two standpoints are subtle, but they do result in a different control strategy.

The control strategy presented is constraited to the case of EV with batteries as the master energy source and ultracapacitors as the slave energy source. The objective is to examine how the system would behave if UCs are used as the source that observes the behavior of the master source and reacts on demand or when necessary. The results obtained from the simulations and presented in this paper and some further experimentation could show us that this approach facilitates the optimization of master energy sources and allows UCs to successfully complement master energy delivery characteristics.

The paper is divided into six chapters. After the introduction, a short overview of electrical energy sources is presented, but only to the extent that it is easier to understand why and how energy management is designed. The details that concern traction are omitted because they are of little importance to this work. Chapter III tackles the choice of converters and explains how batteries and UCs are modeled. In chapter IV, energy management and control structure are analyzed. Chapter V deals with simulation results, while chapter VI contains conclusions and future work considerations.

II. ENERGY SOURCES IN ELECTRIC VEHICLES

This chapter should provide a good starting point for the understanding of the ideas presented in the paper. A reader that is not familiar with these technologies would otherwise find it hard to follow and understand. Fuel cells are presented for educational purposes only, but the system that uses fuel cells instead of batteries could be easily derived from the system used here and thus they are not completely omitted from the paper.

A. *Fuel cells*

The fuel cell is a galvanic cell in which the chemical energy of fuel, hydrogen or some hydrocarbon like natural gas or methanol, is converted into electrical energy by means of electrochemical processes with air/oxygen. Fuel cell are usually coupled with a tank that stores fuel under high pressure. Essentially, fuel cells act as generators of the constant electromotive force and, as long as a constant flow of fuel and air is supplied, fuels cells should produce energy in consistent manner.

Unlike batteries, the construction of a fuel cell makes it suitable only for electric propulsion, i.e. fuel cells cannot be used to receive electric energy during regenerative braking. Fuel cell powered vehicles have the advantages of a longer driving range without a long battery charging time. In contrast to internal combustion engine powered vehicles, it has the advantage of high energy efficiency and much lower emissions. Currently, a major drawback of fuel cells is that they tend to be very expensive, as they are only starting to be commercially available. Next, the problem with fuel storage is that engineers are still trying to tackle it in economically and practically acceptable way. Also, proper infrastructure for this type of vehicles is nonexistent, which additionally impedes their usage[4][5].

B. *Batteries*

Batteries are electrochemical devices that convert chemical energy into electrical while discharging, and convert electrical energy into chemical while charging.

There are a number of parameters used to characterize the battery, namely specific energy, specific power, safety, performance, life span, and cost. All of these parameters are properties taken into account when designing the battery energy storage unit for an EV or an HEV. In general, batteries have good energy density, which means that they can be used for energy delivery during relatively long periods of time during which the power demand is slowly changing. On the other hand, batteries mostly have poor power density, which means that they are not fit to be used during short power peaks.

Several types of batteries can be found in the market today, mainly lead-acid, nickel based batteries (NiFe, NiCd, NiMH), and lithium based batteries (lithium polymer Li-P, Li-Ion). For high power applications, two main battery technologies of interest are NiMH and Li-Ion.

An important aspect of handling battery technology is battery lifetime and ageing process. There are a number of different factors that influence its lifetime:

- high temperature. High temperature has a negative influence on the battery's state of health.
- battery's lifetime is closely related to the number of charge/discharge cycles the battery has been subjected to. Also, battery life is shortened by frequent surges of energy.
- Battery should not be completely discharged nor overcharged. The rule of thumb is that battery should always work in the state-of-charge between 20 and 80% of its full capacity, although a single overcharge or over-discharge would not harm the battery too much.

Generally, battery life span is thus extended by slow, controlled flow of energy without fluctuations and within over-discharge/overcharge boundaries[4].

C. *Ultracapacitors*

Compared to batteries, UCs have a good power density, but poor energy density. This makes them unsuitable to use alone as an energy storage unit. However, their characteristics make them an appropriate auxiliary power source. Particularly, when used as energy storage units along with batteries, UCs can be used to smooth out power to the batteries and to relieve batteries from stress. Additionally, UCs are long lasting, and can go through a very large number of charge/discharge cycles (over one million). UCs are thus more resilient to fluctuations of energy often seen in cars (acceleration and deceleration)[6][7].

III. BATTERY AND UC MODELING AND CHOICE OF CONVERTERS

A. *Battery model*

A battery is modeled as a controlled voltage source in series with an internal resistance (Figure 1.). The controlled voltage source is calculated as follows:

$$E = E_0 - K\frac{Q}{Q-it} + Ae^{-B\cdot it} \quad (1)$$

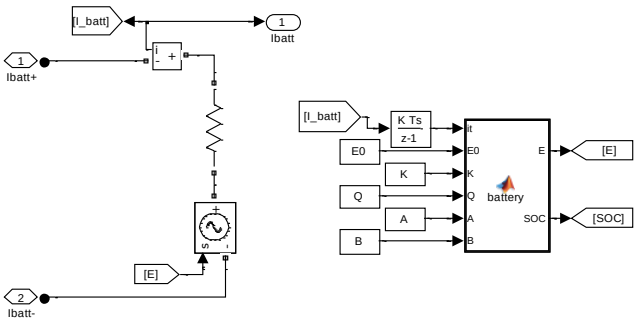


Figure 1. Simulink implementation of battery model.

This equation allows modeling of battery voltage as a function of state of discharge *it.* Quantities used in equation (1) are:

- $it = \int_0^t i_{bat}dt$ battery state of discharge [Ah]
- $A = E_{full} - E_{exp}$ voltage drop during exponential zone [V]
- $K = \frac{\left(E_{full} - E_{nom} + A\left(e^{-B\cdot Q_{nom}} - 1\right)\right)\cdot(Q - Q_{nom})}{Q_{nom}}$ polarization voltage [V]
- $B = \frac{3}{Q_{exp}}$ charge at the end of exponential zone [(Ah)$^{-1}$]
- $E_0 = E_{full} + K + Ri - A$ battery constant voltage [V]

Internal resistance R is calculated using :

$$R = V_n\frac{1-\eta}{0.2\cdot Q_n} \quad (2)$$

where η is battery efficiency, Q is maximum battery capacity and Q_n is rated capacity. To extract battery model parameters, three points on the battery discharge curve are used (Figure 2.):

- fully charged voltage
- end of exponential zone (voltage and charge)
- end of nominal zone (voltage and charge)

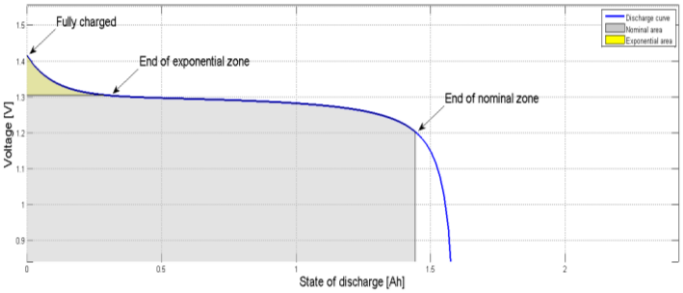


Figure 2. Typical battery discharge curve.

