Proceedings of the 8th Small Systems Simulation Symposium 2020

February 12-14
Faculty of Electronic Engineering
Niš, Serbia

Organized by

Faculty of Electronic Engineering Niš

iCAT
Innovation Centre of Advanced Technologies

Proceedings of the 8th Small Systems Simulation Symposium 2020

Publisher:

University of Niš, Faculty of Electronic Engineering
Aleksandra Medvedeva 14,
18000 Niš
http://elfak.ni.ac.rs

Editor:

Miona Andrejević Stošović

Printed by:

Perins inženjering, Novi Sad

Number of copies printed: 40
Year of publication: 2020

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

519.876.5(082)
004.942(082)

SMALL Systems Simulation Symposium (8th ; 2020 ; Niš)
Proceedings of the 8th Small Systems Simulation Symposium 2020, February
12-14, Niš, Serbia / organized by Faculty of Electronic Engeneering, Niš and
Innovation Center of Advanced Technologies ; [editor Miona Andrejević Stošović].
- Niš : Faculty of Electronic Engineering, 2020 (Novi Sad : Perins inženjering). -
149 str. : ilustr. ; 30 cm

Tekst štampan dvostubačno. - Tiraž 40. - Bibliografija uz svaki rad. - Registar.

ISBN 978-86-6125-220-4

a) Симулација -- Зборници

COBISS.SR-ID 282783756

--

Preface

To the Proceedings of the 8th Small Systems Simulation Symposium

Dear colleagues,

This year it is 20 years from the First Small Systems Simulation Symposium. The Symposium was established in the Laboratory for Electronic Design Automation (LEDA), on the Faculty of Electronic Engineering, University of Niš.

LEDA laboratory, since its establishment in mid-eighties of the last century, went through several turbulent periods due to the political changes and wars conducted around and against the country. Nevertheless, the researchers within it were trying all the time to keep pace with the running edge of the scientific research in the field. One may say with admirable success, having in mind the circumstances. There were several areas in which particular achievements were reached: electronic modelling and simulation, CMOS IC design, theory and implementation of artificial neural networks, and renewable energy sources and circuit characterization and design. All that was only possible thanks to strong support obtained from outside the Electronic faculty especially from the University of Southampton and Middlesex University from England, The Technische Universitaet Ilmenau from Germany, The Universidad Politécnica de Madrid from Spain and others, particularly including the Technische Universitaet Wien and the Fachhochschule Wiener Neustadt. Probably the decisive support was given (first at the end of the eighties and the beginning of the nineties) by the British Council office from Belgrade (supporting the visit of Prof. Litovski to Southampton) and the Serbian Council for Intentional Cultural and Scientific Collaboration (supporting the visit of Prof. Zwolinski to Niš).

Having that in mind it is understandable that the research staff of LEDA started contemplating the necessity of deepening the international cooperation with several main goals among which the most important were acquiring new knowledge and promoting research results. One of the opportunities was to start regular meetings with the research centres with which the collaboration already lasted for more than ten years. So the idea of Small Systems Simulation Symposium (SSSS) was born in the autumn of the year 1999 in Southampton and one may claim that Prof. Robert Damper was the one who mainly contributed to the final spelling of the name.

Among the large group of authors we would like to especially appreciate the contributions of some who made the symposium to start and pertain during these two decades. In the attempt not to exaggerate we will most surely miss some names but still we wouldn't like to miss all of them. These are Prof. M. Zwolinski, Prof. R. Damper, Dr I. Bushehri, Prof. S. Bojanić, Prof. O. Nieto, Prof. V. Zerbe, Prof. D. Trajanov, Prof. V. Melikyan, and with particular respect Prof. T. Kazmierski.

It is worth to note that during this turbulent period the SSSS was trying to find the best way of survival so for some time it was organized by the Yugoslav Simulation Society. Later on, the organization was taken aback by the Faculty of Electronic Engineering with the support of the Innovation Centre of Advanced Technologies (ICNT) from Niš.

We really hope that the tradition established will last for long, especially having in mind the recognition (by the Serbian society) of the importance of the advanced technologies for the national wealth and progress. That may be seen from the fact that the surroundings of the Faculty of Electronic Engineering now is literally a construction site for buildings intended to support the research and education in the IT field. We believe that SSSS, by its actions, did in some (no matter how small) extent influenced these new believes and that it will flourish in the future together with the Serbian IT sector as a whole.

SSSS 2020 Organizing Committee

The 8th Small Systems Simulation Symposium

is supported by:

Ministry of Education, Science, and Technological Development of the Republic of Serbia

ELSYS Eastern Europe, Belgrade, Serbia

HDL Design House, Belgrade, Serbia

Veriest Venture, Belgrade, Serbia

NiCAT - Nis Cluster of Advanced Technologies, Niš, Serbia

IT Centar, Niš, Serbia

STEERING COMMITTEE OF THE SSSS2020

M. Dimitrijević, University of Niš (Serbia)
M. Andrejević Stošović, University of Niš (Serbia)
B. Blanuša, University of Banja Luka (Bosnia and Herzegovina)
S. Bojanić, Universidad politécnica de Madrid (Spain)
I. Bushehri, Lime Microsystems, (United Kingdom)
N. Cvetković, University of Niš (Serbia)
B. Dokić, University of Banja Luka (Bosnia and Herzegovina)
N. Janković, University of Niš (Serbia)
T. Kazmierski, University od Southampton (United Kingdom)
LJ.Kocarev, FCSE (FINKI) Skopje (Macedonia)
V. Litovski, University of Niš (Serbia)
V. Melikyan, Synopsys (Armenia)
S. Milenković, Lime microsystems (United Kingdom)
D. Milovanović, University of Niš (Serbia)
O. Nieto, Universidad politécnica de Madrid (Spain)
D. Pantić, University Niš (Serbia)
P. Petković, University of Niš (Serbia)
B. Reljin, University of Belgrade (Serbia)
M. Simić, School of Aerospace, Mechanical and Manufacturing Engineering (Australia)
M. Savić, Lime Microsystems (Serbia)
D. Trajanov, FCSE (FINKI) Skopje (Macedonia)
V. Zerbe, Technical University of Erfurt (Germany)
M. Živanov, University of Novi Sad (Serbia)
M. Zwolinski, University of Southampton (United Kingdom)

ORGANIZING COMMITTEE OF THE SSSS2020

M. Andrejević Stošović, University of Niš (Serbia)
M. Dimitrijević, University of Niš (Serbia)
M. Milić,University of Niš (Serbia)
S. Đorđević, University of Niš (Serbia)
V. Litovski, University of Niš (Serbia)
P. Petković, University of Niš (Serbia)
B. Jovanović, University of Niš (Serbia)
D. Mirković, University of Niš (Serbia)
V. Arsić, University of Niš (Serbia)
D. Stevanović, University of Niš (Serbia)
J. Milojković, University of Niš (Serbia)
M. Stanojlović Mirković, University of Niš (Serbia)

SYMPOSIUM SECRETARY

Miljana Milić, PhD
Faculty of Electronic Engineering
Aleksandra Medvedeva 14, Niš, Serbia

Srđan Đorđević, PhD
Faculty of Electronic Engineering
Aleksandra Medvedeva 14, Niš, Serbia

Proceedings of

The 8th Small Systems Simulation Symposium

Faculty of Electronic Engineering, Niš, Serbia, 12-14 February, 2020

Contents

Plenary Session

Session 1

Session 2

Session 3

Session 4: Poster Session

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Control Methods for Efficiency Optimization of Electrical Drives-Present Trends and Perspectives

Invited Paper

Branko Blanuša, Đorđe Lekić, Predrag Mršić

Abstract - The paper gives an overview of methods for efficiency optimization in electric drives. A historical overview of these method is given and in a more detailed actual trends are described through presented control models, simulation and experimental results. Possible directions for further research and the challenges in this field are also given in the paper.

Keywords - efficiency optimization; model based methods; search control; hybrid methods; fuzzy logic controller; artificial neural networks; energy storage

NOMENCLATURE

R_s, R_r	Resistance of stator and rotor winding
L_s, L_r	Self inductance of stator and rotor
L_m	Magnetizing inductance
P	Number of pole pairs
ϑ_s, ω_s	Rotor flux angle and angular speed
ϑ_m, ω_m	Rotor angle and angular speed
ω_r	Slip speed
s_m, L	Position
T_{em}	Electromagnetic torque
T_L	Load torque
Ψ_{sd}	Magnetizing flux
i_{sd}, i_{sq}	d and q component of stator current vector

I. INTRODUCTION

In the last 10 years, the use of electric drives has been increasing, from simple drives such as pumps, compressors, fans (heating, ventilation, air conditioning-HVAC), to high-performance servo drives which are characterized by fast response, precision and a wide range of regulated speed. Due to its simple control characteristics, for a long time the DC motor was irreplaceable in the controlled electrical drives. The relatively complex control algorithms and the need to calculate the rotate transformations made it impossible to use induction motor (IM) in servo applications. However, since the late of 1990s, the use of IM and permanent magnet synchronous motor (PMSM) in servo systems has been made possible by the use of digital systems such as microprocessors and microcontrollers, and modern power converters based on pulse-width modulation and space vector modulation [1-3]. Concept vector control, or direct torque control is generally used in IM and PMSM drives [4-7]. Also, in IM drives scalar control is also used.

It enables speed control of the by varying the supply frequency, so the ratio voltage per frequency (V/f) is constant. In this way, the torque characteristic of the motor is displaced and it is possible to regulate the rotor speed over a wide range of speeds. But with scalar control, torque cannot be directly controlled, so it is not applicable to high-performance electric drives. In high-performance drives, vector control and direct torque control are predominantly used. So, the torque and magnetizing flux can be directly controlled. The implementation of such control requires complex calculations in order to set torque and flux commands, and more complex control of the drive converter.

Very interesting research topic is efficiency improvement of electric motor drives. Considering that in industrialized countries, these drives consume more than of 2/3 of the total consumed electrical energy in industry, this issue becomes more relevant. This trend of energy efficiency improvement and the increasing presence of regulated electric drives in the industry have influenced the development of numerous methods for increasing efficiency in drives with an induction motor.

Electrical losses in an electric motor drive are calculated as the difference between the input power to the drive and the output power delivered to the mechanical subsystem. Function of these losses in the general case is a nonlinear and non-stationary function. It depends on several factors: frequency, heating, groove shape, power supply, etc. There are also additional losses, which cannot be predicted and calculated in advance. These are all reasons why optimization model should be non-linear and adaptive. In the recent years, in the implementation of such algorithms, fuzzy controllers, neural networks, or their integration in the form of neuro-fuzzy controllers have been increasingly used [8-11]. Their advantage in such applications is reflected in their non-linear structure and adaptive character. Also hybrid methods which combine different strategies for efficiency optimization become more interesting [12-16].

Today, there are a numerous number of strategies for efficiency improvement in electric motor drives. Regardless of the used method, the goal is unique; For the given operating conditions, determine the drive control so the drive operates with minimal losses. Algorithm should be fast and independent of the motor parameters variations.

Energy efficiency improvement in industrial processes is very significant in the field of horizontal and vertical

Branko Blanuša, Đorđe Lekić and Predrag Mršić are with the Faculty of Electrical Engineering, University of Banja Luka, Patre 5, 78000 Banja Luka, Republika Srpska, Bosnia and Herzegivina, E-mail: branko.blanusa@etf.unibl.org, djordje.lekic@etf.unibl.org, predrag.mrsic@etf.unibl.org

transportation. So, there are significant opportunities for increasing efficiency in electric motor drives, since the movement of the controlled object is along a known trajectory. Examples are the various conveyors and elevator drives [17 - 26]. These drives often operate in generator mode, so there is considerable potential for energy recovery and storage. Although there is a significant number of the papers that addresses these issues, the problem has not been adequately solved and it is an open subject for research.

Organization of paper is as follows: Short description of power losses is given in second section. Review and comparison of different method for efficiency optimization are described in third section. Application of fuzzy controller and neural networks in methods for efficiency improvement of electric drives are presented in fourth section. Energy efficiency improvement in elevator drives and the possibility for energy recuperation and energy storage are given in the fifth section. Some results and interesting areas for further research are summarized in the conclusion.

II. DESCRIPTION OF POWER LOSSES IN ELECTRIC DRIVES

During the energy conversion in electrical motor drive, one part of input energy converts in power losses in the converter and the electrical motor. So, power losses in electrical drives consist of converter losses, and motor losses (copper losses, iron losses and other "secondary" losses). From the practical reasons these losses are described in d-q coordinative system.

Converter losses: Main constituents of converter losses in the typical electrical drive topology are the rectifier, DC link and inverter losses. Rectifier and DC link losses are proportional to output power and these losses are not dependant from the magnetization flux. The flux-dependent losses in converter are inverter losses. These are usually given by:

$$P_{INV} = R_{INV} \cdot i_s^2 = R_{INV} \cdot \left(i_{sd}^2 + i_{sq}^2\right). \tag{1}$$

Motor losses: These losses consist of hysteresis and eddy current losses in the magnetic circuit (iron losses), losses in the stator and rotor conductors (copper losses) and stray losses. The main core losses can be modeled by

$$P_{Fe} = c_h \Psi_{sd}^2 \omega_e + c_e \Psi_{sd}^2 \omega_e^2, \tag{2}$$

where c_h is hysteresis and c_e eddy current core loss coefficient.

Copper losses depend on the effective value of current through the stator and rotor windings and these can be expressed as:

$$p_{Cu} = R_s i_{sd}^2 + R_r i_{sq}^2, \tag{3}$$

The total secondary losses (stray flux, skin effect and shaft stray losses) usually don't exceed few percent of the overall losses [3]. Therefore, these losses are usually omitted in the methods for efficiency optimization based on the loss model.

Based on previous consideration, total flux dependent losses in the drive can be given by the following expression:

$$P_\gamma = \left(R_{INV} + R_s\right) i_{sd}^2 + \left(R_{INV} + R_s + R_r\right) i_{sq}^2 + c_e \omega_e{}^2 \psi_{sd}^2 + c_h \omega_e \, \psi_{sd}^2. \tag{4}$$

In the methods for efficiency optimization motor works with the flux which is less or equal to its nominal value. So we can suppose that during efficiency optimization procedure magnetization characteristic is linear. Based on expression (4) it can be concluded that losses in the electrical drive depend on the current in stator and rotor windings, electrical frequency and level of magnetization flux.

III. METHODS FOR EFFICIENCY OPTIMIZATION IN ELECTRICAL DRIVES

Efficiency optimization algorithms are based on the adjusting magnetization flux to drive load. In this way, at lower loads, magnetization flux is also lower so the balance between the copper losses and the iron losses is achieved and total losses are reduced.

Also, it should be taken into account that the reduction of the magnetization flux degrades the dynamic characteristics of the drive. If maximum drive performance is required, the flux should be at nominal level. This should be taken into account when methods for efficiency optimization are applied.

There are many methods for efficiency optimization in electrical drives and they are usually classified in 4 strategies [2]:

- Simple State Control (SSC) ,
- Loss Model Control (LMC) and
- Search Control (SC)

Also, there are methods which can't be classified in one strategy, but combine characteristics of two strategies. These methods are usually known as hybrid methods (HM).

A. *Simple state control*

The first strategy is based on the control of one of the variables in the drive. This variable must be measured or estimated and it is used in the feedback control of the drive. Slip frequency or power factor displacement are often used variables in this control strategy. This strategy is simple but gives good results only for a narrow set of working conditions. Also, it is sensitive to variations of motor parameters due to temperature changes and nonlinearity of the magnetic circuit.

B. Loss Model Control

A good feature of loss-based model methods is that they are fast and the optimal flux is determined directly from the model. Because of this, they are often applied, especially in electric drives that have faster dynamics. Taking into account expression for electromagnetic torque in steady-state:

$$T_{em} = \frac{3}{2} P \frac{L_m^2}{L_r} i_{sd} i_{sq} \,, \tag{5}$$

and expression (4), power losses can be expressed as a function of magnetization current i_{sd} and working conditions ω_e, T_{em}:

$$\begin{aligned} P_\gamma &= \left(R_{INV} + R_s + c_e \omega_e^2 L_m^2 + c_h \omega_e^2 L_m^2\right) i_{sd}^2 \\ &+ \left(R_{INV} + R_s + R_r\right) \frac{T_{em}^2}{\left(\frac{3}{2} p \frac{L_m^2}{L_r}\right)^2} \frac{1}{i_{sd}^2} \end{aligned} \,, \tag{6}$$

Putting $A = R_{INV} + R_s + c_e \omega_e^2 L_m^2 + c_h \omega_e L_m^2$ and $B = \left(R_{INV} + R_s + R_r\right) \frac{T_{em}^2}{\left(\frac{3}{2} p \frac{L_m^2}{L_r}\right)^2}$ from (6) expression for power losses can be given as follows:

$$P_\gamma = A i_{sd}^2 + \frac{B}{i_{sd}^2} \,. \tag{7}$$

Minimization of losses is calculated from the condition $\frac{\partial P_\gamma}{\partial i_{sd}} = 0$ which gives:

$$2 A i_{sd} - \frac{2B}{i_{sd}^3} = 0 \,. \tag{8}$$

Based on (8), the reference value of the magnetization current (i_{sd}^*), which gives minimal losses for the given working conditions is:

$$i_{sd}^* = \sqrt[4]{\frac{B}{A}}. \tag{9}$$

Block diagram of loss model control implemented in induction motor drive is presented in fig 1.

Based on loss model, and working conditions, values of control variables i^*_{sd} and i^*_{sq} are generated. The main drawback of the model is that it takes into account only the most significant sources of losses in the electric drive. Also, it is sensitive to parameter variations in the model especially due to changes in temperature and nonlinearity of the magnetic circuit. Beside all, this strategy is very popular. Control variables are generated fast and directly from loss model so this method can be use even in cases frequent change of working conditions.

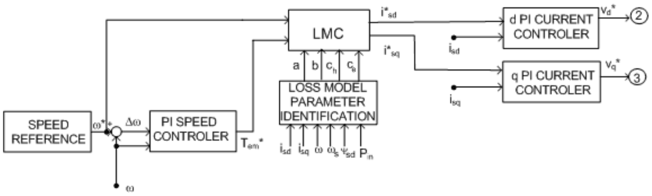


Fig. 1. Block diagram of loss model strategy implemented in induction motor drive.

C. Search Control

Application of SC methods is very simple. For two successive values of the i_{sd} current, power losses are determined as difference between input and output power of the drive:

$$P_\gamma(n) = P_{in}(n) - P_{out}(n). \tag{10}$$

$P_\gamma(n)$ is value of power losses in last sample (nT_s) , where T_s is sample time.

Sign of Δi_{sd} is maintained if power losses are reduced. Otherwise, the sign of Δi_{sd} is opposite in the next step:

$$i_{sd}(n) = i_{sd}(n-1) - \operatorname{sgn}\left(\Delta P_\gamma(n-1)\right)\Delta i_{sd} \,. \tag{11}$$

Value of Δi_{sd} is usually determined to be proportional of P_γ. It means if $P_\gamma(n)$ is greater, then $\Delta i_{sd}(n)$ is greater and vice versa. Block diagram of search control strategy is presented in fig.2.

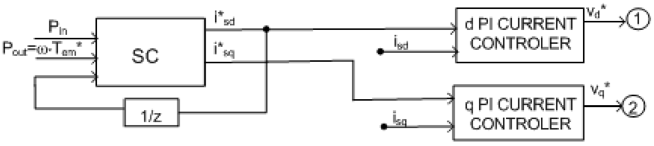


Fig. 2. Block diagram of search control strategy implemented in induction motor drive.