This way of battery modeling gives the same charge and discharge curves. The model used in this paper gives very good results when compared to actual batteries and its ease of implementation makes it suitable for usage. Interested reader can find more detailed explanation of this model in [8]. Figure 3. shows the response obtained by simulation.

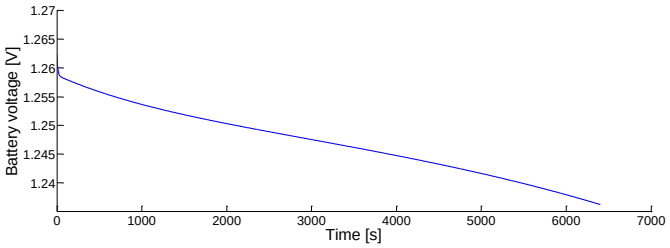


Figure 3. Battery voltage for 0.3 amps discharge current.

B. UC model

Over the last two decades many models of ultracapacitors have been developed. The differences between them are the result of different modeling objectives. Many factors influence the working state of the UC and its lifetime, so different models are trying to capture different characteristics of the UC in order to model its behavior in different circumstances and then draw conclusions about how the UC should be used. This paper uses the model proposed by Zubieta and Bonert[7]. It represents a model that can be easily implemented in some software packages, like Matlab, but still captures basic and most important physical properties of the UC.

Figure 4. shows the implemented Simulink model. This model consists of four parallel branches. The first three branches consist of a capacitor and a resistor, where the first branch has an additional voltage dependent capacitor.

Each branch depicts a different phase of the charging/discharging process – the first has the time constant in order of seconds (thus named immediate branch), the second has the time constant in order of minutes (delayed branch) and the third has the time constant longer than ten minutes (long term branch). The forth branch consists of the big resistor that should model self-discharge process.

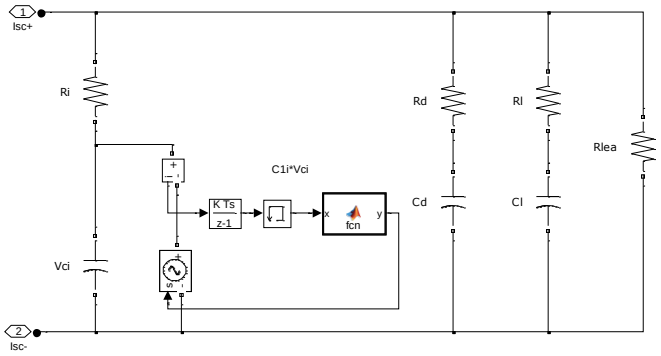


Figure 4. Model of UC.

Figure 5. shows the typical charging curve of the UC obtained by simulation.

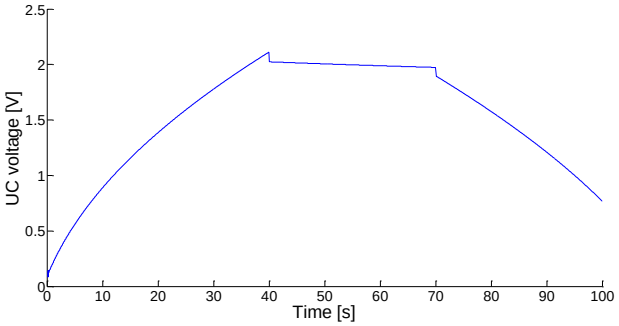


Figure 5. Characteristic UC terminal voltage.

C. Choice of converters

Some papers propose topology that assumes that the battery is connected directly to dc-link, but here both the battery and UC are connected to dc-link via converters, as Figure 6. suggests. The reason for this is better controllability of the battery charging/discharging process. Here the boost converter is used to step up the operating voltage of battery array which is around 200 volts (in Figure 3. only one battery was discharged) to 560 volts, which is standard dc-link voltage for three phase inverters and drives that use 400 volts line voltage. It is thus assumed that the battery is not being used for recuperation.

UCs array is connected to dc-link over the half-bridge converter. This converter is used because it is assumed that both directions of energy flow are permitted, so that UCs can absorb the excess of energy from dc-link and inject energy in case of shortages.

Synchronous drive is used for traction, powered by the three phase inverter that is controlled by standard vector control and space vector modulation. Further details about traction, inverter and control are omitted because they are not, as already explained, of great significance to this topic.

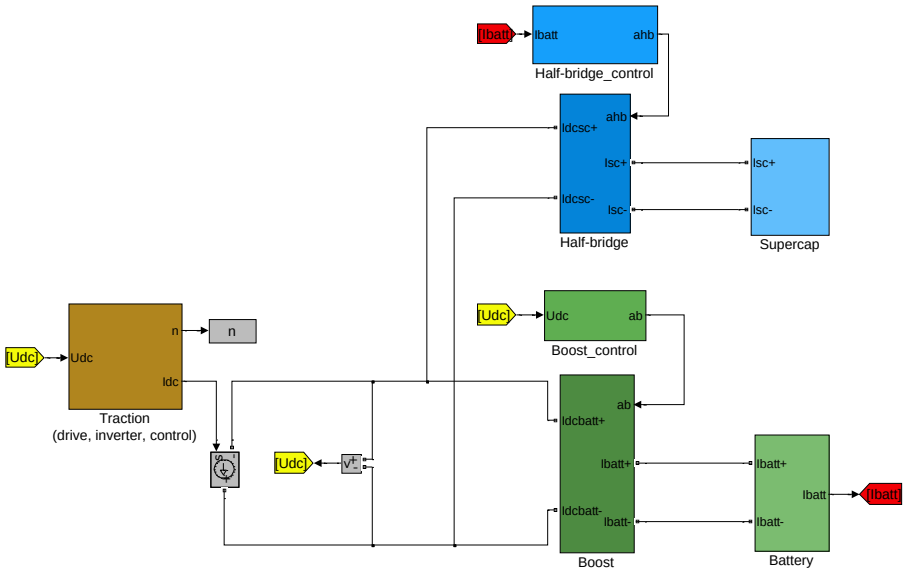


Figure 6. Simulink model of EV.

IV. ENERGY MANAGEMENT

From Figure 6. one can see how Simulink model is organized. Battery and UC banks are on the right, converters and control structures in the middle and traction on the left. It can be observed that there is a current source used in dc-link. This is so because in this simulation a 'bridge' between control of power drive, drive itself, inverter on the one side and the rest of simulation on the other is needed since the former is implemented using Simulink toolbox, while the latter (converters and energy sources) is simulated using SimPowerSystem toolbox. Voltmeter measures dc-link voltage and sends this information to the control of the inverter. On the other hand, the current that should be injected/taken from dc-link is calculated using the next expression:

$$I_{dc} = \sum_{x=1}^{q} S_x * i_x \qquad (3)$$

where q is the number of phases, S_x is the switching function for top transistor of x-th leg of inverter (where $S_x \in \{0,1\}$) and i_x is the current flowing through x-th phase.

In this paper, UCs are seen as the slave energy source, in the sense that they are slave to the master energy source, which is batteries. Their purpose is to ensure the best possible working state for the batteries which, as explained above, is the state of perpetual but constant flow of energy with as little fluctuations as possible. Thus UCs should 'react' every time the fluctuation in energy flow occurs. These fluctuations occur mostly during acceleration and deceleration. In other words, UC should flatten energy demand curve, and by doing so not only improve the performance of the system, but also the expected lifespan of the batteries.

One way of controlling energy flow in the aforementioned manner is by observing battery current and judging on the *change* in the battery current control half-bridge converter in such a way that injected/absorbed energy from the UC tends to lower this change. It can be seen that one of the outputs from the battery block in Figure 6. is the measured battery current, which is then fed to the half-bridge control block. The content of this block is shown in Figure 7.

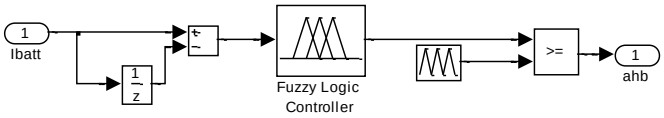


Figure 7. Control for half-bridge converter.

In case of this somewhat unusual control strategy, where the control of one source is dependent on the dynamics in another source, a good approach for designing control strategy is to use human reasoning. For instance, it is clear that when the motor is accelerating it will need a bigger current than when the drive is spinning at a constant speed. Thus, there will be a rise in the current drawn from dc-link, and this sudden rise reflects on the rise in the battery current. To prevent this from happening, UCs should inject energy in dc-link as soon as the change is detected. To do so, duty cycle of half-bridge should change. If the motor is accelerating aggressively, duty cycle should be dramatically lowered or enlarged depending on the regime. It is quite hard to quantify this change and some conventional control strategies would not be as practical as the control based on fuzzy logic.

Fuzzy logic controllers can already be found in numerous applications[9][10][11][12]. Their robustness, simplicity and ability to easily tackle nonlinear and very complex control problems make them very popular among researchers. Here they are used both for the control of the boost converter and control of the half-bridge converter, but only the fuzzy controller for half-bridge is explained in some detail.

The range of the change of the battery current without the usage of UCs was experimentally determined to be between 0.06 amperes and -0.06 amperes and thus this range is used in fuzzification block for the input variable, as it was reasonable

to expect that the range of change should stay within these boundaries with UCs used. It should be noted that difference between current and previous value of battery current can be subjected to high influence of measurement noise but this problem wasn't addressed here, partially because it didn't cause any problems during simulation. The range of output variable, duty cycle, was set to be from zero to one. Figure 8. shows the equidistant distribution of membership functions for input and output variables.

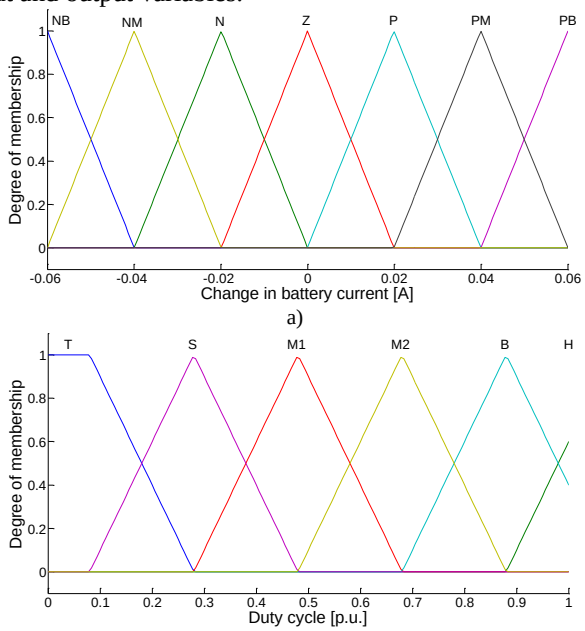


Figure 8. Membership functions: a) change in battery current, b) duty cycle.

The distribution of input variable was fixed at the beginning of the tuning process, while the distribution of membership functions for duty cycle is the result of the tuning process. Number of membership function was chosen intuitively. Rule base that consist of IF-THEN rules was generated by reasoning how duty cycle should change with the change in the battery current. For example, if the change in the battery current is negative big (NB), duty cycle for half-bridge converter should be huge (H), because this duty cycle would stop this change in current sufficiently quickly. The following set of rules was obtained:

- 1. If (dIbatt is NB) then (a is H) (1)
- 2. If (dIbatt is NM) then (a is B) (1)
- 3. If (dIbatt is N) then (a is M2) (1)
- 4. If (dIbatt is Z) then (a is M1) (1)
- 5. If (dIbatt is P) then (a is S) (1)
- 6. If (dIbatt is PM) then (a is S) (1)
- 7. If (dIbatt is PB) then (a is T) (1)

For deffuzification centroid method is used, while for 'and' method min operation is used and for 'or' method max operation is used. In the next section simulation results will demonstrate how the system behaves with and without UCs.

V. SIMULATION RESULTS

Figure 9. a) shows dc-link voltage without UCs being used. Fluctuations are rather big and are not caused only by the oscillating load but also by the poorly tuned boost controller.

Fuzzy controller for boost converter was purposefully left poorly tuned because in this way time consuming tuning was avoided, but also half-bridge control could have been put to a test to maybe even unrealistic circumstances. This was reasonable because from the battery's standpoint there was no difference in fluctuations caused by the variable load and those caused by the poorly tuned controller. Of course, after the tests were finished, boost converter was properly tuned, but tests with this configuration are omitted. Figure 9. b) shows dc-link voltage with the same control for boost converter as used in picture under a), but now with the UCs used and after the tuning of half-bridge controller.