Search strategy methods have an important advantage compared to other strategies. These methods are not sensitive to parameter variations in the motor. Also, these methods are applicable universally to any motor. Besides all good characteristics of search strategy methods, there is an outstanding problem in their use. Flux convergence to the value which gives minimal losses can be slow for drives with higher dynamic requirements. Also, magnetization current oscillates in small steps around its optimal value it.

D. Hybrid Methods

Hybrid method usually combines characteristics of two optimization strategies SC and LMC [14-15]. It is an interesting solution and it is expected to be more used in the future. During transient process LMC is used, so fast flux changes and good dynamic performances are achieved. In a steady state search algorithm is applied. In that case parameter variations has not significant influence to method for efficiency optimization. Control diagram for the hybrid optimization method is presented in fig 3.

In hybrid methods block which determines steady state of the drive is usually used. So, for a steady state this block changes optimization control from LMC to SC and vice versa when working conditions are changed.

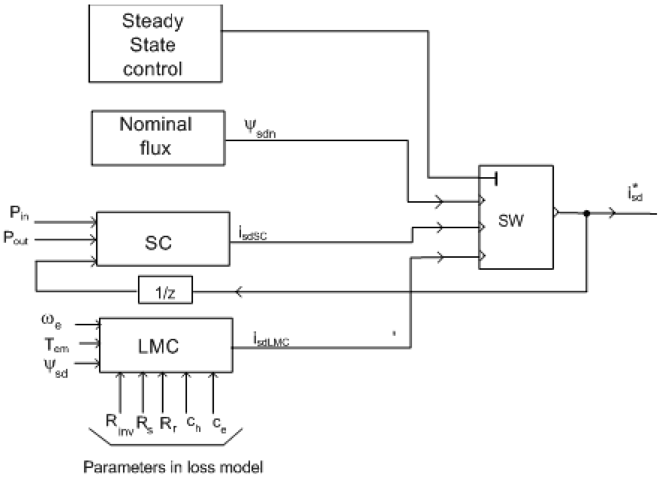


Fig. 3. Block diagram of hybrid control method implemented in induction motor drive.

IV. IMPLEMENTATION OF FUZZY LOGIC AND ARTIFICIAL NEURAL NETWORKS IN METHODS FOR ENERGY EFFICIENCY IMPROVEMENT OF ELECTRICAL DRIVES.

One of the outstanding problem in all methods for efficiency improvement is that they are sensitive to fast change of working condition. Reduction of the magnetization flux degrades the dynamic characteristics of the drive. This is presented in fig 4. For a step change of load torque speed drop is much more expressed when motor works with lower then for nominal flux. Speed response on the step change of load torque (from 0.5 p.u. to 1.1 p.u.) at t=25s, for nominal flux and when LMC method is applied, is presented in the fig. 4. That is reason why it is necessary maintain torque reserve in the methods for efficiency improvement, especially in the applications where output power changes are more frequent. The loss model is nonlinear. It is not possible to predict and take into account all the factors that affect to losses. Also, many parameters in the model are sensitive to parameter variations in drive due to temperature, saturation of the magnetic circuit, skin effect, etc. The reserve of the electromagnetic moment also should be taken into account. Based on this, it can be concluded that it is convenient to implement fuzzy controller in methods for efficiency improvement (fig. 5).

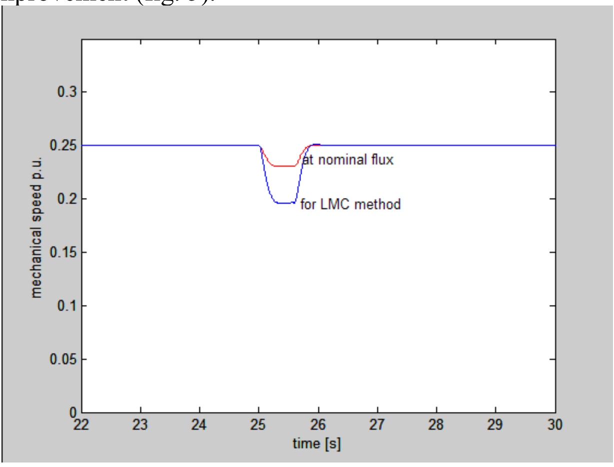


Fig. 4. Block diagram of hybrid control method implemented in induction motor drive.

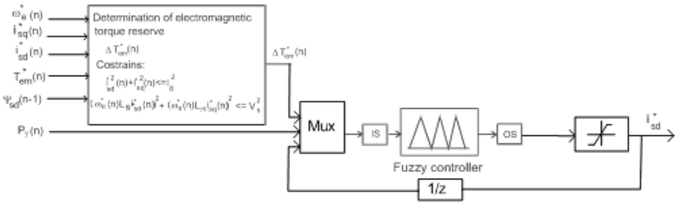


Fig. 5. Implementation of fuzzy controller in methods for efficiency optimization of electrical drives.

Fuzzy controller can be applied for both SC and LMC methods. Based on electromagnetic torque reserve, last sample of the calculated power losses and magnetization current, new value of magnetization current is determined. Electromagnetic torque is determined based on maximum torque which can be achieved for the current flux level in the motor and taking into account constrains for current and voltage (fig. 5). Fuzzy set has 12 rules, 3 inputs and 1 output. Mamdani type of fuzzy inference and centroid defuzzification method are used. For the working condition presented in fig. 6. obtained speed response and power losses are presented in Figs. 7 and 8.

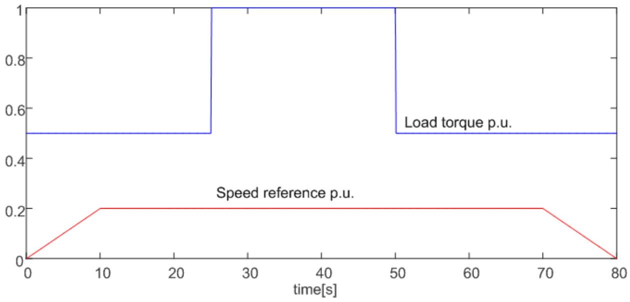


Fig. 6. Graph of mechanical speed and load torque reference.

Reference value of mechanical speed is set to 0.2 p.u. For low speed, drive is more sensitive to step change of load torque. So, control of electromagnetic torque is very important in efficiency optimization methods when drive works with low speed. It is obvious that electrical drive is more robust on the step change of load when electromagnetic torque reserve is included in method for efficiency improvement (fig. 7). But in this case the losses are slightly higher for low loads (fig. 8).

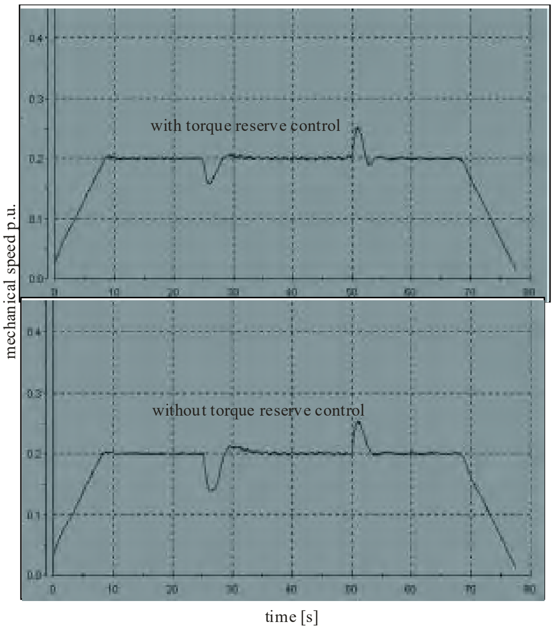


Fig. 7. Speed response for a given working conditions when LMC is applied with torque reserve and without torque reserve.

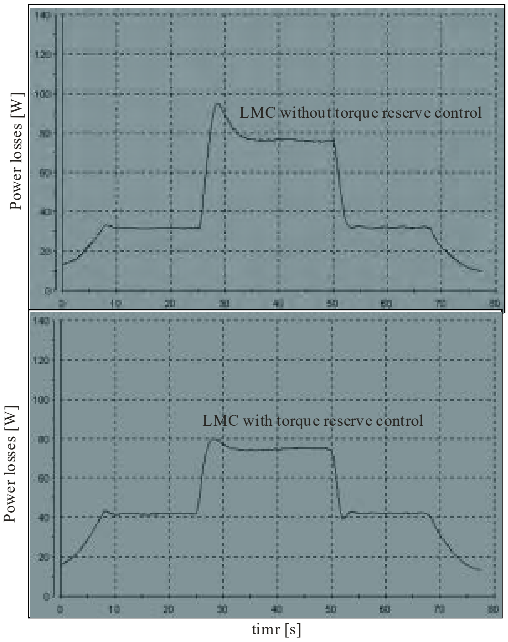


Fig. 8. Power losses for a given working conditions when LMC is applied with torque reserve and without torque reserve.

Therefore, it is necessary to maintain balance between efficiency optimization and torque reserve in order to keep good dynamic characteristics of electrical drive. More recently, neural networks have been used for faster SC methods and maintain electromagnetic torque reserve. The Hybrid method for efficiency optimization with ANN is shown in fig. 9.

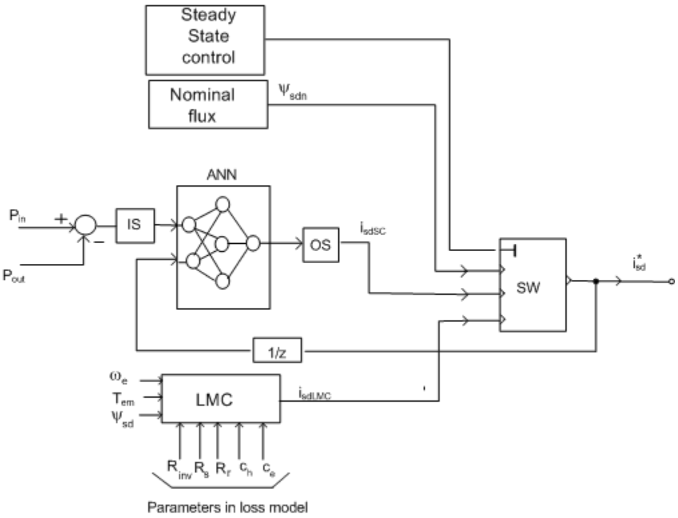


Fig. 9. Hybrid method for efficiency optimization with ANN.

ANN is used for the implementation of SC method in hybrid system for efficiency improvement. The neural network has 2 inputs, 1 output and 2 hidden layers with 7 and 5 neurons respectively. The transfer functions of the neurons in the hidden levels are sigmoidal (tansig) and the output level linear (purelin). The Levenberg-Marquardt method is used to train the network. The performance index target value is set to $5*10^{-4}$. The qualitative analysis of the network operation gives the dependence of the performance index and the mean-square error as a function of the learning epochs number. Also, for the defined working conditions of the drive (fig. 6), a trajectory of the current reference value i_{sdSC} was generated using a neural network (fig. 10), and the speed response and loss power for the nominal flux and for the hybrid method with ANN were compared and obtained results are presented in figs. 11 and 12 respectively.

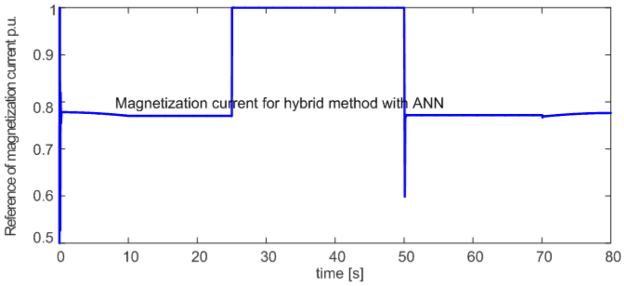


Fig. 10. Magnetization current generated in sc with ANN and for given working conditions presented in fig.6.

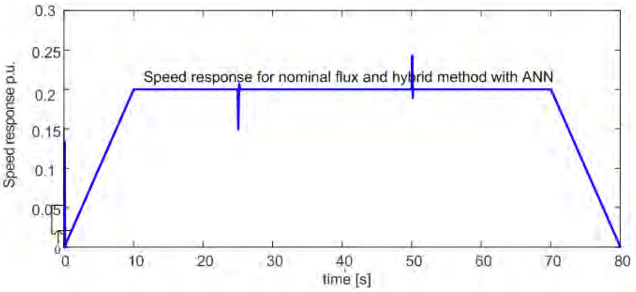


Fig. 11. Speed response for nominal flux and for hybrid method with ANN and for given working conditions presented in fig.6.

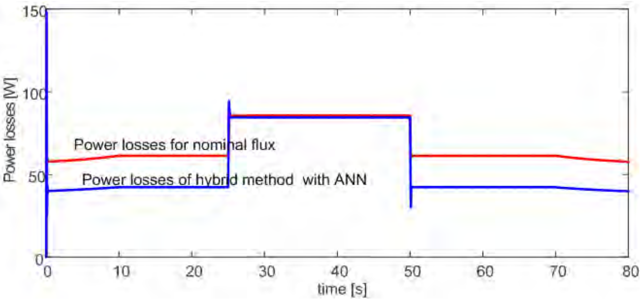


Fig. 12. Power losses for nominal flux and for hybrid method with ANN and for given working conditions presented in fig.6.

The obtained results show that the application of ANN in efficiency optimization is justified. Good results are achieved in maintaining good dynamics of the electric drive, but also in efficiency improvement.

V. EFFICIENCY OPTIMIZATION IN ELEVATOR DRIVES, ENERGY RECUPERATION AND STORAGE

The standard topology of the indirect AC/AC converter controlled elevator drives, consists of a diode rectifier, DC link, and a three-phase PWM controlled inverter. Such topology has many drawbacks and does not allow bidirectional power flow and energy recuperation.

Elevator drives often work in generator mode, so it is convenient object for efficiency optimization improvement and energy recuperation and storage. In that case units for energy storage are necessary. Usually, supercapacitors are used, due to fast charging and discharging. In this application converters which provide bidirectional power flow, and appropriate control units which automatically control the working are required. These applications must not affect to elevator ride, jerk, speed, positioning and the comfort of passengers.

One concept of efficient elevator drive with the unit for energy storage is shown in fig. 13 [22]. It consists of two tree phase back to back converters and a digitally controlled switch unit. Three phase power supply is connected with the voltage oriented control (VOC) PWM rectifier. This rectifier obtain good control of DC voltage and phase control between input voltage and current (PF≈1) (fig.13). Total harmonic distortion of input current is low. This rectifier is connected to unit switch over DC link from one side and three phase inverter from the other side. Space Vector Modulation (SVM) technique is used for inverter control. Such topology provide possibilities for the implementation of advanced control strategies like VC or SVM. The basic parameters are monitored on the storage unit; DC voltage, charging /discharging current and the capacity, i.e. state of charge (SOC), status of the main supply and the working mode of elevator motor (motor, generator).

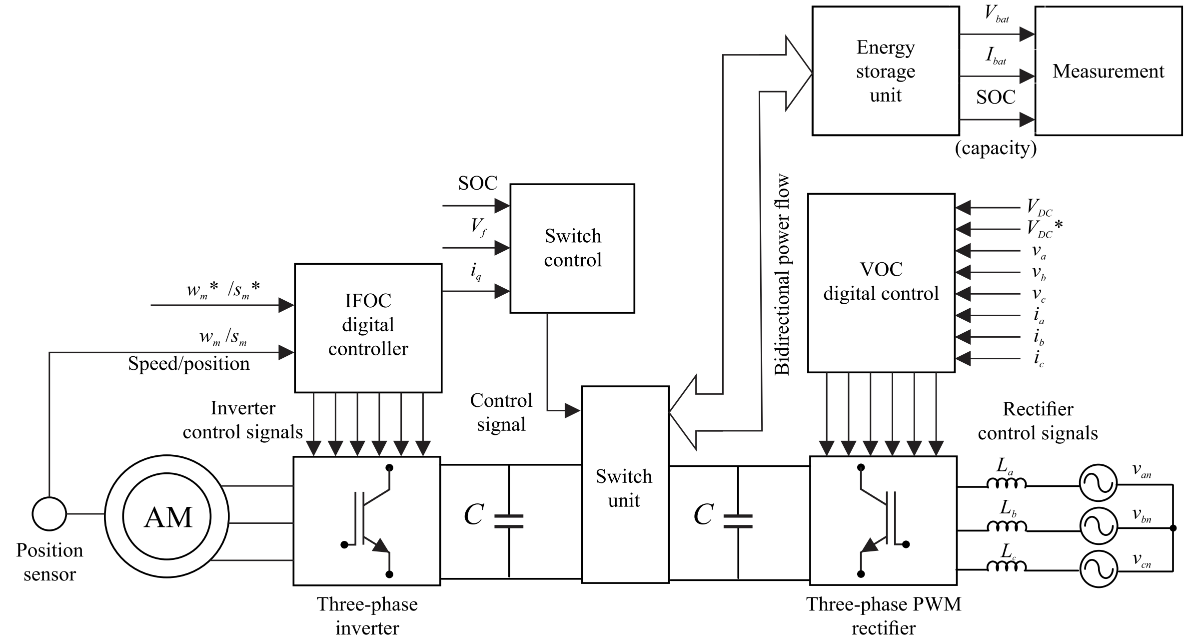


Fig. 13. One concept of elevator drive with the energy storage unit [22].

If the motor works in the generator mode and the storage unit is not full, switch is turned so the storage unit is charged with the regulated current (fig. 15). In the case of a main power failure, storage unit is used as auxiliary power supply. This process is fast and automatically realized (Fig. 16). Any changes in the operation mode are fast and from the standpoint of the elevator user seamlessly.

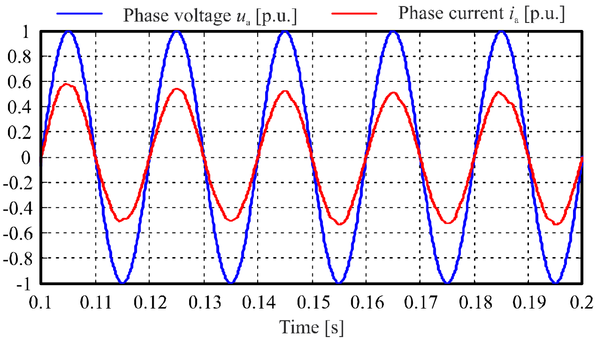


Fig. 14. Input phase voltage and current of PWM rectifier in elevator drive.

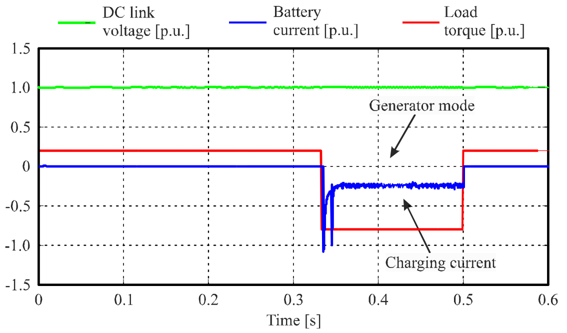


Fig. 15. Charge of energy storage unit in generator mode of elevator drive [22].

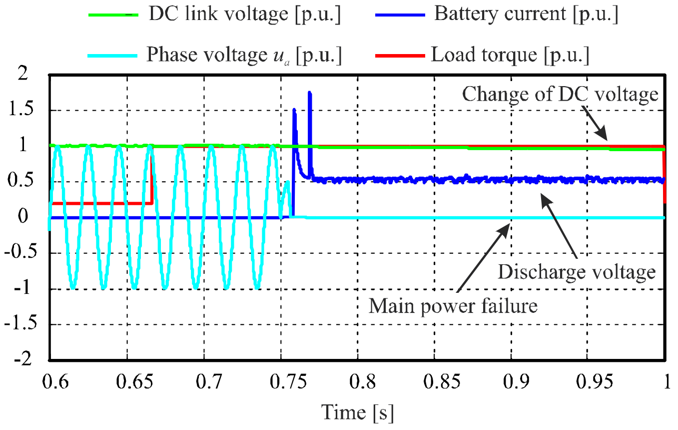


Fig. 16. Discharge of energy storage unit in a case of main power failure [22].