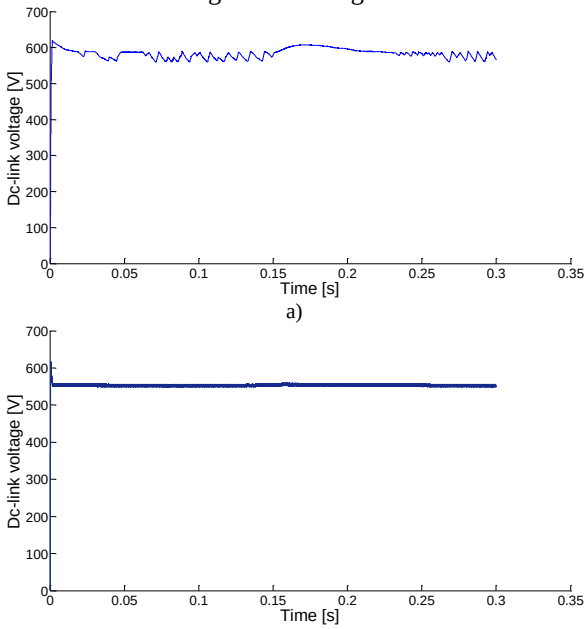
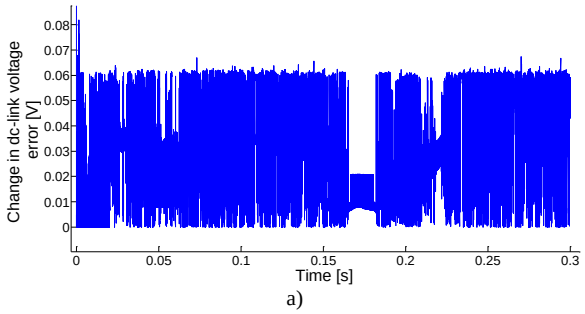
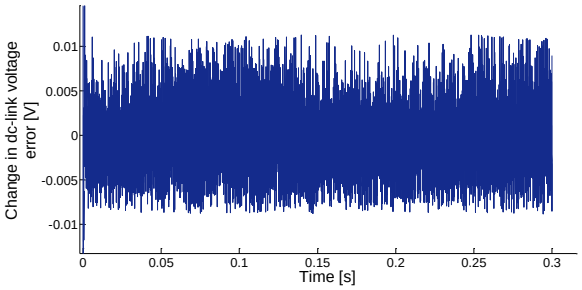


Figure 9. Dc-link voltage without a) and with b) UCs.

It is obvious that voltage is now much steadier and this certainly reflects on battery stress. Figure 10. shows the change in dc-link voltage error. Again, significant improvement is visible. Without the UCs, the change in error oscillates within the range of 0.1 volts, while with UCs within the range of 0.01 volts.

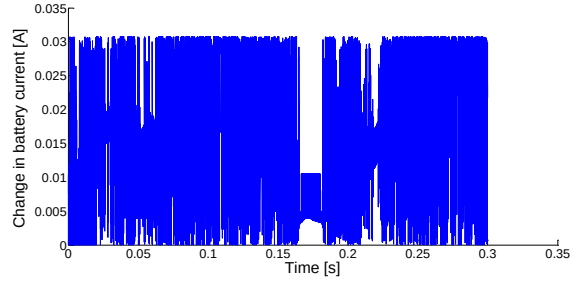




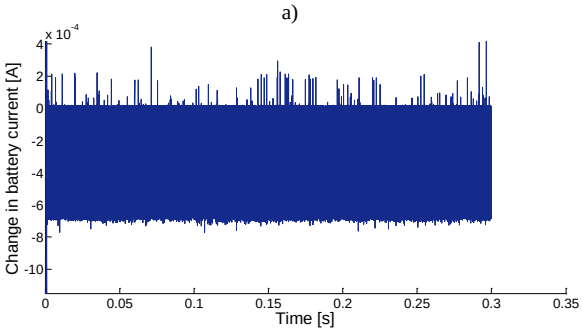
b)

Figure 10. Change in dc-link voltage error without a) and with b) the UCs.

Figure 11. depicts the change in the battery current. In the case without the UCs, it oscillates within the range of 0.03 amperes, while in the case with the UCs used the change oscillates within the range of $6*10^{-4}$ amperes and thus UCs have successfully helped in optimizing battery working conditions and thus improved their life expectancy.



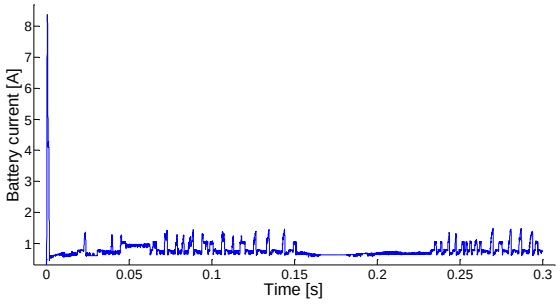
a)



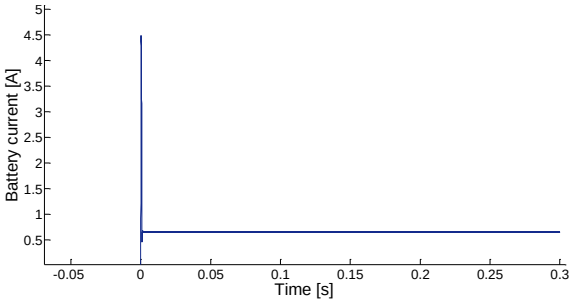
b)

Figure 11. Change is battery current without a) and with b) the UCs.

Figure 12. shows the battery current. It can be seen that the peak current is lowered. In cases of sudden change in the battery current, fuzzy controller of half-bridge is tuned to dramatically change duty cycle and thus inject energy in dc-link and stop the change. In this way surges of energy caused by braking and acceleration are mitigated.



a)



b)

Figure 12. Battery current without a) and with b) the UCs.

Figure 13. show a peak in the UCs' current of over 40 amperes which suggests that the most of the energy needed is supplied from UC bank rather than from battery. Mean value then drops but not to zero which suggests that additional tuning should be done in order to make UCs' current drop to zero so that they save energy for acceleration.

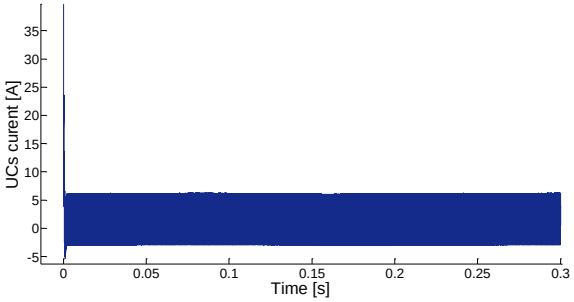


Figure 13. The UCs current.

Figure 14. shows the motor speed curve, which suggests that after acceleration the speed reference is changed (at 0.15 seconds), so that breaking could be simulated.

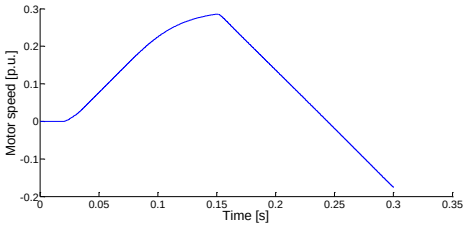


Figure 14. Motor speed.

VI. CONCLUSIONS AND FUTURE WORK CONSIDERATIONS

Based on the simulation results shown in the previous chapter it can be concluded that UCs improve working conditions of the batteries and thus, by judging what causes the batteries to deteriorate, increase their life expectancy and can be used to optimize system performance by using simple and straightforward control strategy. By observing battery voltage, it has been shown that fuzzy control can govern UCs energy flow in the required manner. In some future work, though, these improvements could be quantized in more detail. Next, a bigger motor drive could be used, so that more serious braking and acceleration conditions could be simulated. Also, the research of how easily this control strategy could be used in case of HEV or EV with fuel cells can be pursued.

VII. REFERENCES

[1] Phatiphat Thounthong, Stephane Rael "The benefits of hybridization", IEEE Ind. Electron. Mag., vol. 3,issue 3, pp. 25-37, Sep. 2009.

[2] M. Ortúzar, J. Moreno, J. Dixon, "Ultracapacitor-Based Auxiliary Energy System for an Electric Vehicle: Implementation and Evaluation", IEEE Trans. on Ind. Electron., vol. 54, no. 4, pp. 2147-2156, Aug. 2007.

[3] S. Lukic, S. Wirasingha, F. Rodriguez, J. Cao, A. Emadi "Power Management of an Ultracapacitor/Battery Hybrid Energy Storage System in a HEV", IEEE Vehicle Power and Propulsion Conference, pp. 1-6, VPPC '06, 2006..

[4] Mehrdad Eshani, Yimin Gao, Sebastien E. Gay, Ali Emadi "Modern electric, hybrid electric and fuel cell vehicles", CRC Press LCC, New York, 2005.

[5] Srdjan M. Lukic, Jian Cao, Ramesh C. Bansal, Fernandor Rodriguez, Ali Emadi "Energy storage systems for automotive applications", IEEE Trans. on Ind. Electron., vol 55. Issue 6., pp. 2258-2267, June 2008.

[6] Dimitri Torregrossa, Maryam Bahramipanah, Emil Namor, Rachid Cherkaoui, Mario Paolone, "Improvement of dynamic modeling of supercapacitor by residual charge effect estimation", IEEE Trans. on Ind. Electron., vol. 61, no. 3, pp. 1345-1354, March 2014.

[7] L. Zubieta, R. Bonert, "Characterization of Double-Layer Capacitors (DLCs) for Power Electronics Applications", IEEE Trans. on Ind. App., vol. 36, issue 1., pp. 199-205, Jan.2000.

[8] Olivier Tremblay, Louis-A. Dessaint, Abdel-Illah Dekkiche "A generic battery model for the dynamic simulation of hybrid electric vehicles", IEEE Vehicle Power and Propulsion Conference, VPPC'07, pp. 284-289, 2007.

[9] E. Adzic, Z. Ivanovic, M. Adzic, V. Katic, "Maximum Power Search in Wind Turbine Based on Fuzzy Logic Control", Acta Polytechnica Hungarica, vol. 6, no. 1, pp. 131-149, 2009.

[10] T. Kottas, Y. Boutalis, A. Karlis, "New Maximum Power Point Tracker for PV Arrays Using Fuzzy Controller in Close Cooperation With Fuzzy Cognitive Networks", IEEE Trans. on Energy Conv., vol. 21, no 3, pp. 793-803, Sept. 2006.

[11] Jiangtao Cao, Ping Li, Honghai Liu "An Interval Fuzzy Controller for Vehicle Active Suspension System", IEEE Trans. on Intelligent Transp. Systems, vol.11, no. 4, pp. 885-895, Dec. 2010.

[12] X. Li, "Fuzzy adaptive Kalman filter for wind power output smoothing with battery energy storage system", IET Renew. Power Gener., 2012, Vol. 6, Iss. 5, pp. 340–347.

Reducing the Active Power Losses in Transmission Network by Using Phase Shifting Transformer

Bojan Erceg, Petar Matić, Čedomir Zeljković
Faculty of Electrical Engineering
University of Banja Luka
Banja Luka, Republic of Srpska, BiH
bojan.erceg@etfbl.net

Abstract—**This paper analyses the possibility of decreasing the active power losses in 110 kV transmission network in BiH/RS by using phase shifting transformer. The network before and after connection of phase shifting transformer on the line with the largest power flow is simulated by a commercially available software package. The optimal phase shift required to minimize the overall losses in the transmission network is determined by a search through simulation results.**

Keywords - phase shifting transformer, power flows, power losses.

I. INTRODUCTION

The main role of Phase Shifting Transformer (PST) is to control active power flows in a steady state in transmission networks with many line loops and parallel lines [1].

The PSTs are devices of a great significance in current tendency for increasing transmitted power through existing lines and improving the efficiency of the network. Modern transmission networks are usually loaded close to their transmission limits. The PST is proposed as one of the most cost-effective solution that efficiently controls the flows of active power through network, and therefore optimizes network operation, improves system stability and thus reduces the active power losses.

Apart from avoiding overloading and instability problems in transmission networks, PSTs can help to decrease operational costs by performing optimized delivery of energy to consumers in post-failure state (state after damage). Furthermore, by using PSTs, investments in electrical power system (EPS) can be postponed by using the existing transmission lines operating on the limits of their thermal capabilities without overloading.

Phase shifting transformers are used for [2]:

- controlling active power flows in parallel lines,
- increasing of transmission capacities without violation of *n*-1 safety criterion,
- increassing the system reliability,
- decreasing the post-failure line overloads,
- eliminating unwanted power flows,
- removing bottlenecks in the network caused by concentrated power injection.

This paper describes the importance of using PSTs in the regulation of active power flows through lines of 110kV network in the part of electric power system of Bosnia and Herzegovina, with the aim of decreasing overall active power losses.

The paper is organized in two parts. The first part gives the basic principles of PSTs, while in the second part a computer model of 110kV network in program package PowerWorld Simulator (PWS) [3] is developed, in order to calculate power flows and active power losses in an EPS for the case with and without using the PST. Results of the calculation are given at the end of the paper.

II. PHASE SHIFTING TRANSFORMER

Due to mainly inductive character of transmission lines of the EPS, the active power flow between the source and consumer is defined by phase difference between voltages on the input and output of the line. In order to control active power flows, the phase shifting transformer can be used. The voltage phase shift is based on proper connection of the primary and secondary phase windings of the PST.

The active power P that is transmitted by transmission line from the sending node k to the receiving node m, (Figure 1a), is calculated by [4]:

$$P = \frac{U_k \cdot U_m}{X_L} \cdot sin(\theta_k - \theta_m). \quad (1)$$

From (1) it can be seen that active power P is proportional to the modules of voltage at the beginning and the end of transmission line, the sine of the phase angles difference between the beginning and the end of the line, and also inversely proportional to the reactance of the line X_L.

Relation (1) gives the basics for possible techniques of the active power flow control through a network branch. As voltage modules do not change significantly in normal operation (values are defined by standards), and reactance of transmission line X_L is constant in normal operation modes, the difference of phase voltage angles θ_k - θ_m remains a possibility for controlling the active power flows.