Studies and measurements on such elevator drive in the elevator operation show that the concept is interesting and it gives good results. Also, load perturbations or changes in operating modes have no effect on maintenance of the elevators' performance. Beside the step change of load and changes in operating modes, the elevator keeps defined speed and motion trajectories.

VI. CONCLUSIONS

The paper describes some of the interesting topics related to efficiency optimization of electric drives.

Based on content presented in the paper, the following conclusions can be drawn:

1. Controlled electric drives are suitable for implementation of methods for energy efficiency improvement.

2. In many HVAC applications it is possible to apply simple methods for efficiency y improvement, whose implementation is purely software.

3. In applications where speed and torque change more frequently, methods based on loss model and search algorithms may be used.

4. Hybrid methods, which combine the merits of loss-model based methods and search algorithms are also become more popular (fig. 3).

5. Elements of artificial intelligence, neural networks and fuzzy logic are increasingly used in methods for efficiency optimization and improve characteristics of these methods (figs. 7, 8, 10-12)

6 Horizontal and vertical transport systems are also of interest from the standpoint of energy efficiency. In addition to increasing efficiency in motor mode, it enables energy recovery and storage in generator mode. A model of one elevator system with the elements for efficiency improvement and energy storage is also presented in the paper (figs. 14-16).

APPENDIX

U_n = Δ220/Y380 V
P_n = 0.75 kW
n_n = 1390 rpm
R_s = 10 Ω
$L_{\gamma s}$ = 43.067 mH
L_{mn} = 0.4212 H

I_n = 3.6/2.1 A
$\cos\varphi$ = 0.72
f_n = 50 Hz
R_r = 6.3 Ω
$L_{\gamma r}$ = 40.107 mH
J_m = 0.00442 kgm^2

REFERENCES

[1] Abrahamsen F, Pedersen JK, Blaabjerg F: "State-of-Art of optimal efficiency control of low cost induction motor drives", *In: Proceedings of PESC'96*, pp. 920-924, 1996.

[2] F. Abrahamsen, F. Blaabjerg, J.K. Pedersen, P.Z. Grabowski and P. Thorgensen," On the Energy Optimized Control of Standard and High Efficiency Induction Motors in CT and HVAC Applications", *IEEE Transaction on Industry Applications*, Vol.34, No.4, pp.822-831, 1998.

[3] Vukosavic S.N., "Controlled electrical drives - Status of technology". In: Proceedings of XLII ETRAN Conference, No. 1, pp. 3-16, Vrnjacka Banja, Yugoslavia, 1998.

[4] Matić P., Blanuša B., Vukosavić S., "A Novel Direct Torque and Flux Control Algorithm for the Induction Motor Drive", *Proceedings of IEEE - IEMDC International Electrical Machines and Drives Conference*, pp. 965-970, Medison, USA, 2003

[5] Matić P., Raca D., Blanuša B., Vukosavić S., "A Direct Torque Controlled Induction Motor Drive Based on Control of both Tangential and Radial Component of Stator Flux Vector", *Proceedings of Applied Power Electronics Conference – APEC 2008*, D7.8., Austin, USA, 2008.

[6] Blanuša B., Matić P., Vukosavić S., " Vektra – A Test Bench for Student Exercises and Development of Digital Control Algorithms for AC Drive Control", *Electronics*, Vol.6, No.2, 2002, 26-30, 2002.

[7] Matić P., Blanuša B., Vukosavić S., "A Novell Approach to Direct Torque Control of Induction Machine Drive", *Electronics*, Vol. 6, No.1, pp. 18-21, 2002.

[8] Liwei Z., Jun L., Xuhui W.: "Systematic Design of Fuzzy Logic Based Hybrid On-Line Minimum Input Power Search Control Strategy for Efficiency Optimization of IMD", IPEMC 2006.

[9] Prymak B., Moreno-Eguilaz J.A., Peracaula J., "Neural network based efficiency optimization of an induction motor drive with vector control", *In: Proceedings of the IECON 02 Industrial Electronics Society, IEEE 2002 28th Annual Conference* , Vol. 1, pp. 146-151, 2002.

[10] Sergaki E. S., Stavrakakis G.S., "Online Search Based Fuzzy Optimum Efficiency Operation in Steady and Transient States for Dc and Ac Vector Controlled Motors", *Proceedings of the 2008 International Conference on Electrical Machine*, Paper ID 1190, 2008.

[11]Uddin M.N., Nam S.W.: "Development of a nonlinear and Model Based Online Loss Minimization Control of an IM Drive", *IEEE Transaction on Energy Conversion*, Vol.23, No.4, pp. 1015-1024, 2008.

[12] Blanuša B., Dokić B., "Efficiency Optimized Control of Induction Motor Drive in a Dynamic Operations", *Proceedings of IX National Conference with International Participation-ETAI*, E4-01, 2009.

[13] Blanuša B., Matić P., Dokić B., "New Hybrid Model for Efficiency Optimization of Induction Motor Drives", *Proceedings of 52nd International Symposium ELMAR-2010*, pp. 313-317, Zadar, Crotaia, 2010.

[14] Branko Blanuša, Petar Matić, Branko Dokić, "Hybrid Model for Efficiency Optimization of Induction Motor Drives", *Proceedings of 54th ETRAN Conference*, EL5.1-1-4, 2010.

[15] Blanusa B., Vukosavic S. N., "Efficiency Optimized Control for closed-cycle Operations of High Performance Induction Motor Drive", *Journal of Electrical Engineering*, Vol.8/2008-Edition: 3, pp. 81-88, ISSN 1582-4594, 2008.

[16] Branko Blanusa and Bojan Knezevic,"Simple Hybrid Model for Efficiency Optimization of Induction Motor Drives with Its Experimental Validation", *Advanced in Power Electronics*, Hindawi Publishing Corporation, Volume 3, Article ID 371842, 2013.

[17] I. Georgiev, S. Mircheski, "Analysis of electrical energy consumption in elevator drives", EPE-PEMC, Ls3d.3.-1-Ls3d.3.-6, 2012.

[18] P. Kubade, S,K.Umathe, Dr.D.R. Tutakne, "Regenerative Braking in an Elevator Using Supercapacitor", *International Research Journal of Advanced Engineering and Science*, pp.62-65, 2017.

[19] B. Plangklang S. Kantawong, S Dangeam, Y. Kumsuwan, "Elevator Energy Regenerative Unit (EERU) for Energy Saving in a Permanent Magnet Motor Elevator System", *Proceedings of 17th International Conference on Electrical Machines and Systems,* pp. 2120-2124, 2014.

[20] W. Wenlong, W. Fei, W. Yong and G. Qiang, "Two Parallel Converters with Small Inter Converter Inductor for Elevator Drive System",*Proceedings of 9th International Conference on Power Electronics-ECCE Asia*, pp. 334-336, 2015.

[21] Q. Xun, P. Wang, L. Quan, Y. Xu, "Research on Control Strategy of Super Capacitor Energy Storage System in Traction Elevator", *Proceedings of International Conference on Engineering Science and Management (ESM 2016)*, pp. 0160-0154, 2016.

[22] Branko Blanuša, Bojan Knežević, Đorđe Lekić, Predrag Mršić i Bojan Erceg, "Elevator Drives - Present Trends and Perspectives", *Proceedings of 11-th Conference on Electromechanical and Power Systems (SIELMEN 2017)*, pp. 1-9, 2017.

[23] B. Z. Knezevic, B. Blanusa, D. P. Marcetic, "A synergistic method for vibration suppression of an elevator mechatronic system", *Journal of Sound and Vibration*, *Elsevier*, 406C, pp. 29-50, 2017.

[24] Tomislav B. Sekara, Branko Blanusa, Bojan Knezevic," A new analytical approach for design of current and speed controllers in electrical motor drives", *Electrical Engineering, Springer*, DOI: 10.1007/s00202-018-747-z, 2018.

[25] Zhen Guo, Jisheng Zheng, Qinguei Zhong, "Research on efficiency of current-fed asynchronous motor drive based onhybrid search method", *System Science& control Engineering*, Vol 7, 2019-Issue1, pp. 89-96, 2019.

[26] Jorge Resa, Domingo Kortes, Juan FranciscoMarques-Rubio, David Navaro," Reduction of Induction Motor Energy Consumption via Variable Velocity and Flux References", *Electronics*, Vol. 8, Issue 7, 8(7), 740, 2019.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Linked Data Based Power Grid Management

Invited Paper

Vladimir Zdraveski, Dimitar Trajanov and Ljupco Kocarev

Abstract - A data model for a power delivery system is developed in this work, that combines the power grid data with the general geo-political and socio-economical data from web and data hubs. It is based on the power grid ontology, that covers the required entities and connection properties for topology and measurements annotation in a power system. In order to proof the proposed model, several data queries and reports are presented and evaluated on recently published data sets. Few end-user tools are conceptually explored in order to note the usability for power companies as well as citizens.

Keywords - power grid, ontology, query, monitoring

I. INTRODUCTION

Smart grids have already explored the benefits of information and communications technology (ICT) applied in the power systems. However, the large amount of data generated in the smart grids requires appropriate data modeling, data processing and data publishing to result in a real smart grid data understanding for the power supplying companies as well as the energy consumers. The grid of the future would tend to transform the manual operations, along with the electro-mechanical components, into a ”smart grid” and automated processes, that will require further data collection and processing, so a comprehensive data modeling will be essential [1]. The smart grid is always followed by the smart energy management, that is open for the challenges of big data-driven smart energy management in IT infrastructure, data collection and governance, data integration and sharing, processing and analysis, security and privacy, and professionals [2].

The knowledge retrieved from the collected and processed data is useful in any part of the power grid. The data-driven approach is applied to wind turbine power generation performance monitoring using supervisory control and data acquisition (SCADA) data in [3] and the results demonstrate high accuracy in detecting the abnormal power curve profiles. An optimal power flow paired with the load control is achieved using data-driven concept in [4], resulting in computational efficacy of the distributionally robust approach and trade off between cost and robustness of solutions driven by data. Power grid data modeling has been considered in [5] and a central ontology layer has been introduced. It uses a common interconnection layer for all agents and it offers unified access to the stored data, infor- mation and knowledge. An electric power knowledge theory model is proposed in [6]. It was based on ontologies and the semantic web technologies and focused on the solution of the problem of normalized modeled knowledge for the management and analysis of electric power big data.

In the field of data science there has been produced an advanced set of standards and tools for data modeling, that could also apply in the field of power system’s data management. One such initiative is the Schema.org [7], promoted by Google and Microsoft. It unifies the different existing data schemes and becomes a standard. Schema.org is actually considered as one of the main drivers for the adoption of the semantic Web principles worldwide, by a broad number of organizations and individuals in their real businesses [8].

Schema.org is already applied in different fields. A health information representation, querying, and visualization system by using Linked Data tools has been developed in [9] and have imported more than 20,000 HIV-related data elements on mortality, prevalence, incidence, and related variables. Semantic tools have been applied in the home health care systems and a cloud-based reasoning and mapping system has been built in [10] with Ontology Web Language (OWL), Resource Description Framework (RDF), Simple Protocol and RDF Query Language (SPARQL) and SPARQL Inferencing Notation (SPIN). An approach that provides integrated and situational information on different tourism-related topics is presented in [11] and the authors introduce an adaptation concept based on semantic descriptions of user context and integrated information sources and describe a prototype implementing that concept. The ongoing effort on using Linked Data technologies to improve the online visibility of touristic service providers from Innsbruck and its surroundings is presented in [12]. A contribution in the field of transportation applied to the city of London, UK is explored in [13] and a novel framework to address accessibility information barriers by establishing a linked data repository for publishing, linking and consuming the open accessibility data.

The community has already launched data portals that offer data from different fields, such as DBPedia[1] and its SPARQL endpoints[2] and WikiData[3]. DBpedia is a crowd-sourced community initiative to extract structured information from Wikipedia and make this information available on the Web. By combining the power grid data with the general geo-political and socio-economical data from such web and data hubs, certain indicators for a city or country could be calculated, that will simplify power grid quality evaluation and possibly

Vladimir Zdraveski, Dimitar Trajanov, and Ljupco Kocarev are with the Faculty of Computer Science and Engineering ss. Cyril and Methodius University - Skopje, Ruger Boskovik 16 Skopje, North Macedonia, E-mail: vladimir.zdraveski, dimitar.trajanov, ljupco.kocarev@finki.ukim.mk

Ljupco Kocarev is also with the Macedonian Academy of Science and Art, Bul. Krste Misirkov 2, P.O. Box 428, Skopje, North Macedonia, E-mail: lkocarev@manu.edu.mk

---

[1] DBPedia, http://wiki.dbpedia.org/

[2] SPARQL endpoints, http://live.dbpedia.org/sparql and https://dbpedia.org/sparql

[3] WikiData, https://www.wikidata.org/

enable ranking of the cities and countries by that indicators [14].

Authors in [15] address the energy performance of a household and the ontology of a household micro-system, while taking into account the possibility of it being controlled via energy management systems (EMS). A step towards using ontologies to describe the knowledge, concepts, and relationships in the domain of solar irradiance forecasting is proposed in [16] with a goal to develop a shared understanding for diverse stakeholders that interact with the domain. A preliminary ontology on solar irradiance forecasting, SF-ONT, was created and validated on three use cases. The OEMA (Ontology for Energy Management Applications) ontology network is presented in [17]. This ontology is an attempt to unify existing heterogeneous ontologies that represent energy performance and contextual data. The paper describes the OEMA ontology network development process, which has included ontology reuse, ontology engineering and ontology integration activities. Scholars in [18] focus on power grid modelling based on open and publicly available data from OpenStreetMap using open source software tools. A complex systems representation as tightly integrating components in the physical space (sensors, actuators) is proposed in [19], with advanced software algorithms in the cyber-space, that were called Cyber-Physical Systems (CPS). A Key Performance Indicator (KPI)-based, linked data methodology is proposed in [20] to systematically support the identification and analysis of stakeholders, the extraction of key performance information and master data that underpin stakeholders' goals. The research explored in [21] describes an open linked dataset containing data on energy efficiency improvements, i.e., recommendations and measures taken based on energy audits, from both Sweden and the US, i.e., from the Swedish Energy Agency and the US Department of Energys Industrial Assessment Centers (IAC), respectively. Authors in [22] provide a comprehensive overview of the state-of-the-art and related work for the theory, distribution, and use of the Smart Grid Architecture Model (SGAM), an approach that has been developed during the last couple of years, provides a very good and structured basis for the design, development, and validation of new solutions and technologies. The feasibility of a data integration approach that uses available ontologies and avoids ontology alignment is explored in [23]. The approach is based on the RDF representation of diverse datasets, the semantic description of these using available ontologies and the integration of these by matching literals between datasets, instead of establishing semantic links, is presented. A conceptual framework is developed and tested in the context of a simple but typical scenario where Building Information Modeling (BIM) has to be integrated with heterogeneous data sources in order to perform several analyses.

An implementation of the Universal Smart Energy Framework (USEF) through a multiagent system and a novel semantic web ontology is presented in [24]. It aligns and enriches relevant existing standards. A collective platform for raising awareness on climate change is developed in [25]. The platform collects data from smart plugs, and exports appliance consumption information and community generated energy tips as linked data, that enables users to view and compare the actual energy consumption of various appliances, and to share and discuss energy conservation tips in an open and social environment.

Considering the existing problems in the power grid systems and encouraged by the achievements of schema.org in other fields, in this paper we extend schema.org to cover the field of power grid systems. We develop the power grid ontology (PGO), i.e. a data scheme for data annotation, based on the schema.org vocabulary.

This is the outline of the paper. In Section II we explore the power grid ontology and explain the data entities and relations. The usage benefits, such as reports and queries, from a data repository based on the ontology are presented in Section III. Section IV concludes the paper and addresses the future work opportunities.

II. POWER GRID ONTOLOGY

The power grid ontology (PGO), shown in Fig. 1, introduces a data model for power distribution system's data annotation. PGO is developed on the top of schema.org and reuses and inherits many entities and properties (marked with s: in Fig. 1). However, for the most specific domain requirements new entities and properties are introduced (marked with pgo: in Fig. 1). The core entity is *Node* that represents a node in the power network, such as *generator, substation, pillar* of a *transmission line* and a *power meter*. Generators could be *renewable* such as *windturbine, solar, biomass, geothermal* and *hydroturbine* and *nonrenewable* such as *nuclear, coal, natural gas, crudeoil* and *petroleum*.

Every node has its geographical location, represented by *GeoShape* and *GeoCoordinates*, including longitude, latitude and optionally a postal address. In every node several measurements (such as voltage, current, frequency, active and reactive power) could be conducted periodically or continuously, defined by the *dateCreated* for the measurement record. In multi-phase systems, the measurements could be taken by different phase, represented by (*Phase*) in the ontology.

Two nodes could be connected with a connection (*Link)* that is addition- ally described by *length*, *operator*, *impedance*, *frequency*, number of *cables* and number of *wires*. Every link could be a part of a *Transmission Line*, thus providing a more comprehensive big-picture of a whole geographical area or a city, towards a smart grid and a smart city.

After a sufficiently large data set is annotated with the PGO, it could be used to generate (periodically and continuously) a set of reports, such as to find the most over-loaded node or transmission line or to find the nodes with a variable frequency or the most frequent voltage drops. Few other reports/queries could be as follows:

- List the transmission lines with a voltage under certain level;

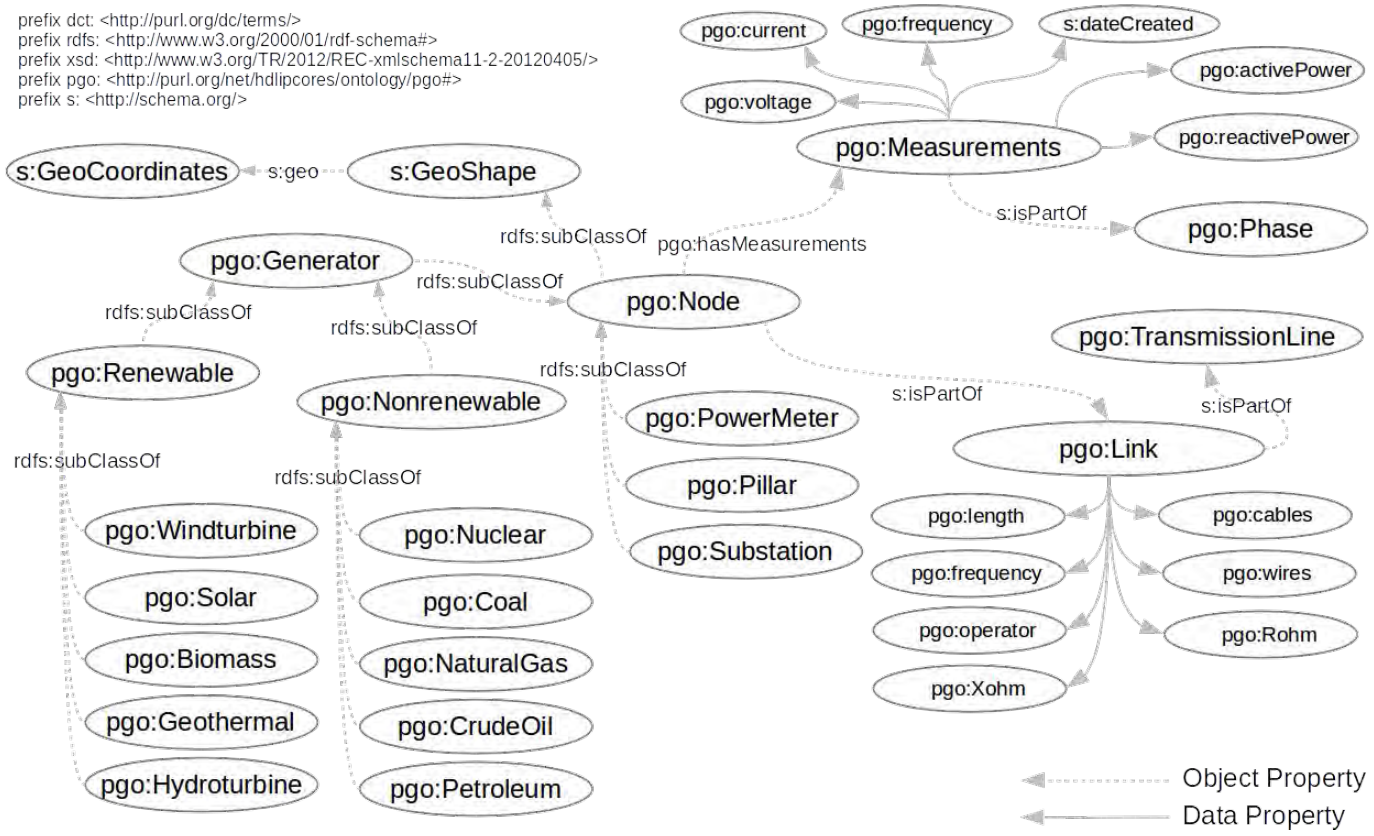


Fig. 1. Power grid ontology (PGO) provides a data model for power distribution system's data (such as network topology and continuous and periodic measurements) annotation.