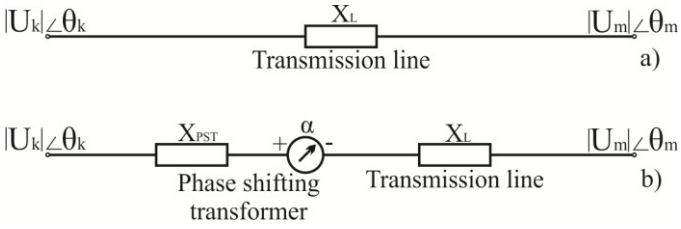


Figure 1. Transmission line with and without PST [4]

If the PST is placed between the nodes k and m on a transmission line (Figure 1b), the equation (1) is modified to:

$$P = \frac{U_k \cdot U_m}{X_L + X_{PST}} \cdot sin(\theta_k - \theta_m + \alpha), \qquad (2)$$

where X_{PST} is the reactance of PST, and α is the phase angle of PST.

Figure 2 gives a comparison of the transmission of active power through the line with and without PST. The line parameters are $\underline{U_k} = 1.050\angle 0°$ [p.u.], $\underline{U_m} = 0.997\angle 0.475°$ [p.u.], and $X_L = 0.205$[p.u.], while the parameters of PST are $X_{RFT} = 0.125$ [p.u.] and $\alpha = 20°$.

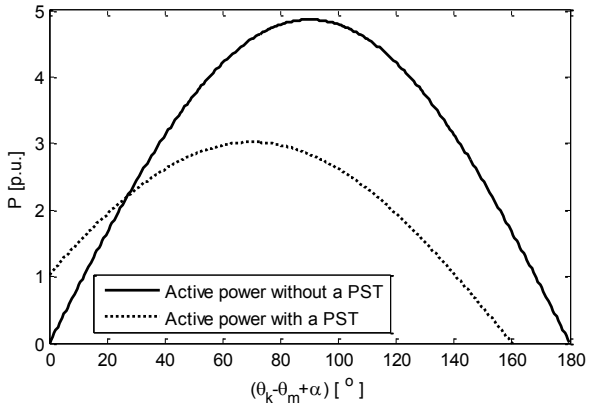


Figure 2. Active power flows without and with a PST

From the Figure 2 it can be seen that by using PST the transmitted active power can be increased if the angle $\theta_k - \theta_m + \alpha$ ranges between 0° and 26.93°.

A. Principles of Operation of Phase Shifting Transformer

The general principle of operation of PST is based on the connection of a part of one phase winding to the winding of another phase. In order to obtain a quadrature voltage ΔU for the phase shift regulation, the simplest solution is to use delta-connected winding on the secondary side [5]. Figure 3 shows connection of the secondary phase windings of phase L_2 and L_3. The secondary windings are split into two halves and connected in series with phase winding L_1. The regulating winding is connected by using on-load tap changer which provides regulation voltage ΔU and the phase-shift angle [5].

The phase diagram (Figure 4) shows the change of phase angle between input and output voltages for no load conditions, i.e., without considering the voltage drops in the transformer. It also should be noted that the currents in the two halves of the series winding I_{L1} and I_{L2} are not in phase. This is different from normal power transformers and has consequences with respect to the internal flux distribution [5].

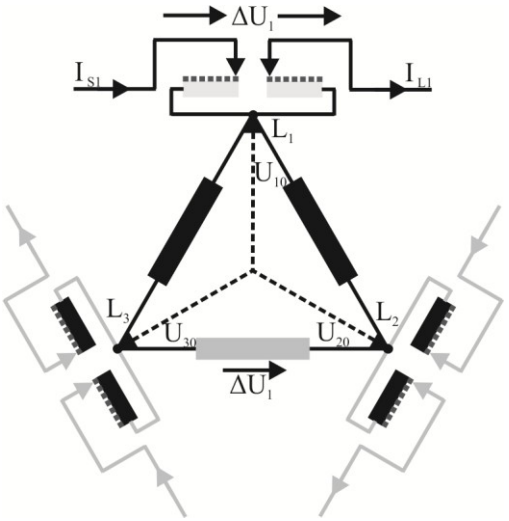


Figure 3. PST in delta-connection [5]

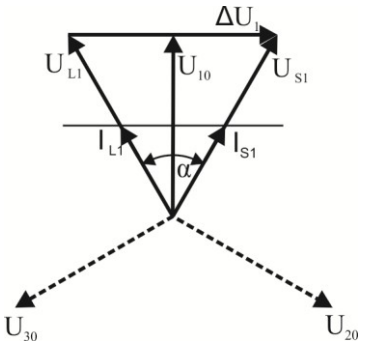


Figure 4. Phase diagram [5]

III. MINIMIZATION OF ACTIVE POWER LOSS BY APPLICATION OF PHASE SHIFTING TRANSFORMERS

In order to minimize the losses of active power in 110 kV transmission network of Bosnia and Herzegovina (Figure 5), a computer simulation of power flows before and after inserting PST is made in PWS.

The considered 110kV network is supplied from substation Banja Luka 6 and hydropower plants HPP Bocac and HPP Jajce, and has both single and dual lines.

The powerful 400kV network that supplies the substation Banja Luka 6 is modeled as a slack generator. Dual lines of 110kV network are modeled by the single line equivalents. The line parameters are given in the Table I.