- Fetch the critically overloaded transmission line;
- Fetch the critically overloaded generator;
- List nodes with the most frequent current changes;
- Fetch the average load of the nodes on a street in a city and
- List the nodes with the current under certain level.

On the other hand, advanced end-user friendly tools could be provided, such as a power distribution system city map, by analogy with the maps in [14], representing all transmission lines and nodes and possibly alarming people for any stability issues periodically or in real time. These tools would be helpful for both the power delivery companies as well as the citizens, since they will provide a high-level power network map with all required parameters for the companies. That will allow them to make in-time decisions. Whereas, there will be a real time information for a possible overload for the citizens, so they will react by switching off some of the devices, at least to protect them, and it will result in a negative feedback behavior, that will return the the power system in a stable state. Few possible types of smart-phone notifications are listed in Fig. 2.

The power grid data annotation using the PGO would be a strait forward process, since most of the node types already contain measurement units and even communication modules. For example, a small part of a street already contains tens of smart power meters, that could measure the required values and send them over network or Internet. All those nodes have their coordinates and belong to a transmission line, that is a concept and structure also provided in our ontology.

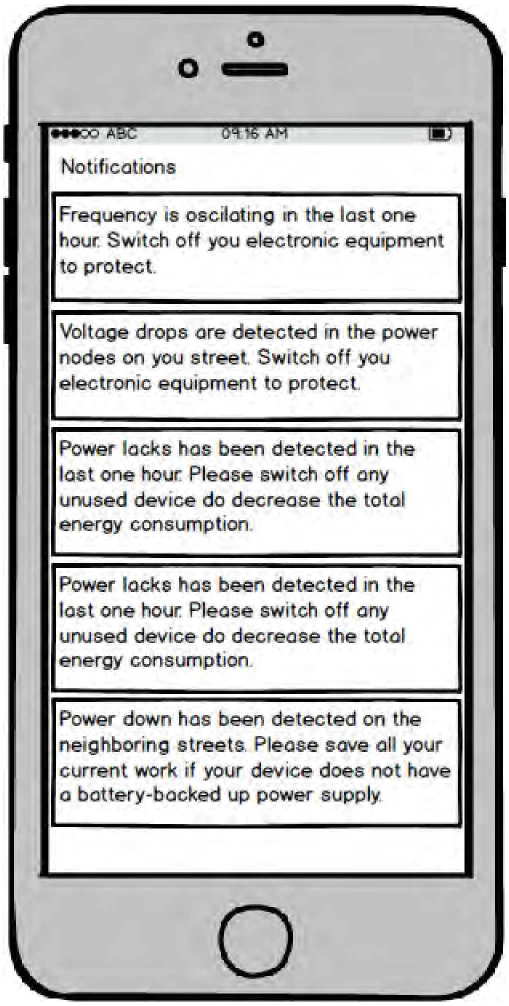
A. Annotation tool

To test the ontology, an annotation software tool[4] was developed and released as an open source code. There we provide the basic ontology entities and relations and it could be used as a Java API to build any application logic on the top of our ontology. In the API there are PGO.java and DATAREPO.java that cover entities and relations as well as the URL pattern definitions and also Anotator.java that contains the annotation methods for the data set we work with, but could be also extended for any other dataset.

III. REAL CASE SCENARIO

Several power grid data sets have been published recently, such as SciGRID[5] and GridKit[6], that we have used in order to evaluate our concept and the ontology itself. The GridKit dataset explores the vertices and the links in Germany as

[4] Software tool, https://bitbucket.org/zdrave/pgo/overview
[5] SciGrid, http://scigrid.de/
[6] GridKit, https://github.com/bdw/GridKit



Notifications

Frequency is oscilating in the last one hour. Switch off you electronic equipment to protect.

Voltage drops are detected in the power nodes on you street. Switch off you electronic equipment to protect.

Power lacks has been detected in the last one hour. Please switch off any unused device do decrease the total energy consumption.

Power lacks has been detected in the last one hour. Please switch off any unused device do decrease the total energy consumption.

Power down has been detected on the neighboring streets. Please save all your current work if your device does not have a battery-backed up power supply.

Fig. 2. Smart phone notifications.

separate files. We have annotated them and merged them in a single RDF model using the entities and the relations of our ontology. Data such as the name, geo-coordinates, voltages and links length are provided in the data set, available through our SPARQL endpoint[7].

Different types of queries could be executed against that repository. One simple illustration, would be to get a list of power plants in Germany and draw them on a map, as shown in Fig. 3. Moreover, if the power plants' data is available in real time, then appropriate status data could be also displayed, and notifications could be sent respectively. The details for the SPARQL query and results are shown in the appendix Section A in Listing 1 and Table VI.

The top ten operators in Germany, sorted by the total length of all the power distribution system links that they own is shown in Table I and the SPARQL query is presented in Listing 2.

TABLE I

TOP 10 OPERATORS BY LINKS' LENGTH.

operator	links length (m)
RWE	2.44718e+06
TenneT	2.12231e+06
TenneT TSO GmbH	1.48374e+06
50Hertz Transmission	682354.0
50Hertz	671870.0
Amprion	479220.0
50Hertz Transmission GmbH	364219.0
EnBW	318911.0
TenneT TSO GmbH;E.ON Netz GmbH	215094.0
EnBW;Amprion	202539.0

More complex queries may result in a deeper statistic as well as real time data reports. One such example is the query that returns the critical links as shown in Fig. 4. A click on the specific link would result with a pop-up window showing important data for the selected link. At the moment, there is no measurements data available, thus the data shown in Fig. 4 is only a show-case example that should illustrate the behavior. The query details are listed in Listing 3.

Furthermore, combining the power grid data with other existing data from DBPedia, could result in indicators and numbers such as power links length per square meter or number of generators and generated power per square meter or per capita. For example, the total power lines length in Germany is 10 100 *km* and the total population of 82 175 700 would result in a new indicator of power links length per capita with a value of $0,\ 12290738\ m$.

By combining our PGO repository (via our SPARQL endpoint) and the DBPedia repository (via the DBPedia SPARQL endpoint[8], as shown in Fig. 5) and introducing a more granular indicators, we could find the cities placed in radius of 20*km* (or any other distance) from a power plant and calculate the total population living around the power plants, as shown in Table II. The complete SPARQL query is presented in Listing 4.

Using the data available for the substations on our SPARQL endpoint and the city population data from DBPedia's SPARQL endpoint, we can introduce a new indicator as citizens per substation. The top ten cities with population over 100 000 in Germany, ranked by population per substation are listed in Table III and the full SPARQL query is presented in Listing 5. On the other side of the ranking table, the flop (worst) ten cities ranked by this indicator are listed in Table IV.

Another indicator could be the number of power operators per city. In this case, the PGO data for the power nodes' operators has been used and combined with the geographical data (city position, population and area) from DBPedia. To improve the precision (compared to the queries in previous examples/indicators) in this indicator we did an approximation of the city with a circle, deriving the radius from the city area. The results are shown in Table V and the whole SPARQL query is presented in Listing 6.

---

[7] PGO SPARQL endpoint, http://hdlipcores.finki.ukim.mk/sparql

[8] DBPedia SPARQL endpoint, http://live.dbpedia.org/sparql

Fig. 3. Power Plants in Germany.

Fig. 4. Critical links in Germany.

IV. CONCLUSIONS AND FUTURE WORK

A data model for a power delivery system was presented with the power grid ontology (PGO), that covered the required entities and connection properties for topology and measurements annotation in a power system. In order to proof the proposed model, several data queries and reports were presented, and few end-user tools were conceptually explored in order to emphasize the usability for power companies as well as citizens. Such a formal data model could simplify many previous power grid solutions related to the data and could significantly improve our previous work in the field of dynamic intelligent load balancing in power distribution networks [26] and reactive power compensation switch embedded in power meters [27] by semantic data annotation, that would allow semantic tools usage.

TABLE II
CITIES AND POPULATION IN RADIUS OF 20 KM FROM A POWER PLANTS.

Power Plant	Distance (m)	City	Population
Gemeinschaftskraftwerk Kiel	3021.53	Kiel	240832
Koepchenwerk	5607.79	Hagen	191241
Koepchenwerk	11579.8	Dortmund	575944
Koepchenwerk	18027.1	Bochum	361876
Kraftwerk Scholven	10144.6	Bottrop	117450
Kraftwerk Scholven	11287.4	Gelsenkirchen	260900
Kraftwerk Scholven	15564.4	Herne, North Rhine-Westphalia	166187
Kraftwerk Scholven	15747.0	Oberhausen	214990
Kraftwerk Scholven	16571.7	Essen	589075
Kraftwerk Scholven	19545.8	Bochum	361876
Statkraft Kraftwerk Knapsack II	11083.7	Cologne	1057327
TOTAL			4137698

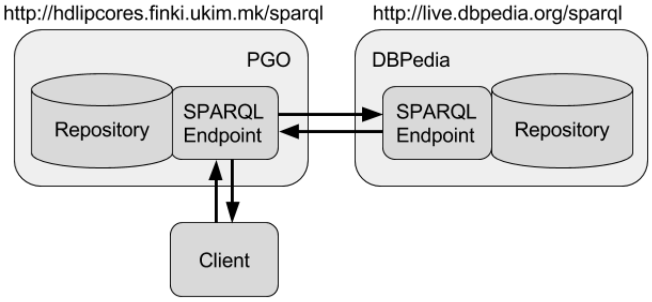


Fig. 5. Remote SPARQL endpoints. One part of the SPARQL query is executed on PGO repository and another part is forwarded to DBPedia. Afterwards, the data received from DBPedia is merged to the PGO data and additionally filtered before to be sent to the client.

TABLE III
TOP TEN CITIES (WITH OVER 100 000 POPULATION) ORDERED BY POPULATION PER SUBSTATION.

City	Population	Substations	Citizens/Substation
Hamburg	1774242	3	591414
Munich	1517868	7	216838
Nuremberg	498876	3	166292
Potsdam	161468	1	161468
Oldenburg	160907	1	160907
Mnster	300000	2	150000
Chemnitz	243521	2	121760
Jena	105192	1	105192
Hanover	518386	5	103677
Kassel	195530	2	97765

TABLE IV
FLOP TEN CITIES (WITH OVER 100 000 POPULATION) ORDERED BY POPULATION PER SUBSTATION.

City	Population	Substations	Citizens/Substation
Bottrop	117450	28	4194
Herne, North Rhine-Westphalia	166187	31	5360
Mlheim	168956	29	5826
Oberhausen	214990	30	7166
Gelsenkirchen	260900	30	8696
Braunschweig	250556	22	11388
Leverkusen	161279	14	11519
Solingen	161366	13	12412
Bochum	365406	28	13050
Krefeld	235860	15	15724

REFERENCES

[1] A. Ipakchi and F. Albuyeh, “Grid of the future,” *IEEE Power and Energy Magazine*, vol. 7, pp. 52–62, March 2009.

[2] K. Zhou, C. Fu, and S. Yang, “Big data driven smart energy management: From big data to big insights,” *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 215 – 225, 2016.

[3] H. Long, L. Wang, Z. Zhang, Z. Song, and J. Xu, “Data-driven wind turbine power generation performance monitoring,” *IEEE Transactions on Industrial Electronics*, vol. 62, pp. 6627–6635, Oct 2015.

[4] Y. Zhang, S. Shen, and J. L. Mathieu, “Data-driven optimization approaches for optimal power flow with uncertain reserves from load control,” in *2015 American Control Conference (ACC)*, pp. 3013–3018, July 2015.

[5] I. Budinska, V. Oravec, and B. Frankovic, “Central ontology layer for power grid scheduling,” in *2006 IEEE International Conference on Mechatronics*, pp. 267–271, July 2006.

[6] Y. Huang and X. Zhou, “Knowledge model for electric power big data based on ontology and semantic web,” *CSEE Journal of Power and Energy Systems*, vol. 1, no. 1, pp. 19–27, 2015.

[7] R. V. Guha, D. Brickley, and S. Macbeth, “Schema. org: Evolution of structured data on the web,” *Communications of the ACM*, vol. 59, no. 2, pp. 44–51, 2016.

[8] M. Hepp, “The web of data for e-commerce: Schema. org and goodrelations for researchers and practitioners,” in *International Conference on Web Engineering*, pp. 723–727, Springer, 2015.

[9] B. Tilahun, T. Kauppinen, C. Keßler, and F. Fritz, “Design and development of a linked open data-based health information representation and visualization system: potentials and preliminary evaluation.,” *JMIR medical informatics*, vol. 2, no. 2, pp. e31–e31, 2013.

[10] M. Hennessy, C. Oentojo, and S. Ray, “A framework and ontology for mobile sensor platforms in home health management,” in *Engineering of Mobile-Enabled Systems (MOBS), 2013 1st International Workshop on the*, pp. 31–35, IEEE, 2013.

[11] C. Keller, R. Pöhland, S. Brunk, and T. Schlegel, “An adaptive semantic mobile application for individual touristic exploration,” in *International Conference on Human-Computer Interaction*, pp. 434–443, Springer, 2014.

[12] I. Toma, C. Stanciu, A. Fensel, I. Stavrakantonakis, and D. Fensel, “Improving the online visibility of touristic service providers by using semantic annotations,” in *European Semantic Web Conference*, pp. 259–262, Springer, 2014.

[13] C. Ding, M. Wald, and G. Wills, “Open accessibility data interlinking,” in *International Conference on Computers for Handicapped Persons*, pp. 73–80, Springer, 2014.

[14] V. Zdraveski, K. Mishev, D. Trajanov, and L. Kocarev, “Iso-standardized smart city platform architecture and dashboard,” *IEEE Pervasive Computing*, vol. 16, no. 2, pp. 35–43, 2017.

[15] J. Kott and M. Kott, “Generic ontology of energy consumption households,” *Energies*, vol. 12, no. 19, p. 3712, 2019.

TABLE V
CITIES (WITH OVER 100 000 POPULATION) ORDERED BY THE NUMBER OF POWER OPERATORS.

City	Population	Area	Approx. Radius	Operators	Citizens/Operator	Area/Operator
Hamburg	1774242	755.0	15502.4	6	295707	125.833
Cologne	1057327	405.15	11356.2	5	211465	81.03
Dortmund	575944	280.4	9447.44	4	143986	70.1
Duisburg	488005	232.82	8608.66	4	122001	58.205
Bielefeld	327199	257.8	9058.72	4	81799	64.45
Leverkusen	161279	78.85	5009.87	3	53759	26.2833
Bochum	365406	145.4	6803.11	3	121802	48.4667
Salzgitter	106077	223.96	8443.27	3	35359	74.6533
Dresden	536107	328.8	10230.4	3	178702	109.6
Hamm	182022	226.26	8486.51	3	60674	75.42

[16] A. Kantamneni and L. E. Brown, "An ontology for solar irradiation forecast models.," in *KEOD*, pp. 261–268, 2018.

[17] J. Cuenca, F. Larrinaga, and E. Curry, "A unified semantic ontology for energy management applications.," in *WSP/WOMoCoE@ ISWC*, pp. 86–97, 2017.

[18] W. Medjroubi, U. P. Müller, M. Scharf, C. Matke, and D. Kleinhans, "Open data in power grid modelling: new approaches towards transparent grid models," *Energy Reports*, vol. 3, pp. 14–21, 2017.

[19] M. Sabou, S. Biffl, A. Einfalt, L. Krammer, W. Kastner, and F. J. Ekaputra, "Semantics for cyber-physical systems: A cross-domain perspective,"

[20] Y. Li, *Linked data for enhanced energy management at district and building levels: an indicator-based methodology*. PhD thesis, Arquitectura, 2017.

[21] E. Blomqvist, P. Thollanderb, R. Keskisärkkä, and S. Paramonovab, "Energy efficiency measures as linked open data," 2014.

[22] M. Uslar, S. Rohjans, C. Neureiter, F. Pröstl Andrén, J. Velasquez, C. Steinbrink, V. Efthymiou, G. Migliavacca, S. Horsmanheimo, H. Brunner, *et al.*, "Applying the smart grid architecture model for designing and validating system-of-systems in the power and energy domain: A european perspective," *Energies*, vol. 12, no. 2, p. 258, 2019.

[23] D. Patsias, "Addressing the limited adoption of semantic web data integration in aec industry: Exploration of an rdf-based approach depending on available ontologies and literals matching," Master's thesis, University of Twente.

[24] J.-L. Hippolyte, S. Howell, B. Yuce, M. Mourshed, H. A. Sleiman, M. Vinyals, and L. Vanhée, "Ontology-based demand-side flexibility management in smart grids using a multi-agent system," in *2016 IEEE International Smart Cities Conference (ISC2)*, pp. 1–7, IEEE, 2016.

[25] G. Burel, L. S. Piccolo, and H. Alani, "Energyuse-a collective semantic platform for monitoring and discussing energy consumption," in *International Semantic Web Conference*, pp. 257–272, Springer, 2016.

[26] V. Zdraveski, M. Todorovski, and L. Kocarev, "Dynamic intelligent load balancing in power distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 157–162, 2015.

[27] V. Zdraveski, M. Todorovski, D. Trajanov, and L. Kocarev, "Dynamic load balancing and reactive power compensation switch embedded in power meters," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2016.

APPENDIX A
SPARQL QUERIES AND REPORTS

The following SPARQL queries could be directly executed at our PGO SPARQL endpoint[9] and some of them, in the background, will also use data from DP Pedia[10].

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo:
<http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s:
<http://schema.org/>

SELECT ?name ?lat ?lng
WHERE {
        ?generator rdf:type pgo:Generator .
        ?generator s:name ?name  .
        ?generator s:geo ?geoCoordinates .
        ?geoCoordinates s:latitude ?lat .
        ?geoCoordinates s:longitude ?lng
}
```

Listing 1. SPARQL to list all power plants.

TABLE VI
RESULTS: LIST OF ALL POWER PLANTS.

name	lat/lng
Gemeinschaftskraftwerk Kiel	10.1789348237874 54.3388198256183
Kernkraftwerk Grohnde	9.40954428358131 52.0343295468096
Koepchenwerk	7.45130176031297 51.4130023629132
Kavernenkraftwerk Sckingen	7.95921958358891 47.5783600004331
Solarpark	10.6123957294136 48.0313556441031
Kavernenkraftwerk Wehr	7.94281008665733 47.6530214005328
Kraftwerk Scholven	7.00616826157462 51.5997745902162
Gode Wind I	6.98646123119149 54.0178365789963
Grubengas-Heizkraftwerk	7.23233534324937 49.3579040643531
Pumpspeicherkraftwerk Hohenwarte II	11.4747302318416 50.6039905453239
Statkraft Kraftwerk Knapsack II	6.84810730266128 50.8617173194521

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo:
<http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s:
<http://schema.org/>

SELECT
?operator
SUM(xsd:float(?length)) AS ?total_links_length
WHERE {
        ?link rdf:type pgo:Link .
        ?link s:name ?name  .
        ?link pgo:length ?length .
        ?link pgo:operator ?operator
}
GROUP BY ?operator
ORDER BY DESC(?total_links_length)
LIMIT 10
```

Listing 2. SPARQL to show the top 10 operators by links’ length.

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo:
<http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s:
<http://schema.org/>

SELECT DISTINCT
(?link) ?link_name ?node1_lat ?node1_lng
?node2_lat ?node2_lng ?length
WHERE {
?link rdf:type pgo:Link .
?link s:name ?link_name  .
?link pgo:length ?length .
?link pgo:voltage '380000' .
?link pgo:operator 'RWE' .
?link s:hasPart ?node1 .
?link s:hasPart ?node2 .
?node1 s:geo ?gc1 .
?gc1 s:latitude ?node1_lat .
?gc1 s:longitude ?node1_lng .
?node2 s:geo ?gc2 .
?gc2 s:latitude ?node2_lat .
?gc2 s:longitude ?node2_lng .
?node1 pgo:hasMeasurements ?n1Measurements .
?n1Measurements pgo:reactivePower ?reactivePower .
FILTER
(
(
xsd:float(?activePower)
/
math:sqrt(
xsd:float(?activePower)*xsd:float(?activePower) +
xsd:float(?reactivePower)*xsd:float(?reactivePower)
)
) < 0.9
)
}
```

Listing 3. SPARQL to show the critical links with high reactive power.

---

[9] PGO SPARQL endpoint, http://hdlipcores.finki.ukim.mk/sparql

[10] DB Pedia SPARQL endpoint, http://live.dbpedia.org/sparql

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo:
<http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s:
<http://schema.org/>

#remote prefixes
PREFIX  dbo:
<http://dbpedia.org/ontology/>
PREFIX dbr:
<http://dbpedia.org/resource/>
PREFIX dbp:
<http://dbpedia.org/property/>
PREFIX geo:
<http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX afn:
<http://jena.apache.org/ARQ/function#>

SELECT
?genName
MAX (
bif:acos(bif:sin(xsd:float(?lat) * 3.14159 / 180) *
bif:sin(xsd:float(?genLat) * 3.14159 / 180)+
bif:cos(xsd:float(?lat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLong) * 3.14159 / 180 -
xsd:float(?long) * 3.14159 / 180))*6371000
) as ?distance
xsd:string(?cityName) as ?city
MAX(?pop) as ?population

WHERE {

  SERVICE <http://live.dbpedia.org/sparql> {
    # The remote part of the query
       ?city rdf:type dbo:City;
             rdfs:label ?cityName;
             dbo:country ?country;
             dbo:country dbr:Germany;
             geo:lat ?lat;
             geo:long ?long;
             dbo:populationTotal ?pop .
             FILTER (lang(?cityName) = 'en')
    # End of The remote part of the query
  }#End of Service

  SERVICE <http://localhost:8890/sparql> {
    # The remote part of the query
    ?generator rdf:type pgo:Generator .
    ?generator s:name ?genName  .
    ?generator s:geo ?geoCoordinates .
    ?geoCoordinates s:latitude ?genLat .
    ?geoCoordinates s:longitude ?genLong .
    # End of The remote part of the query
  }#End of Service


FILTER(
(
bif:acos(bif:sin(xsd:float(?lat) * 3.14159 / 180)*
bif:sin(xsd:float(?genLat) * 3.14159 / 180)+
bif:cos(xsd:float(?lat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLong) * 3.14159 / 180 -
xsd:float(?long) * 3.14159 / 180))*6371000
) < 20000.0
) .

}
group by ?genName ?cityName
order by ?genName ?distance
```

Listing 4. SPARQL to show the cities and population in radius of 20 km from a power plant.

```
PREFIX rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs:
<http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo:
<http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s:
<http://schema.org/>

#remote prefixes
PREFIX  dbo:
<http://dbpedia.org/ontology/>
PREFIX dbr:
<http://dbpedia.org/resource/>
PREFIX dbp:
<http://dbpedia.org/property/>
PREFIX geo:
<http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
PREFIX afn:
<http://jena.apache.org/ARQ/function#>

SELECT
xsd:string(?cityName) as ?city
MAX(?pop) as ?population
COUNT(?node) as ?sustations
(MAX(?pop)/COUNT(?node)) as ?citizensPerSubstation

WHERE {

  SERVICE <http://dbpedia.org/sparql> {
    # The remote part of the query
       ?city rdf:type dbo:City;
             rdfs:label ?cityName;
             dbo:country ?country;
             dbo:country dbr:Germany;
             geo:lat ?lat;
             geo:long ?long;
             dbo:populationTotal ?pop .
             FILTER (lang(?cityName) = 'en')
             FILTER (?pop > 100000)
    # End of The remote part of the query
  }#End of Service

  SERVICE <http://localhost:8890/sparql> {
    # The remote part of the query
    ?node rdf:type pgo:Substation .
    #?node s:name ?nodeName  .
    ?node s:geo ?geoCoordinates .
    ?geoCoordinates s:latitude ?genLat .
    ?geoCoordinates s:longitude ?genLong .
    # End of The remote part of the query
  }#End of Service


FILTER(
(
bif:acos(bif:sin(xsd:float(?lat) * 3.14159 / 180)*
bif:sin(xsd:float(?genLat) * 3.14159 / 180)+
bif:cos(xsd:float(?lat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLong) * 3.14159 / 180 -
xsd:float(?long) * 3.14159 / 180))*6371000
) < 20000.0
) .

}
group by ?cityName
order by DESC(MAX(?pop)/COUNT(?node))
LIMIT 10
```

Listing 5. SPARQL to show the top ten cities (with over 100 000 population) ordered by population per substation.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX pgo: <http://purl.org/net/hdlipcores/ontology/pgo#>
PREFIX s: <http://schema.org/>

#remote prefixes
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX dbp: <http://dbpedia.org/property/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX afn: <http://jena.apache.org/ARQ/function#>

SELECT
xsd:string(?cityName) as ?city
MAX(?pop) as ?population
(MAX(?areaTotal) / 1000000) as ?area
MAX(bif:sqrt(?areaTotal / 3.14159 )) as ?radius
COUNT(distinct ?operator) as ?operators
(MAX(?pop)/COUNT(distinct ?operator)) as ?citizensPerOperator
((MAX(?areaTotal) / 1000000)/COUNT(distinct ?operator)) as ?areaPerOperator

WHERE {

  SERVICE <http://dbpedia.org/sparql> {
    # The remote part of the query
       ?city rdf:type dbo:City;
             rdfs:label ?cityName;
             dbo:areaTotal ?areaTotal;
             dbo:country dbr:Germany;
             geo:lat ?lat;
             geo:long ?long;
             dbo:populationTotal ?pop .
             FILTER (lang(?cityName) = ’en’)
             FILTER (?pop > 100000)
    # End of The remote part of the query
  }#End of Service

  SERVICE <http://localhost:8890/sparql> {
    # The remote part of the query
{    ?node rdf:type pgo:Pillar .
    ?node pgo:operator ?operator .
    ?node s:geo ?geoCoordinates .
    ?geoCoordinates s:latitude ?genLat .
    ?geoCoordinates s:longitude ?genLong .
}
     UNION {
?node rdf:type pgo:Substation .
    ?node pgo:operator ?operator .
    ?node s:geo ?geoCoordinates .
    ?geoCoordinates s:latitude ?genLat .
    ?geoCoordinates s:longitude ?genLong .
    }

    # End of The remote part of the query
  }#End of Service



FILTER(
(
bif:acos(bif:sin(xsd:float(?lat) * 3.14159 / 180)*
bif:sin(xsd:float(?genLat) * 3.14159 / 180)+
bif:cos(xsd:float(?lat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLat) * 3.14159 / 180)*
bif:cos(xsd:float(?genLong) * 3.14159 / 180 -
xsd:float(?long) * 3.14159 / 180))*6371000
) < bif:sqrt(?areaTotal / 3.14159) #10000.0 #bif:sqrt(?area / 3.14 ) * 1000 #
) .

}
group by ?cityName
order by DESC(COUNT(distinct ?operator))
LIMIT 10
```

Listing 6. SPARQL to show the Cities (with over 100 000 population) ordered by the number of power operators.