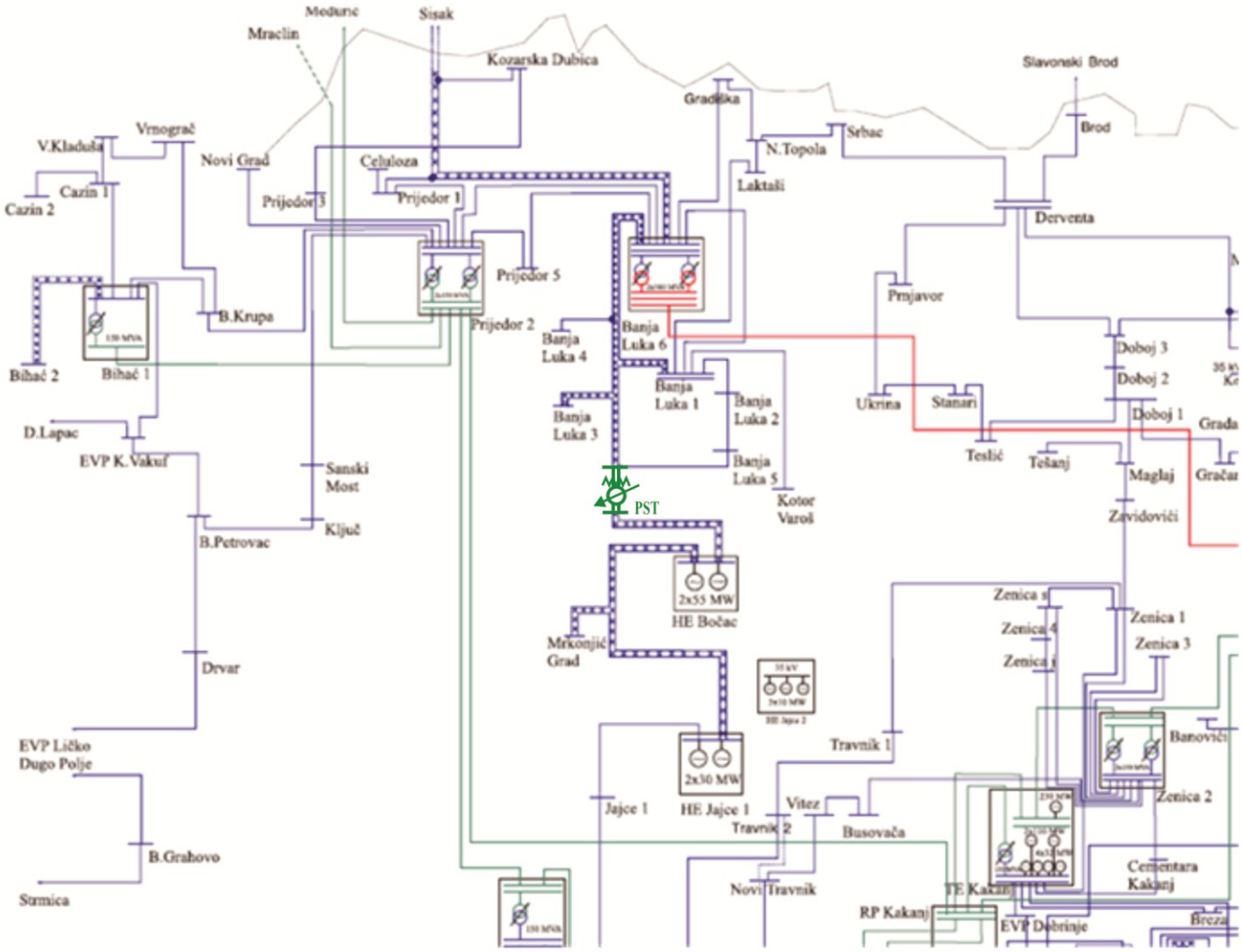


Figure 5. Part of 110kV network in BiH/RS [6]

The active power flows and losses for 110kV network from Figure 5 are at first calculated without the PST and this is considered as the base state. After that, the PST is placed on the line with the largest active power flows, in order to control active power flows through it and to decrease overall losses in the network.

By placing a PST to a series with the line with the highest active power flow, it is expected that active power flows are changed in both the observed line and the surrounding lines. Due to active power flows change, the losses also change and they are calculated by:

$$P_{loss} = R \cdot \frac{P^2 + Q^2}{U^2}, \tag{3}$$

where R is the line resistance, P and Q are the active and reactive power flows at the sending (receiving) node of the line, and U is the voltage at the sending (receiving) node of the line.

The model of the observed part of 110kV network in PWS is shown in Figure 6.

TABLE I. PARAMETERS OF TRANSMISSION LINES

Transmission line	Length [km]	x [p.u.]	r [p.u.]	Note
BL6 - Gradiška	38.2	0.12660	0.06330	Single line
BL 6 – Prijedor 1	62.6	0.10373	0.05187	Dual line
BL 6 – Prijedor 2	42.5	0.14085	0.07043	Single line
BL 6 – BL 3	15.1	0.02502	0.01251	Dual line
BL 6 – BL 1	12.7	0.02104	0.01052	Dual line
BL 3 – HE Bočac	33.1	0.05485	0.02743	Dual line
BL 1 – HE Bočac	37.2	0.06164	0.03082	Dual line
BL 1 – BL 2	4.50	0.01491	0.00746	Single line
BL 2 – BL 5	10.4	0.03447	0.01724	Single line
BL 5 – Grbići	1.80	0.00597	0.00300	Single line
HE Bočac – MG	15.4	0.02552	0.01276	Dual line
HE Bočac – Jajce 1	23.3	0.03861	0.01931	Dual line
MG – HE Jajce 1	19.3	0.03198	0.01599	Dual line
Prijedor 1 – Prijedor 2	3.70	0.01226	0.00613	Single line