```
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:pgo="http://purl.org/net/hdlipcores/ontology/pgo#"
    xmlns:s="http://schema.org/"
    xmlns:datarepo="http://purl.org/net/hdlipcores/ontology/pgo/data#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <pgo:Link rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#link-175370">
    <pgo:length>458.244020145921</pgo:length>
    <pgo:operator>RWE</pgo:operator>
    <s:name>
    380 kV Hessen Ost; Trebur Nord+ S d ;380kV Trebur S d; 380 kV Trebur Nord;380 kV Hessen Ost
    </s:name>
    <pgo:voltage>380000;380000;380000;380000</pgo:voltage>
    <s:hasPart>
      <pgo:Pillar rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-48536">
        <s:geo>
          <s:GeoCoordinates rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-48536-gs-gc">
            <s:latitude>8.47704075248088</s:latitude>
            <s:longitude>50.0935645616639</s:longitude>
          </s:GeoCoordinates>
        </s:geo>
        <s:name>380 kV Hessen Ost; Trebur Nord| S d; 110 kV Kriftel Nord;380 kV Hessen Ost;
        Trebur Nord+ S d ;110 kV Kriftel Nord+ S d
        </s:name>
        <pgo:operator>RWE</pgo:operator>
        <pgo:hasMeasurements>
          <pgo:Measurements
          rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-48536-measurements">
            <pgo:voltage>380000;110000;380000;110000</pgo:voltage>
            <pgo:frequency>50;50;50</pgo:frequency>
          </pgo:Measurements>
        </pgo:hasMeasurements>
      </pgo:Pillar>
    </s:hasPart>
    <s:hasPart>
      <pgo:Substation rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-18221">
        <s:geo>
          <s:GeoCoordinates
          rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-18221-gs-gc">
            <s:latitude>8.47130677635285</s:latitude>
            <s:longitude>50.0972089406338</s:longitude>
          </s:GeoCoordinates>
        </s:geo>
        <s:name>Umspannwerk Kriftel</s:name>
        <pgo:operator>Amprion</pgo:operator>
        <pgo:hasMeasurements>
          <pgo:Measurements
          rdf:about="http://purl.org/net/hdlipcores/ontology/pgo/data#node-18221-measurements">
            <pgo:voltage>380000;110000</pgo:voltage>
            <pgo:frequency>50</pgo:frequency>
          </pgo:Measurements>
        </pgo:hasMeasurements>
      </pgo:Substation>
    </s:hasPart>
  </pgo:Link>
  </rdf:RDF>
```

Listing 7. System generated RDF description. Using the ontology, we generate appropriate RDF code to describe the power system components.

Accelerated simulation of passive analogue VLSI interconnect on GPUs

Tom J. Kazmierski and Gines Domenech-Asensi

Abstract—Analogue models of VLSI interconnect in complex digital systems pose significant design challenges due to their size, often exceeding thousands of discrete nodes, and their tightly coupled structure. Such models need to be formulated and solved in the analogue domain to ensure high accuracy of interconnect effects such as signal delays and signal-to-signal interference. Currently available design tools are inadequate for simulating such systems due to prohibitive CPU times. This paper presents a technique which takes an advantage of the passive nature of interconnect such that simulations of large interconnect systems can be accelerated by about an order of magnitude compared with equivalent SPICE simulations. The acceleration is possible due to the use of explicit integration of the interconnect state equations were a fast estimate of the maximum allowed step-size is used to guarantee numerical stability. We show that a parallel implementation of the proposed algorithm is straightforward on GPU architectures and compare the results with those obtained from both standard and GPU implementations of SPICE. Several case studies are presented to illustrate the speed of the proposed method and to show that a good match can be obtained between the accuracy of standard SPICE-like simulations and the proposed approach. The proposed method has already been tested in RC interconnect simulations but here, for the first time, we present the performance of our method when applied to RLC interconnect which poses significant restrictions on the maximum allowed simulation step-size necessary to maintain numerical stability,

Index Terms—Simulation acceleration, state-space technique, many-core computer, GPU.

I. INTRODUCTION

Classical SPICE-like simulators used to analyse the behaviour of analogue circuits likerely on the modified nodal analysis and use implicit integration techniques based on the Newton-Raphson linearisation method to solve the circuit analogue equations at each time step. These methods have proven to be reliable and numerically stable, but on the other hand, they lead to long CPU times, often hours or even days and weeks for large circuits. These long simulation times contribute to delays in the design cycle time. The main reason for extensive computation times is due the necessity to build and factorise the Jacobian matrix of the analogue system multiple times at each time step. In contrast to implicit integration methods, the computational workload of explicit integration techniques is lighter. The main disadvantage of explicit methods is the need to limit the step size to ensure numerical stability. In a general case of a non-linear analogue system, where equations are stiff due to the large disparity of time constants, the step-size limitation may be very severe and in such cases implicit methods perform better. However, in the case of VLSI interconnect, equations are not stiff and what is more, interconnect is passive and usually linear which simplifies the state equation formulation for explicit methods, and makes estimates of the maximum allowed step-size easier and therefore faster. Different works have proved that the use of state-space equations combined with explicit integration methods is a suitable technique to speed up transient simulations of many types of analogue circuits [1] or mixed systems [2]. However, given the increasing complexity of analogue circuits and systems, new techniques are required to speed transients simulations, besides the use of alternative integration algorithms. Among these techniques, those based on exploiting the parallelisation of analogue integration methods running on parallel computer architectures are gaining more and more attention in the recent years. The Compute Unified Device Architecture (CUDA) [3] proposed by NVIDIA in 2006, is a programming model that allows engineers to use a high level programming language such as C to develop algorithms for general purpose Graphics Processing Units (GPUs). This has provided design engineers with software tools to use relatively cheap parallel architecture computers which now can perform fast simulations in different types of scientific and engineering applications. Thus, in the last decade, there have been many proposals to accelerate the simulation of analog circuits using GPUs e.g. [4]–[7]. Some works have focused on sparse matrix solvers [8] or LU factorization matrix solver [9]–[11], which have achieved substantial speedups compared with traditional parallel sparse solvers like PARDISO [12] or KLU [13]. NVIDIA also released an official sparse matrix solver, cuSolver [14], but the LU factorization in it is still performed on the CPU rather than GPU.

A common characteristic of these works is that they are focused on the traditional implicit integration methods used in SPICE-like simulators like. Recently, an explicit integration method parallelisable over a many-core processors has been proposed [15]. This method combines space state equations with a fixed-step explicit scheme based on the Adams-Bashforth integration formula to speed up the simulation of passive circuits of a complexity up to 1000 nodes.

In this paper we explore the method proposed in [15] further. We improve our fast estimate of the maximum allowed step size and we apply the technique to the RLC interconnect which poses a harder challenge on the step-size requirements than the RC interconnect considered in [15] . We show that in the case of RLC interconnect our method is not only faster than the

T.J. Kazmierski is with the Department of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom e-mail: tjk@ecs.soton.ac.uk

G. Domenech-Asensi is with the Departmento de Electronica y Tec. de Computadoras, Universidad Politecnica de Cartagena, Cartagena, 30201, Spain e-mail: gines.domenech@upct.es

parallel CU-SPICE by up to one order of magnitude for large systems but it is also significantly more accurate as it avoids spurious numerical ringing which characterises the implicit trapezoidal method used as the default integration scheme in CU-SPICE.

II. STABILITY ANALYSIS

Let (1) be the general state equation of a nonlinear, passive dynamic system:

$$\dot{x}(t) = f(x_t, t); x(0) = x_0 \tag{1}$$

As interconnect models are usually linear, the state equation at time point $t_k, k = 0, 1 \ldots$ can be formulated as:

$$\dot{X}(t_k) = A_k X(t_k) + E e_x \tag{2}$$

where X the vector of N state variable wave-forms, e_x a vector of excitations and A_k and E are coefficient matrices, where A_k is the Jacobian of the linearised model at the time point t_k. As the state equation 1 represents a passive system, the eigenvalues of the Jacobian A_k are guaranteed to have negative real parts, so explicit methods can be applied easily to provide a fast integration process. The step-size in explicit integration must be limited to ensure stability besides controlling the accuracy of the numerical solution. The computation of the maximum allowed step-size requires the computation of the spectral radius of $||A||$, a process for which time-consuming operations such as matrix multiplications and eigenvalue calculations are performed. In this work we take advantage of a recently developed fast method to calculate approximate step-size bounds for stability [2]. Although step sizes obtained using such approximate techniques are expected to be smaller than the maximum allowed step sizes calculated from the exact values of the Jacobians eigenvalues, the advantage of using approximate estimates is speed.

The stability of fixed-step Adams-Bashforth methods is defined by the well known stability plots shown in figure 1, where the values of maximum λh which guarantee stability are plotted in the complex plane. So, while for a first order method the maximum acceptable absolute value of λh is 2, for a fourth order method it is decreased to only 0.3. Thus, although higher order methods are more accurate than lower order ones, they are also more unstable.

Stability becomes harder to achieve when variable-step integration is used. Figure 2 shows a finite-difference grid for a q-order Adams-Bashforth method, where t_k is the current time point, t_{k+1} is the next time point, $P_q(t)$ is the interpolation polynomial of order q, and Δx is the unknown in the integration problem.

Let $h_i = t_{i+1} - t_i$ be the time step between two consecutive time points t_i and t_{i+1}. In a fixed-step integration method, all the h_i values are equal and invariable in time. However, in a variable-step method, the values of h_i are different, and change with time. The expressions for the general variable-step method can be obtained by integrating the divided difference polynomial approximation between the current variable value $x_k \equiv x(t_k)$ and the predicted one $x_{k+1} \equiv x(t_{k+1})$.

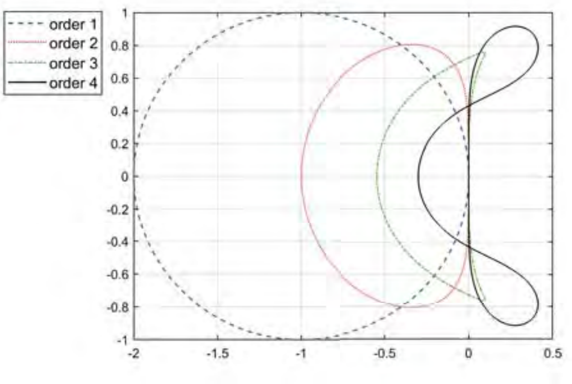


Fig. 1. Stability regions for Adams-Bashforth methods of order 1 to 4.

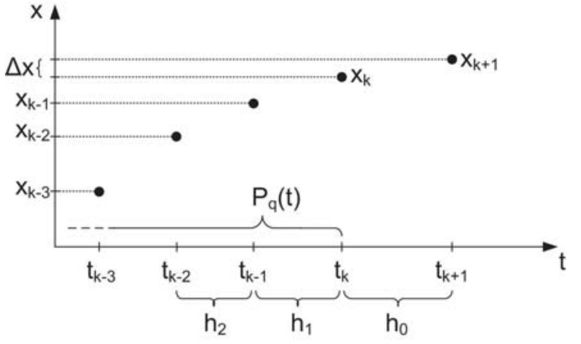


Fig. 2. Finite difference grid for the q^{th}-order Adams-Bashforth method.

$$\begin{aligned} \Delta x = \int_{x_k}^{x_{k+1}} dy = \int_{t_k}^{t_{k+1}} [P_q(t)]_k dt \Rightarrow \\ \Rightarrow x_{k+1} - x_k = \int_{t_k}^{t_{k+1}} \Big(f_0 + (x - x_k) f_k^{(1)} + ... \\ ... + (x - x_k)...(x - x_{k-p}) f_k^{(q)} \Big) dt + O \end{aligned} \tag{3}$$

where $f_k^{(q)}$ is the q^{th} divided difference of function f at t_k [18] and O is the truncation error. Taking as example the second order method, the state variable at time t_{k+1} is computed as:

$$x_{k+1} - x_k = f_k h_0 \left(1 + \frac{h_0}{2h_1}\right) - f_{k-1} h_0 \left(\frac{h_0}{2h_1}\right) \tag{4}$$

being $f_k = f_k^{(0)}$. For the third order method, the following term is added to (4):

$$\begin{aligned} \Big(\frac{t_{k+1}^3 - t_k^3}{3} - (t_k + t_{k-1}) \frac{t_{k+1}^2 - t_k^2}{2} \\ + t_k t_{k-1} (t_{k+1} - t_k) \Big) f_k^{(2)} \end{aligned} \tag{5}$$

Figure 3 shows the stability plots for equations (4) and (5) integrated with different step-size expansion ratios from 1 to 2. The plots show how the variation of the step-size expansion ratio $r = h_{i+1}/h_i$ affects the stability for both the second and

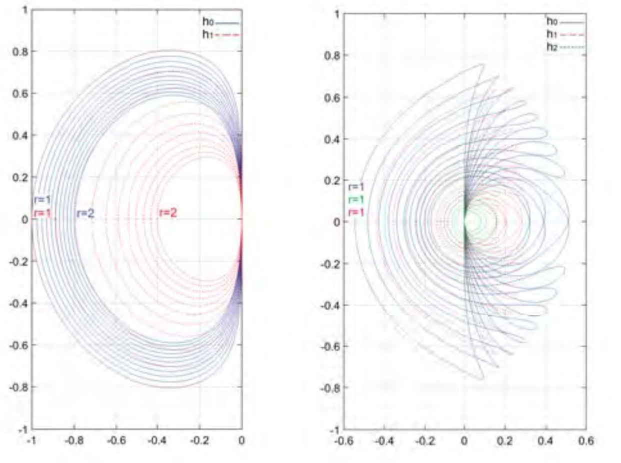


Fig. 3. Stability regions for the second and third order AB methods.

TABLE I
MAXIMUM REAL NEGATIVE VALUE OF THE STABILITY PLOTS

r	1	1.2	1.4	1.6	1.8	2
$L_r(h_0)$	-1	-0.952	-0.909	-0.869	-0.833	-0.8
$L_r(h_1)$	-1	-0.794	-0.649	-0.544	-0.463	-0.4

the third order AB methods. The values of the integration step h_i which guarantee stability decrease as the ratio is increased, being clearly smaller for the third order method. Moreover the stability is more sensitive to the integration step-sizes further from the reference point $tk+1$ than to h_0. So, in order to be able to manage larger values of h_i, in this work the second order variable step integration method has been used. Table I shows the intersection (L_r) of the stability plots with the negative real axis for different values of r for the second order method.

III. FAST APPROXIMATION OF MAXIMUM STEP-SIZE FOR NUMERICAL STABILITY

Given a set of linear ordinary differential equations (ODEs)

$$\dot{X}(t_k) = A_k X(t_k) \tag{6}$$

where A is negative definite and diagonally dominant, the integration method is numerically stable if the integration step size h is

$$h \leq \frac{1}{\max\limits_{r=1,\ldots,N}(\beta_{max}|a_{r,r}|)} \tag{7}$$

where $a_{r,r}$ the diagonal element in row r of A and $\beta_{max} = max(|\beta_0|, \ldots, |\beta_p|)$ the modulus of the maximum coefficient of the pth-order AdamsBashforth formula.

This technique was proposed recently [2] for fast numerical integration of state equations representing many passive systems. In such systems, the proposed fast estimate of the step-size h guarantees stability, but there is a trade off. Step sizes obtained from eq. (7) are expected to be smaller than the maximum allowed step sizes that would be obtained from the exact calculation of the Jacobians eigenvalues.

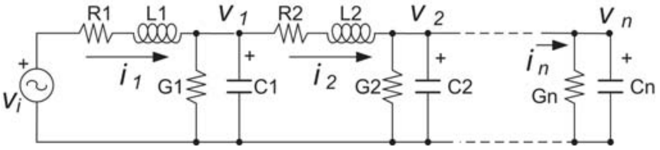


Fig. 4. Transmission line modeled as finite RLC segments.

In the case the Jacobian A is not negative definite but is symmetric, then $\max\limits_{i=1,\ldots,N} \lambda_i = ||A||$ and according to the Gershgorin theorem the values of the eigenvalues are bounded to $\lambda \leq a_{i,i} + \sum_{j=1}^{N} |a_{i,j}|$ for $i \neq j$ [2]. Given that all the Jacobian entries are real numbers, we finally obtain that [20]:

$$h \leq \frac{L_r}{\sum_{j=1}^{N} a_{i,j}} \tag{8}$$

where L_r is the intersection of the stability plot for a given step-ratio r with the negative semi-axis of the complex plane.

IV. INTERCONNECT WITH INDUCTANCE

Figure 4 shows an interconnect modelled as a series of finite RLC segments. Given that the currents through the inductors are state variables, the total number of state variables in an interconnect track is twice the number required for an RC interconnect model.

So, the matrix formulation of the transmission line is given by:

$$\frac{d}{dt}\begin{pmatrix} i_1 \\ v_1 \\ i_2 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = R \begin{pmatrix} i_1 \\ v_1 \\ i_2 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} \frac{1}{L_1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} v_i \tag{9}$$

where the matrix R is:

$$R = \begin{pmatrix} \frac{-R_1}{L_1} & \frac{-1}{L_1} & 0 & 0 & 0 & 0 & \ldots & 0 \\ \frac{1}{C_1} & \frac{-1}{C_1 G_1} & \frac{-1}{C_2} & 0 & 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{L_2} & \frac{-R_2}{L_2} & \frac{-1}{L_2} & 0 & 0 & \ldots & 0 \\ 0 & 0 & \frac{1}{C_1} & \frac{-1}{C_2 G_2} & \frac{-1}{C_2} & 0 & \ldots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & \frac{1}{C_n} & \frac{-1}{C_n G_n} \end{pmatrix} \tag{10}$$

We have performed a number of simulation tests for RLC tracks of different lengths using the proposed method and CU-SPICE, the CUDA version of SPICE [19], using the following component values per discrete section: $C = 1fF$, $L = 100pH$, $R = 10\Omega$, $G = 400\Omega^{-1}$. The excitation was a 1V step and the responses in the first RLC segment are shown in figures 5 and 6 for CU-SPICE and our method respectively.

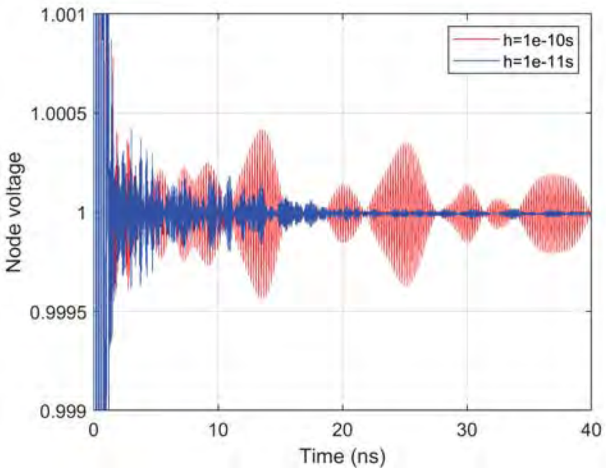


Fig. 5. SPICE simulation results for the RLC interconnect line in fig 4.

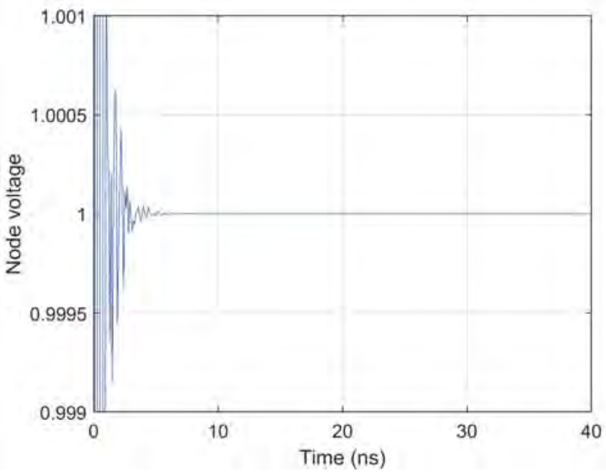


Fig. 6. Simulation results using the proposed technique for the RLC interconnect line in figure 4.

The step-size in the explicit integration was 10^{-13}sec and CU-SPICE simulations were performed using two different step-size limits: 10^{-10}sec and 10^{-11}sec. The spurious numerical ringing is evident in the CU-SPICE results and is absent, after the initial transient process, from the results obtained by the proposed method. In table II we show the CPU times for both methods, the proposed method with the step size of 10^{-13} sec and CU-SPICE with the step-size of 10^{-11}. CU-SPICE, despite using a step size larger by two orders of magnitude than that of our method, is significantly slower with the proposed method reaching a speed up of almost an order of magnitude for 10,000 RLC segments.

V. CONCLUSION

This paper shows promising results obtained when solving equations of large VLSI interconnect by means of explicit integration and state equations. The inclusion of analogue interconnect models in digital VLSI simulations is increasingly important as clock frequencies reach 10 GHz and more. At such speeds the analogue transients which occur in the interconnect cannot be ignored. As interconnect may constitute a large part of an analogue or a mixed-signal VLSI system, the presented method, which accelerates simulations on GPUs can be a useful approach in the development of modern VLSI design tools. For the first time, we present here results of applying the proposed method to interconnect with inductance. These results are preliminary, as this is still work in progress. The presented results are merely meant to illustrate the potential of explicit integration in the solution of vast numbers of equations representing VLSI interconnect.

TABLE II
GPU SIMULATION TIME FOR PROPOSED EXPLICIT METHOD AND CUSPICE

Segments	Explicit (s)	Implicit (s) CUspice	Speedup
100	80.294	40.537	0.504
200	85.387	48.562	0.568
500	98.019	64.159	0.654
1000	96.768	101.833	1.052
2000	100.729	196.108	1.946
5000	100.522	619.070	6.158
10000	123.041	1068.640	8.685

VI. ACKNOWLEDGEMENTS