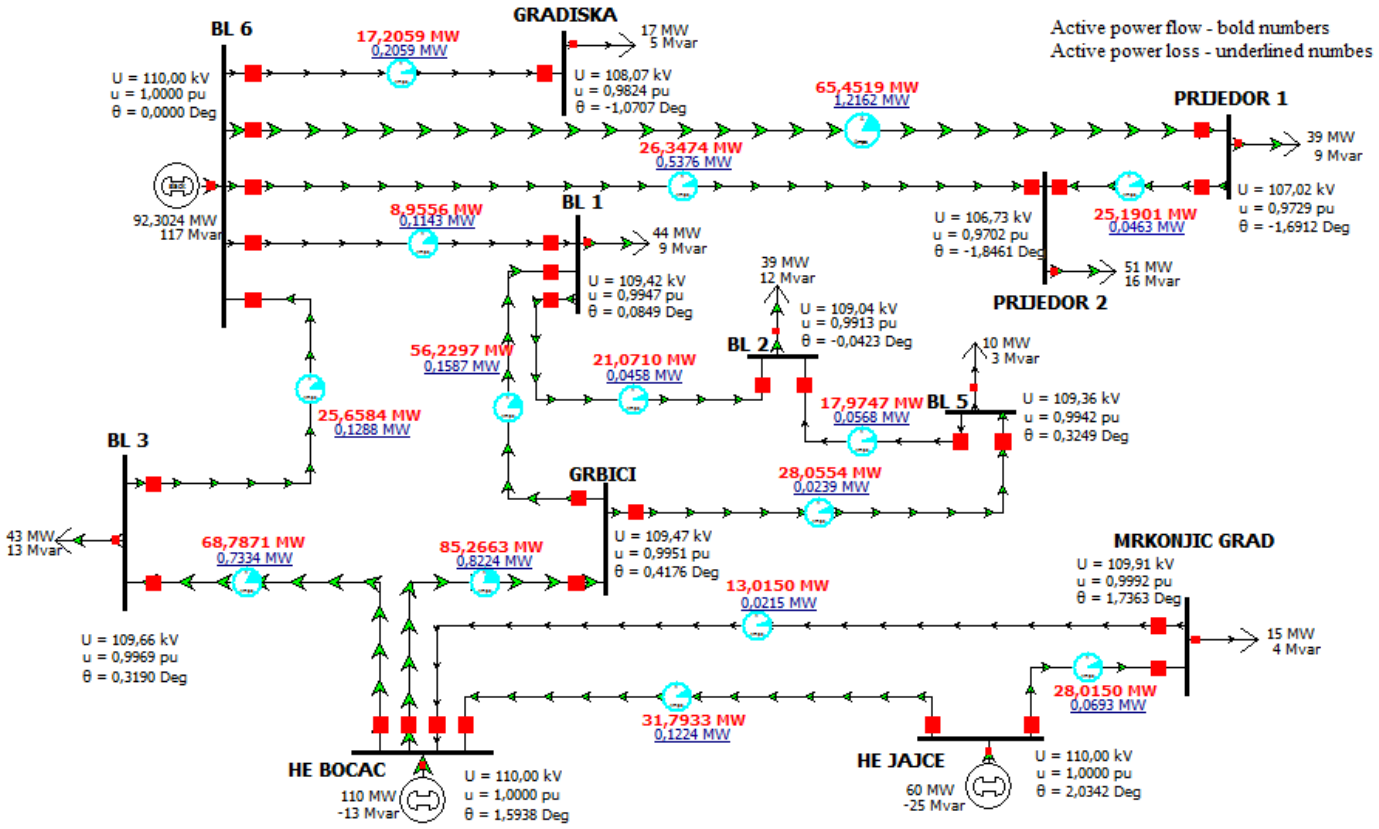


Figure 6. The PWS model of the system under consideration

The overall active power losses for the base case (without PST) for the observed network are 4.3033 MW.

In order to reduce the active power losses, the PST is placed in series with a line with the highest active power flow (Hydropower plant Bočac - Grbići, Figure 5). The phase angle of PST is changed from -20° to 20° in steps of 1° and for each phase shift the overall active power losses in the network are calculated (Figure 7).

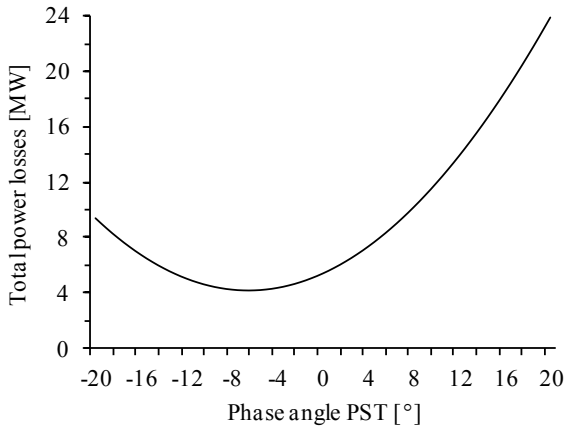


Figure 7. Overall active power losses as a function of the PST phase angle shift

The minimum overall active power losses are obtained for the PST shift angle equaling approximately -6°. The losses are reduced down to 4.1965MW, which is 2.48% lower than in the base case.

If the phase angle is further changed within the range of -6° to -7° (Figure 8) in steps of -0.1°, the minimum active power losses are obtained for the phase angle of -6.5° and they equal to 4.1878MW, or 2.68%. Relative decrease of overall active power losses in network as function of the PST phase angle shift is given in the Figure 8. Certainly, the resolution of the optimization search should be in a compliance with the resolution of the PST shift angle change.

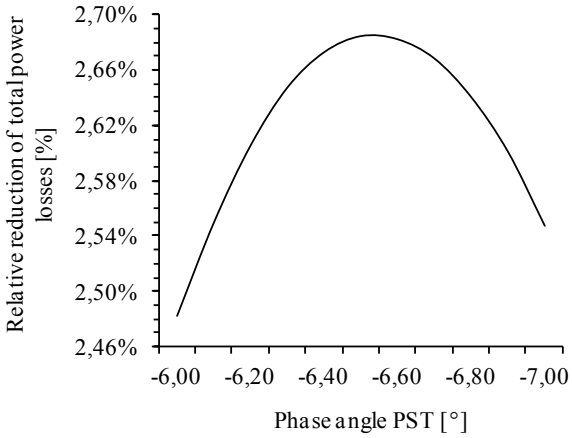


Figure 8. Relative reduction of overall active power losses as function of the PST phase angle shift

The resulting change of the active power flows does not cause overload of any line in the observed network. The voltage conditions in the network upon installing the PST with set phase angle of -6.5° are also within their permissible limits.

IV. Simplified Cost-Benefit Analysis

A simplified cost-benefit analysis for the proposed solution for minimizing power losses can be conducted by comparing the investment costs of PST and savings obtained through decreasing of the losses. Since the active power flow through the branch in which the PST is located equals 89.6434 MW, the PST with rated power of 100 MVA power can be used. A rough estimation of PST price is based on unit price of 20 BAM/kW, which estimate the costs of PST at 2 millions BAM. The average price of electric power in EPS on 110kV voltage level is 0,035 BAM/kWh, resulting in the savings on losses of approx. 100 BAM per day, which implies that the simple period of return is as high as 55 years. Based on the given analysis it can be concluded that savings of 2.68% in the active power losses are not enough for the investment to be profitable.

V. Conclusion

In this paper a possibility of decreasing overall active power losses in transmission network by using PST is analyzed. The PST is placed on the line with the largest active power flow with the aim of obtaining optimal power flows that shall result in minimum power losses. The results of the simulation of a part of EPS of BiH/RS show that by controlling phase angle by PST, the overall active power losses can be reduced by 2.68%. For this particular situation, the proposed use of PST is not an economically acceptable solution. In further work, a methodology for optimal selection of location for the PST, which will include both technical and economical aspects, should be developed.

References

[1] M. Heathcote, J & P Transformer Book, 13th ed., Newnes, 2007.

[2] Asea Brown Boveri (ABB), "*Phase shifting transformers: Reliable and efficient power flow control*", ABB Brochure 1LAB 000428, 2011.

[3] PowerWorld Simulator, "User's Guide (free version 16)", December 22, 2011.

[4] J. Verboomen, D. V. Hertem, P. H. Schavemaker, W. L. Kling, and R. Belmans, "*Phase Shifting Transformers: Principles and Applications*," IEEE Conf. on Future Power Systems, Amsterdam, November 2005.

[5] G.Preininger, Electric power transformer engineering, 2004 by CRC Press LLC.

[6] B. Nišević and D. Pantić, "*Modelling of the influence of inserting a new power plant in the existing 110kV network by three-phase short circuit current analysis*", Infoteh-Jahorina, vol.10, ref. F-34, pp.1051-1054, March 2011.

AUTHOR INDEX

CIP - Каталогизација у публикацији
Народна и универзитетска библиотека
Републике Српске, Бања Лука

621.3(082)

INTERNATIONAL Symposium Industrial Electronics (10 ; 2014 ; Banja Luka)
Symposium proceedings / X International Symposium Industrial Electronics INDEL - 2014, Banja Luka, Republika Srpska, November 6-8, 2014 ; [symposium chairman Branko Dokić]. - Banja Luka : Faculty of electrical engineering =Elektrotehnički fakultet, 2015. - 313 str. : ilustr. ; 30 cm

Tekst štampan dvostubačno. - Bibliografija uz svaki rad. - Registar.

ISBN 978-99955-46-22-9

COBISS.RS-ID 4712216