This work has been partially funded by Spanish government through project RTI2018-097088-B-C33 (MINECO/FEDER, UE) and by EPSRC (the UK Engineering and Physical Sciences Research Council) under grant EP/N0317681/1. The research stay at University of Southampton (UK) has been supported by by Fundacin Sneca-Agencia de Ciencia y Tecnologa de la Regin de Murcia, Programa Regional de Movilidad, Colaboracin e Intercambio de Conocimiento Jimenez de la Espada under grant 21187/EE/19.

REFERENCES

[1] K. C. A. Lam and M. Zwolinski. ”Circuit simulation using state space equations” in Proc. 9th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME). Villach. 2013. pp. 177-180.

[2] T. J. Kazmierski. L. Wang. B. M. Al-Hashimi and G. V. Merrett. ”An Explicit Linearized State-Space Technique for Accelerated Simulation of Electromagnetic Vibration Energy Harvesters.” IEEE Trans. on Computer-Aided Design of Integ. Circ. and Syst. vol. 31, pp. 522-531. Apr. 2012.

[3] NVIDIA. CUDA C Programming Guide Version 7.0. Accessed: Mar. 5, 2018. [Online]. Available: http://docs.nvidia.com/cuda/cudac-programming-guide/

[4] K. Gulati, J. F. Croix, S. P. Khatri and R. Shastry, ”Fast circuit simulation on graphics processing units,” in Proc. Asia and South Pacific Design Automation Conference, Yokohama, 2009, pp. 403-408.

[5] R. E. Poore, ”GPU-accelerated time-domain circuit simulation” in Proc. IEEE Custom Integrated Circuits Conference, Rome, 2009, pp. 629-632.

[6] L. Han and Z. Feng, ”TinySPICE Plus: Scaling up statistical SPICE simulations on GPU leveraging shared-memory based sparse matrix solution techniques,” in Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD), Austin, TX, 2016, pp. 1-6.

[7] E. Schneider, M. A. Kochte, S. Holst, X. Wen and H. Wunderlich, ”GPU-Accelerated Simulation of Small Delay Faults,” IEEE Trans. on Comp.-Aided Design of Integ. Circ. and Syst., vol. 36, no. 5, pp. 829-841, May 2017.

[8] Y. Liang, W. T. Tang, R. Zhao, M. Lu, H. P. Huynh and R. S. M. Goh, ”Scale-Free Sparse Matrix-Vector Multiplication on Many-Core Architectures,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 36, no. 12, pp. 2106-2119, Dec. 2017.

[9] X. Chen, L. Ren, Y. Wang and H. Yang, ”GPU-Accelerated Sparse LU Factorization for Circuit Simulation with Performance Modeling” IEEE Trans. on Parallel and Distr. Syst., vol. 26, pp. 786-795, Mar. 2015.

[10] K. He, S. X. -. Tan, H. Wang and G. Shi, ”GPU-Accelerated Parallel Sparse LU Factorization Method for Fast Circuit Analysis,” IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 24, pp. 1140-1150, Mar. 2016.

[11] W. Lee, R. Achar and M. S. Nakhla, ”Dynamic GPU Parallel Sparse LU Factorization for Fast Circuit Simulation,” IEEE Transactions on Very Large Scale Integration Systems, vol. 26, pp. 2518-2529, Nov. 2018.

[12] O. Schenk and K. Gartner, Solving unsymmetric sparse systems of linear equations with PARDISO, Future Generat. Comput. Syst., vol. 20, no. 3, pp. 475487, Apr. 2004.

[13] T. A. Davis and E. P. Natarajan, Algorithm 907: KLU, a direct sparse solver for circuit simulation problems, ACM Trans. Math. Softw., vol. 37, no. 3, Sep. 2010, Art. no. 36.

[14] Cusolver Library, document DU-06709-001 v9.0, Jun. 2017.

[15] G. Domenech-Asensi and T. J. Kazmierski, ”An efficient numerical solution technique for VLSI interconnect equations on many-core processors”, in Proc. IEEE Intl. Symp. on Circ. and Systems, Sapporo, Japan, 2019.

[16] Ngspice Circuit Simulator. Accessed: May. 24, 2019. [Online]. Available: www.ngspice.org

[17] L. O. Chua and P. Y. Lin. Computer-Aided Analysis of Electronic Circuits: Algorithms and Computational Techniques. Englewood Cliffs, NJ: Prentice-Hall. 1975.

[18] J. D. Hoffman, Numerical Methods for Engineers and Sciencists 2nd Ed.,CRC Press, 2001.

[19] F. Lannutti, F. Menichelli, M. Olivieri ”CUSPICE The revolutionary NGSPICE on CUDA Platforms”, in Proc. 12th MOS-AK Workshop at the ESSDERC/ESSCIRC Conference, Venice, Italy, 2014.

[20] G. Domenech-Asensi and T. J. Kazmierski, ”Stability and efficiency of explicit integration in interconnect analysis on GPUs”, accepted for publication in Proc. IEEE Intl. Symp. on Circ. and Systems, Sevilla, Spain, 2020.

Radiation impact modelling method for gate level design verification

Artur Petrosyan

Abstract – A radiation impact modelling method is presented in this paper. The method is aimed to be used in gate level design verification. A software tool was designed based on the method to configure the radiation impact parameters and automatically inject the radiation-induced single event upset and single event transient effects in the synthesized gate-level netlist. The work aims to provide a new method to speed up the design process of radiation-hardened digital integrated circuits. Experimental results of the proposed method show accuracy loss about 3-4% and simulation speed-up by 2-3 times compared to existing methods. The proposed method effectively verifies the design resistance to radiation impact.

Keywords - Single event upset (SEU), Single event transient (SET), Radiation impact (RI), Synthesis, Radiation modelling.

I. INTRODUCTION

Radiation impact (RI) on integrated circuits (IC) forms effects such as single event transient (SET) and single event upset (SEU) which lead to failures (soft errors) in the combinational and sequential parts of ICs. It is known that there are three general masking factors (logical, electrical and timing) which may filter out the influence of the SET or SEU effect [1]. The influence is logically masked (filtered) when one of the inputs of the logic gate is the dominant, the influence is electrically masked when the RI does not form enough energy to assert a SET or SEU, and the RI is timing masked if the SET occurs when there is no latching window for the memory elements [2]. However, these masking factors do not grant that the IC will work without failures.

Testing and verification of the ICs with consideration of discussed effects is becoming more important as the safety requirements for ICs used in automotive systems and other spheres are getting higher [3], [4], [5].

Different RI models have been proposed so far. The methods proposed in [6] uses device-level simulations to analyse the RI. As an alternative to the device level simulations, the method proposed in [7] uses a double exponential current pulse in circuit-level simulators for modelling the RI. However, these modelling techniques of RI have poor performance when they are applied to large designs. The technique proposed in [8] partly solves the performance of RI modelling in large designs. This method uses Verilog gate level models to modify the entire logic cell library to characterize the Multiple-Event-Transient (MET) in gate level simulation. The main disadvantages of the method are time-consumption and inefficient insertion of MET effects in all the cells in the library, as, the probability of RI occurrence on all the cells in design is low [9]. Another disadvantage of this method is that it considers simulations of only SET effects.

Thus, new methods are needed in gate level simulations to speed-up the RI simulations and model not only the SETs but also the SEUs.

II. SET AND SEU GATE LEVEL MODELS

The SET and SEU radiation effects are random in nature therefore these effects occur at unpredictable moments on random locations [4]. Two Verilog models are suggested for gate level simulation. The models consider the randomness of the effects.

A. SET model

The model (Fig. 1) contains a random pulse generator, a random number generator, logic to randomly select a cell, and an XOR gate.

The random pulse generator is used for generating *N* number of pulses during a period to make the occurrence of SET random. The other random number generator is used to make the pulse widths of the SET random. To apply the generated pulse (SET effect), a cell is selected from the gate level netlist. The output of the cell and the pulse signal are connected to the XOR gate. If the SETs propagate through combinational circuits with influence of masking factors no failures will be detected.

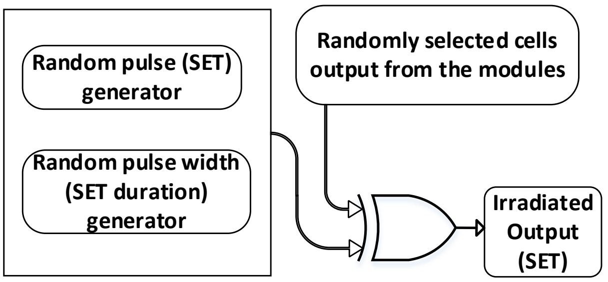


Fig. 1. Block diagram of the SET model

B. SEU model

To make verification process more controllable and make sure that masking factors don't affect error injection

Artur Petrosyan is with Synopsys Armenia Educational Department Synopsys Armenia CJSC, 41 Arshakunyats Ave, Yerevan, Armenia, E-mail: arturp@synopsys.com

process, a model is suggested to test only sequential circuits by inserting SEU directly on the outputs of elements. The Verilog model (Fig. 2) uses an inverter to flip the state of sequential element. The inverter is connected to the output "Q" port.

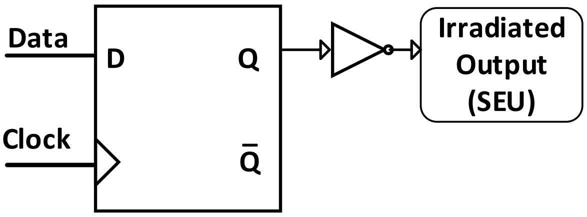


Fig. 2. Block diagram of the SEU model

III. THE PROPOSED METHOD FOR RADIATION MODELLING IN GATE LEVEL SIMULATION

A general overview of the proposed method is shown in the flowchart (Fig. 3), where the first step is the logic synthesis of the RTL design. After this step, the gate-level (synthesized) netlist is generated. The generated netlist is then simulated to save the waveforms as golden results for further comparison and estimation of the RI.

The next step is finding the module and cell instances by passing through the gate-level netlist. The information such as module and cell instances are saved in vectors. For example, module instances are saved in modules vector Eq. (1) and the cell instances are stored in cells vector Eq. (2).

$$M = (m_1, m_2, .., m_i) \qquad (1)$$

$$C = (c_1, c_2, .., c_i) \qquad (2)$$

where i is the number of the cell or module instances in the given netlist.

In gate-level netlists, the module instances contain cells. Thus, the cell vectors are related to the modules. Each module instance vector consists of cell instances vector Eqs. (3), (4) and (5).

$$m_1 = C_1(c_1, c_2, .., c_i) \qquad (3)$$

$$m_2 = C_2(c_1, c_2, .., c_i) \qquad (4)$$

$$m_n = C_n(c_1, c_2, .., c_i) \qquad (5)$$

where n - is the amount of the logic cells in each module instance.

The fourth step is calculating the probability of RI occurrence in a design. This allows modelling the unpredictable RI behaviour (discussed in previous section) and to avoid applying the RI effects on all the cells in the design.

The probability is calculated using exponential density function Eq. (6) [10], [11] for each module in module instances vector and for each cell in cell instances vector.

$$p(t) = \frac{1}{\tau} e^{-t/\tau} \qquad F(t) = \int_0^t p(t)dt = 1 - e^{-t/\tau} \qquad (6)$$

where t – is the SEU or SET assertion period τ – is the time after the assertion of the effects.

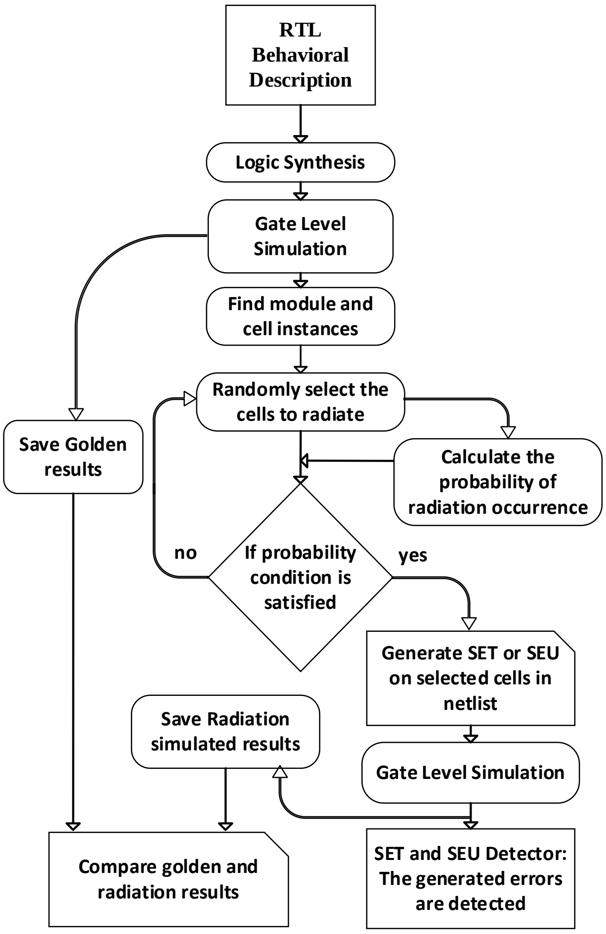


Fig. 3. General overview of the radiation effect (SET, SEU) modelling method

The fifth step is the generation of the SET or SEU effects on module and cell instances vectors.

From the module instance lists the highest RI probability elements are randomly selected. Then based on the type of the instances (combinational or sequential) the SET or SEU (Figs. 1 and 2) models are applied. In order to apply the effects, the selected cell is replaced in the gate level netlist with a modified (radiation effect injected) model.

Next, the gate level simulation of the modified netlist is performed to save the radiation simulated waveforms. During the simulation, the detected errors are extracted using SET and SEU error extractor (Fig. 4).

The error detection and extraction are developed for further error analysis. It is implemented by a separately generated Verilog module for the gate-level netlist. This module counts errors on each SET or SEU occurrence. It also detects the time/moment of each error and extracts its value to help to debug and analyse the results (Table I).

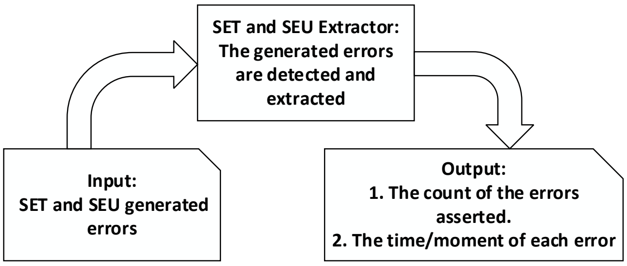


Fig. 4. Block diagram of radiation effect (SET or SEU) extractor

Result analyses are performed by comparing the saved golden simulation waveforms of the gate-level netlist without taking into account RI and the simulation waveforms with consideration of RI on the generated netlist.

IV. EXPERIMENTAL RESULTS

The proposed method was implemented as a software tool in the digital design flow. The tool automates the proposed method (Fig. 3) and allows configuration of the RI parameters. A number of ISCAS89 benchmark circuits [12] were implemented in Verilog and tested using VCS environment [13] to determine the error rates and perform design modification to protect it against RI.

To generate the gate-level netlist synthesis of the s27 benchmark circuits was done using Synopsys Design Compiler [14] in SAED32nm technology [15].

The VCS simulation waveform results of synthesized netlist are saved as golden results (Fig. 5).

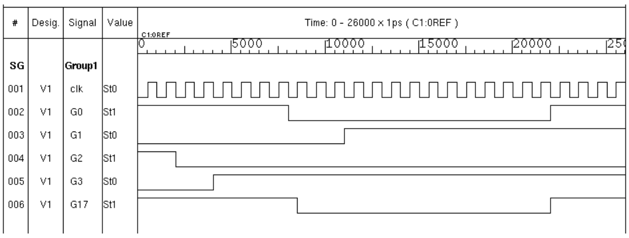


Fig. 5. VCS simulation results of ISCAS89 S27 benchmark circuit without RI as golden results

The synthesized netlist is then used to apply modification to inject SET or SEU using the RI models (Figs. 1 and 2) and the proposed method (Fig. 3). The VCS simulation waveform with RI models is saved (Fig. 6) and errors are extracted (Table I) using the SET (or SET and SEU) error extractor module.

The highlighted signals in Fig. 6 indicate the error which occurred on the output of s27 benchmark circuit with clock capturing edge.

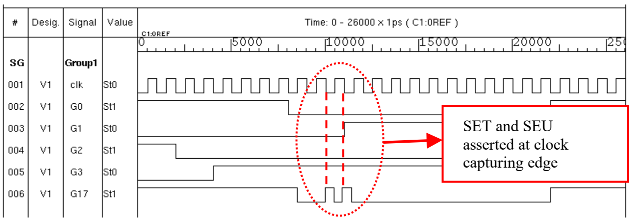


Fig. 6. VCS simulation results of ISCAS89 S27 benchmark circuit with RI

TABLE I
DETECTED ERRORS FOR ISCAS89 s27 BENCHMARK CIRCUIT

Cell Type	RI Type	RI Period(*ps*)	Total RI error count	Total RI reached output
NAND	SET	9958-10434	14	2
NOR	SET	9958-10434		
DFF	SEU	10862-11386		

The result analysis by comparing the golden and modified gate-level netlist waveforms is performed using Synopsys Verdi tool [16]. This tool shows all the differences between the golden and irradiated simulations and allows to debug the design.

A waveform comparison result (Fig. 7) of the ISCAS89 s27 benchmark circuit is done. Signals from the golden waveform are compared to the corresponding signal in the irradiated waveform. The DFF state change is detected on the clock falling edge when the RI occurred. Without the RI the DFF stores logic ‘1’. However, instead of the logic ‘1’ a logic ‘0’ appears at the output of the DFF cell. Which results in SET propagation through the entire design to its output.

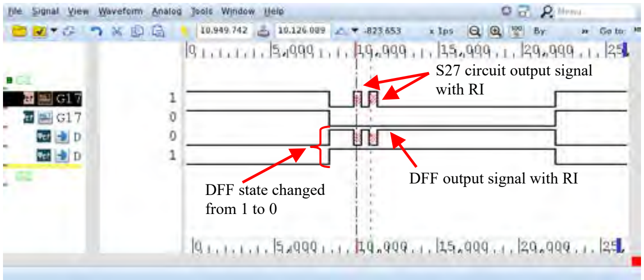


Fig. 7. The comparison results of S27 benchmark circuit with RI and without RI

To verify the proposed method several benchmark circuits from ISCAS89 series were simulated. The same benchmarking circuits are tested with HSPICE simulator [17] and the MET simulation method proposed in [8]. The simulation results show that the accuracy of the proposed method is 3-4% less compared to the HSPICE simulations due to electrical masking factors but it is similar compared to MET simulation method.

From the other hand, the results of CPU runtime (Table III) show that simulation time of the circuits is about 2-3 times faster using the proposed methodology.

TABLE II
COMPARISON RESULTS OF THE PROPOSED METHOD ACCURACY FOR THE ISCAS89 BENCHMARK CIRCUITS COMPARED TO EXISTING METHODS

ISCAS89 circuits	Total RI error count	Total RI reached output		
		HSPICE [17]	MET simulation method [8]	Proposed method
s27	14	3	2	2
s298	28	14	13	12
s344	132	54	50	52
s953	675	86	76	82

TABLE III
SIMULATION RUNTIME COMPARISON RESULTS FOR THE ISCAS89 BENCHMARK CIRCUITS

ISCAS89 circuits	Total RI error count	HSPICE [17] CPU(Sec)	MET simulation method [8] CPU (Sec)	Proposed method CPU (Sec)
s27	14	22,8	7,35	4,15
s298	28	28,4	15,23	6,4
s344	132	52,36	28,9	10,45
s953	675	73,22	35,7	15,35

V. CONCLUSION

The radiation impact modelling and simulation method in gate level design verification is presented. The method was automated as a software tool and tested using the ISCAS89 benchmark circuits. The experimental results show that the runtime of the simulations of digital circuits considering radiation impact is reduced about 2-3 times while the accuracy loss is about 3-4%. The results prove that the proposed method can be integrated into the digital circuit design process for gate level design verification.

REFERENCES

[1] Mariem Slimani, Lirida Naviner, “A tool for transient fault analysis in combinational circuits”, 2015 *IEEE International Conference on Electronics, Circuits, and Systems (ICECS),* pp. 125 - 128, Dec. 2015.

[2] P. Shivakumar, M. Kistler, S.W. Keckler, D. Burger, L. Alvisi, “Modeling the effect of technology trends on the soft error rate of combinational logic”, *Proceedings International Conference on Dependable Systems and Networks*, pp. 427-430, June 2002.

[3] Vazgen Melikyan, “Simulation and Optimization of Digital Circuits”, *Springer International Publishing,* pp. 10-77, Jan. 2018.

[4] Ghaith Bany Hamad, Syed Rafay Hasan, Otmane Ait Mohamed, “Modeling, analyzing, and abstracting single event transient propagation at gate level”, *2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 515-518, Aug. 2014.

[5] Fei Gong, Meenal Vaidya, Rishvanth Kora, Daniel Harshbarger, “A FPGA based prototype verification in automotive mixed signal integrated circuit development”, *2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 1200-1203, Dec. 2013.

[6] Paul E. Dodd, “Physics-based simulation of single-event effects”, *IEEE Transactions on Device and Materials Reliability,* Vol. 5, No. 3, pp. 515-518, Sept. 2005.

[7] William G. Bennett, Nicholas C. Hooten, “Experimental Characterization of Radiation-Induced Charge Sharing”, *IEEE Transactions on Nuclear Science,* Vol. 60, No. 6, pp. 4159 - 4165, Dec. 2013.

[8] Akira Mochizuki, Naoya Onizawa, Akira Tamakoshi, Takahiro Hanyu, “Multiple-event-transient soft-error gate-level simulator for harsh radiation environments”, *TENCON 2015 - 2015 IEEE Region 10 Conference,* Vol. 60, No. 6, pp. 978 - 983, Jan. 2016.

[9] T. Rejimon, S. Bhanja, “A Timing-Aware Probabilistic Model for Single-Event-Upset Analysis”, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 14, No. 10, pp. 1130 - 1139, Oct. 2006.

[10] Grecki, Mariusz, “VHDL Simulation considering Single Event Upsets (SEUs)”, *Technical Proceedings of the 2006 NSTI Nanotechnology Conference and Trade Show*, pp. 717 - 720, Jan. 2006.

[11] Marta Bagatin, Simone Gerardin, Alessandro Paccagnella, “Single Event Upsets Induced by Direct Ionization from Low-Energy Protons in Floating Gate Cells”, *IEEE Transactions on Nuclear Science,* Vol. 64, No. 1, pp. 464 - 470, Dec. 2017.

[12] E. Brglez. D. Bryan, K. Kouninski., “Combinational profiles of sequential benchmark circuits” ISCAS89 Proceedings”, *ISCAS89 Proceedings*, pp. 124 - 198, Dec. 1989.

[13] VCS® User Guide, *Synopsys Inc.* – 2019. pp. 56-254

[14] Design Compiler User Guide, *Synopsys Inc.* – 2019. pp. 138-175

[15] R. Goldman, K. Bartleson, T. Wood, K. Kranen, C. Cao, V. Melikyan, G. Markosyan, “Synopsys' open educational design kit: Capabilities, deployment and future”, *2009 IEEE International Conference on Microelectronic Systems Education*, pp. 20 - 24, Sep. 2009.

[16] Verdi® SystemC Debug User Guide, *Synopsys Inc.* – 2017. pp. 24-76.

[17] HSPICE Application Manual, *Synopsys Inc.*, 2018. pp. 617-712.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Game-based Teaching Approach of Accuracy and Precision

Zivko Kokolanski,
Mare Srbinovska, and Vladimir Dimcev

Abstract – Accuracy and precision are one of the bacis static characteristics of any instrument or sensor system. However, this terms are often interchangeable used and not well understood by the students, regardsles of their different meaning. This paper aims to fill this gap by proposing a game-based approach to teach students the nature of accuracy and precision. The game is realized in LabVIEW environment and supports two gaming scenarios: simulated controls, and realistic controls. The game idea and implementation is described and experimental results are given. It has ben shown that the teaching process and the quality of learning can be improved by involving such interesting and intuitive software and hardware tools.

Keywords – Accuracy, precision, game-based learning, virtual instrument

I. INTRODUCTION

Accuracy and precision are the two important factors when taking measurement data. Both accuracy and precision refer to how close a measurement is to an actual value, but accuracy means how close a measurement is to a known or accepted value, while precision denotes how reproducible measurements are, even if they are far from the accepted value.

The meaning of accuracy and precision can be defined in terms of hitting a bull's-eye. Accurately hitting the target means close to the center of the target, even if all the marks are on different sides of the center.

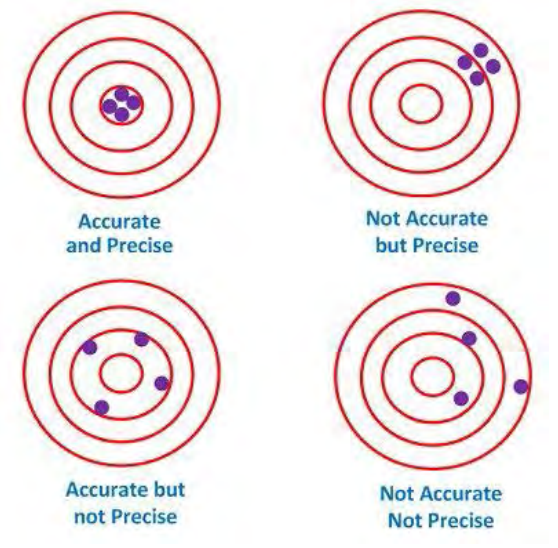


Fig. 1. Examples of accuracy and precision

Precisely hitting a target means all the hits are closely spaced, even if they are very far from the center of the target. Measurements that are both precise and accurate are repeatable and very near true values. The examples on Fig. 1 show four possible combinations of accuracy and precision. The target at the upper lefrt figure shows both accuracy and precision as the shots are clustered together (they are precise) in the target's center-most ring (they are accurate). The next example (right) shows results that are precise, but inaccurate because they are at the target's outer edge instead of its center. The third example is considered accurate because the four shots cluster around the target's center, but they are not precise because the individual shots are quite far apart from each other. The final example shows a dispersion of shots that is both inaccurate and imprecise. Note that the average for a set of measurements may be accurate even if the individual measurements deviate significantly from the desired or theoretical value.

In measurent technique the accuracy is the ability of the instrument to measure accurate value. Or according to the International Organization for Standardization (ISO) definition [1], accuracy is a level of measurement that yields true (no systemic errors) and consistent (no random errors) results.

The term precision means two or more values of the measurement are closed to each other. The value of precision differs because of the observational error. The precision is used for finding the consistency or reproducibility of the measurement. The conformity and the number of significant figures are the characteristics of the precision.

Game-based learning [5-7] aims at students to learn about a certain subject or to acquire certain skills by means of playing. This teaching approach must fulfill both recreational and didactic goals in order to succeed. This paper proposes a game-based learning approach intended to teach students the difference and the meaning of accuracy and precision.

II. ACCURACY VS PRECISION

There are two common definitions of accuracy. The more common definition is that accuracy is a level of measurement with no inherent limitation (ie. free of

Zivko Kokolanski, Mare Srbinovska, Vladimir Dimcev Ss. Cyril and Methodius University of Skopje, Faculty of Electrical Engineering and Information Technologies, Rugjer Boskovic 18, 1000 Skopje, E-mail: kokolanski@feit.ukim.edu.mk

systemic error, another form of observational error).

In science, engineering and math, accuracy refers to how close a measurement is to the true value [2].

The precision of a measurement system, related to reproducibility and repeatability, is the degree to which repeated measurements under unchanged conditions show the same results [2, 3].

The ISO standard applies a more rigid definition, where accuracy refers to a measurement with both true and consistent results.

A measurement system can be accurate but not precise, precise but not accurate, neither, or both. For example, if an experiment contains a systematic error, then increasing the sample size generally increases precision but does not improve accuracy. The result would be a consistent yet inaccurate string of results from the flawed experiment. Eliminating the systematic error improves accuracy but does not change precision. A measurement system is considered valid if it is both accurate and precise.

In addition to accuracy and precision, measurements may also have a measurement resolution, which is the smallest change in the underlying physical quantity that produces a response in the measurement. In numerical analysis, accuracy is also the nearness of a calculation to the true value; while precision is the resolution of the representation, typically defined by the number of decimal or binary digits.

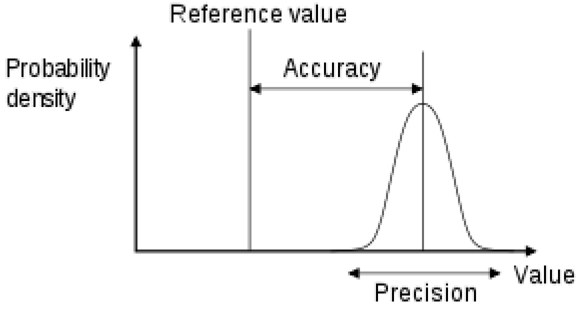


Fig. 2. Accuracy is the proximity of measurement results to the true value, precision, the repatibility or reproducibility of the measurement [4]

According to ISO 5725-1 [1] the general term "accuracy" is used to describe the closeness of a measurement to the true value and precision to repeatability or reproducibility of the measurement, as it is shown in Fig. 2.

When the term is applied to sets of measurements of the same measurand, it involves a component of random error and a component of systematic error. In this case trueness is the closeness of the mean of a set of measurement results to the actual (true) value and precision is the closeness of agreement among a set of results.

III. GAME IDEA AND REALIZATION

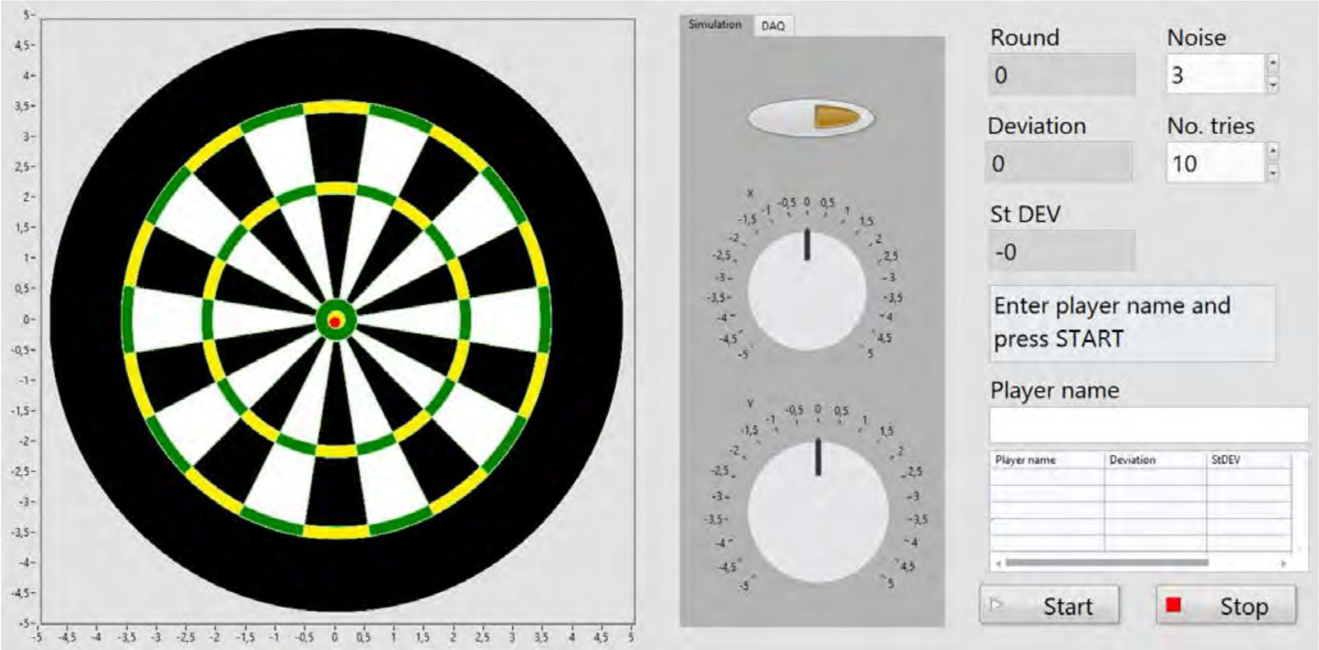
The game idea is bazed on the concept explained in Fig.1, i.e playing a shooting game where the player's performance metrics are described through accuracy and precision. The main goal is to provide a competitive scenario for the students where they can be ranked regarding the geometrical distance from the center of the board (accuracy) and regarding the dispersion of their results (precision). The game is realized in Laboratory Virtual Instrumentation Engineering Workbench (LabVIEW) platform. The front panel and the block-diagram of the virtual instrument are given in Fig.3 and Fig.4 respectively.

When playing the game, the player should try to shoot as close as possible, and as consistent as possible to the center of the board. The game realization is obtained by a red laser spot that can be moved with certain controls. The game supports two modes of playing: simulated or realistic controls. In the first mode, the controls for horizontal and vertical deflection of the laser spot are simulated in LabVIEW, whereas in the second mode the player uses a joystick to move the laser spot. The firing is realized through simulated or realistic digital switch in the simulated or realistic game mode respectively. In the "realistic" mode, the virtual instrument uses a Data Acquisition (DAQ) card to read the joystick and the digital switch status.

At the beginning of the game, the player must select the game mode (simulated or DAQ) and enter it's personal name. Moreover, the user adjusts the number of tries (shots) and the amount of noise (difficulty). Afterwards, the game starts by pressing the digital control "Start". The user now has a predified number of shots to finish the game. However, each time the player shoots, the amount of noise to the laser spot is automatically increased. Therefore, as the shooting progresses, the game becomes more difficult to play. When the predefined number of shots has been reached, the virtual instrument automatically calculates the standard deviation and the geometric mean. The standard deviation is then used as a diameter of a circle which is drawn around the geometrical center of the shots. At the end, the results for the particular player are presented numerically (in a table) and graphically by exporting the target screenshot in Microsoft Word format. The game continues by entering the personal details of the second player and repeating the described procedure all over again.

In order the game to fulfill the didactic goals, it is essential to have more players. The more players are involved, the probability to have more diverse results increases. As the number of results increase, one will eventually obtain samples where the better/worse player accuracy or precision is obvious.

The virtual instrument is realized by using a state-machine programming architecture. The program can be in one of the following five states: *initialize*, *start*, *play*, *end* and *presentation*. In the *initialize* state, the virtual instrument performs reinitializations of all controls and indicators to their default values and immediately enters the *start* state. Now the program wait for player scenario selection.



Round
0

Noise
3

Deviation
0

No. tries
10

St DEV
-0

Enter player name and press START

Player name

Player name	Deviation	StDEV

Fig. 3. Front panel of the virtual instrument used for implementing a game-based scenario for accuracy vs precision

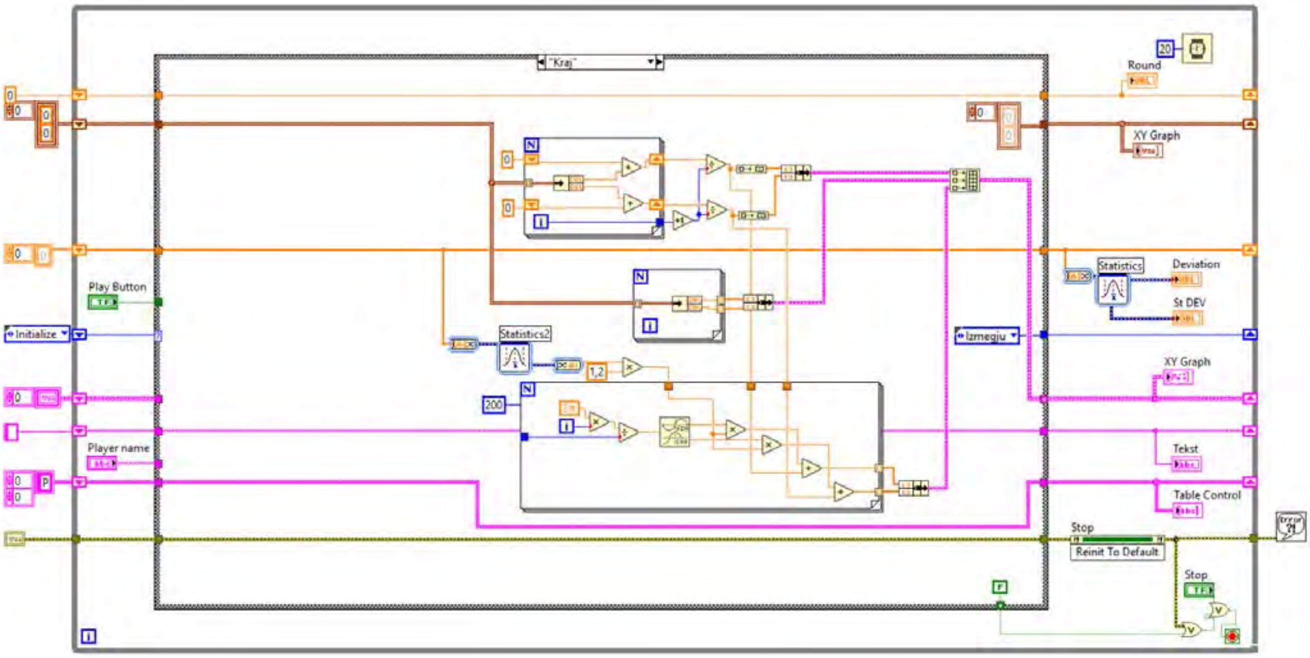
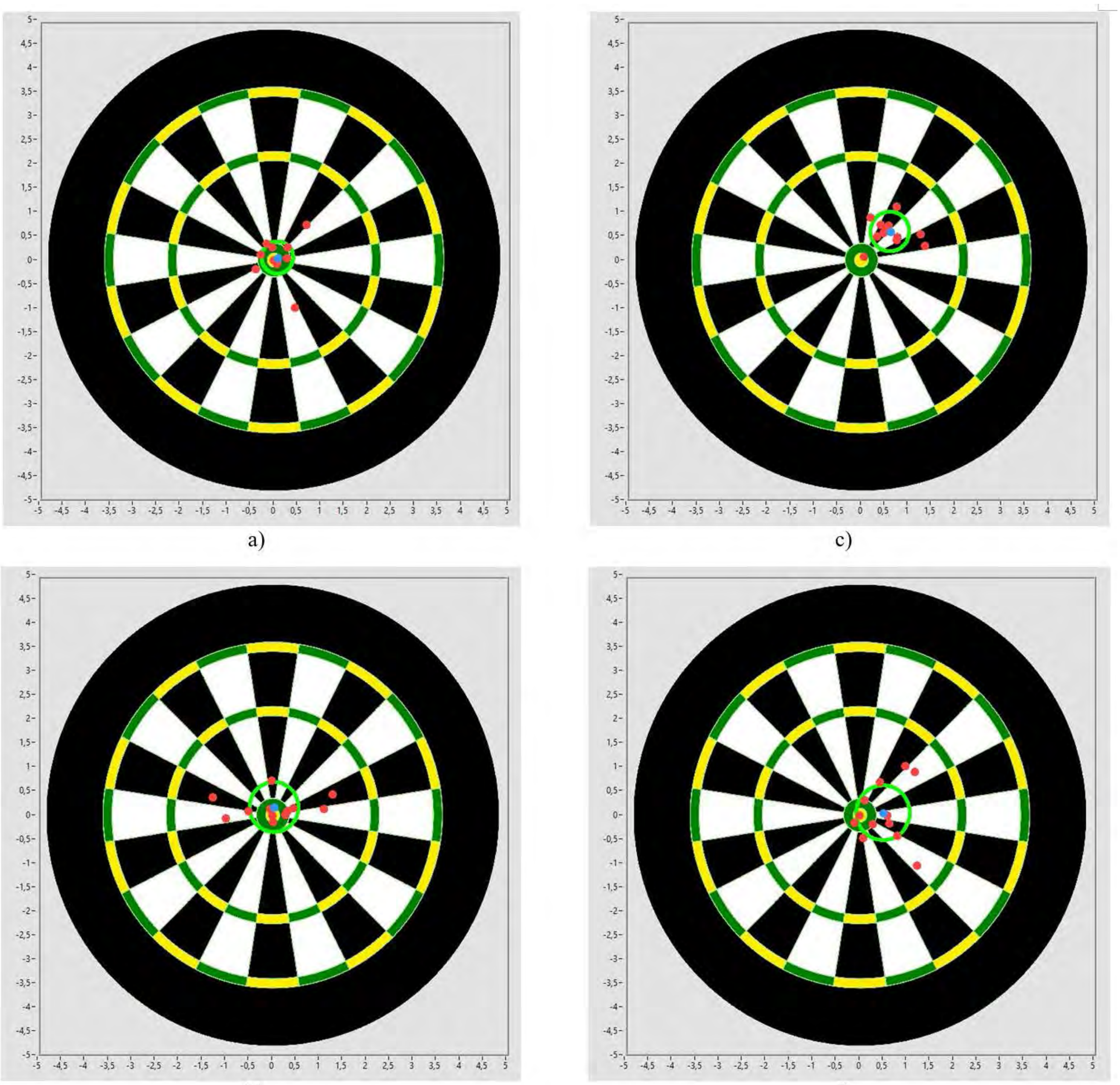


Fig. 4. Block diagram of the virtual instrument (End state)

The user must define the following parameters: game mode (simulation or DAQ), number of tries, noise amplitude and enter personal name. When all data has being defined, the program enters the “Play” state by pressing the *Start* button. Now the virtual instrument initializes the DAQ card channels (in case of DAQ mode) and continually reads the horizontal and vertical controls. In the meantime, the program stores and counts each player shot until the predefined number of tries has being reached. Then, the program enters the *End* state given in Fig. 4. In the *end* state, the program calculates the mean and the standard deviation of the shots and plots all results on the target. Besides each particular shot, the program draws circle whose center is defined by the geometrical mean of the samples, whereas the diameter is defined by the samples standard deviation. The results are simultaneously represented graphically and numerically. Moreover, the target is automatically exported in Word format in the *presentataion* state. Finally, the program is restarted to the *initialize* state and all steps are repeated with the next player. Each player can stop the program at any time by pressing the *Stop* button.



a) c)

b) d)

Fig. 5. Experimental results a) Good accuracy and precision, b) Good accuracy bad precision, c) Good precision bad accuracy, d) Bad accuracy and bad precision

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The game-based approach proposed in this paper was implemented in the regular teaching activities of the course Electrical Measurements at the Faculty of Electrical Engineering and Information Technologies in Skopje. Ten students took part in the game experiment and interpretation of the results. Four of the results, given in Fig.5, that best explain the accuracy and precision concept were selected. The Table I summarizes the calculated standard deviation and absolute error of the selected games. The higher is the absolute error, the worse the accuracy is. Similarly, higher standard deviation suggests worsen precision.

TABLE I
EXPERIMENTAL RESULTS REGARDING ACCURACY AND PRECISION

Player name	Absolute error	Standard deviation
Player a)	0,281	0,326
Player b)	0,294	0,433
Player c)	0,868	0,344
Player d)	0,659	0,474

The analyses of the results reported in Fig.5 and in Table I suggests that players a) and b) were more accurate comparing to the players c) and d). This can be confirmed in two ways: graphically (from observing the results given in Fig.5), and numerically from Table I. It can be clearly seen that the geometrical mean (blue dot) for players a) and b) is closer to the target center. Consequently, the absolute errors of the shots for the players a) and b) are lower than those of the players c) and d). On the other hand, it is obvious that the diameter of the circles and the standard deviations for the players a) and c) is lower comparing to the players b) and d). This suggests that players a) and c) have better precision.

V. CONCLUSION

The goal of this paper is to describe a game-based method for teaching the difference between accuracy and precision. The game was realized in LabVIEW and was implemented in the regular teaching process of the course Electrical Measurements. The initial experience shows that the students gladly accept such teaching approach which in turn positively affects the quality of learning. Ten students took part in the game experiment and selected results were presented. It was shown that the accuracy is reffered to the distance of the mean from the center of the target, i.e the absolute error. On the other hand, the precision is reffered to the dispersion of the samples, i.e the standard deviation.

REFERENCES

[1] BS ISO 5725-1: “Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions.”, p.1, 1994

[2] JCGM 200, International vocabulary of metrology Basic and general concepts and associated terms (VIM), 2008

[3] Taylor, John Robert, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books. pp. 128–129, 1999.

[4] BIPM - Guides in metrology, Guide to the Expression of Uncertainty in Measurement (GUM) and International Vocabulary of Metrology (VIM)

[5] Abdul Jabbar, A. I., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning: A systematic review. Review of Educational Research, 85(4), 740–779.

[6] Barr, M. (2018). Student attitudes to games-based skills development: Learning from video games in higher education. Computers in Human Behavior, 80, 283–294.

[7] Del Blanco, Á., et al. (2010). Integrating adaptive games in student-centered virtual learning environments. International Journal of Distance Education Technologies (IJDET), 8(3), 1–15.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Overview of EV charging stations impact on harmonic distortion in power distribution grid

Željko Hederić, Marinko Barukčić, Toni Varga and Nenad Cvetković

Abstract - In this paper, an overview of the impact of electric vehicle (EV) charging stations on harmonic distortions in the distribution grid are given. Problems that are recognized as a result of the impact of the rapid rise of EV technology and their implementation in modern urban transport are presented through an overview of relevant research primarily related to the EU region and the concentration of EVs and EV charging stations in only a few countries. The effects of harmonics produced by EV chargers are not properly limited by adequate protection causing malfunctions at the facilities and equipment of the residential facilities. The paper emphasizes the need to match the speed of growth of EV technology with the speed of changes in the accompanying standards and the improvement of designers' education.

Keywords - electric vehicles, charging infrastructure, harmonic distortion, power quality.

I. INTRODUCTION

The modern trend of the rapid development of electric vehicle technologies into everyday life brings many changes. The outlines of cities are changing, and there is an urgent effort to adapt city spaces to new trends of the use of electric vehicles that introduce a whole new culture of behaviour. It also requires a change in the system that supplied the energy needed for transportation, changing the fossil fuel delivery system with the power delivery systems. We all around us are accustomed to the power grid systems that bring us all the joys of light, the work of machines, household appliances, and television into our daily lives. It is a general misconception that such a grid system can serve to supply energy to new electric vehicle technology. This system is changing too slowly and is not adapted to new problems [1].

The idea of a smart city environment, energy transfers in all directions, two-way systems for charging vehicles from the grid and supplying power to the grid from vehicles, issues of interconnection of electric vehicles of different power at different points in the distribution grid, as well as generally problems of availability of charging stations for the number of electric vehicles are all this is only part of the problems that are just being recognized. But as no one can stop the development of electric vehicle technology, so we helter-skelter are trying to reshape and adapt power grid systems. This approach has opened the door to many problems.

One such problem that is not clear is the problem of harmonic distortion in distribution grids. In the paper are reviewed some solutions presented through research in various fields as an effort to point out the need to change the paradigms of design and control in distribution grids to ensure the quality of electricity supply to existing consumers.

EV battery chargers have evolved primarily as vehicles, but with increasing demand for energy and current levels, charging stations are nowadays emerging as stand-alone units that convert AC to DC (fast charging). Also, through the development of services, there was a need for two-way energy flow [2].

Vehicle charger connection systems can be conductive or inductive. Conductive charging systems are the most common because they have evolved as a logical sequence of a plug connector in daily use and use direct contact between the plug and the charging port [2]. Inductive chargers use resonance and magnetic coupling technology to ensure energy transfer. This type of EV charger is a novelty that provides charging capability on the moving and has evolved from magnetic levitating train technology [2]. The detailed predictions of harmonic distortions in existing distribution grids are complex because many different categories of portholes appear in both amounts and nonlinearities. Many nonlinear loads have very low power consumption, and although they are relatively more and more present in the distribution grid, the regulatory systems for standards for very low levels of consumer power do not prescribe high quality requirements.

II. THE IMPACT OF EV CHARGERS ON THE DISTRIBUTION GRIDS

Numerous studies addressing the impact of the rapid development of EV technology point out that over the last 10 years, there has been a 200-1000 increase in the Western EU (Fig. 1).

This rapidly grow brings problems affecting everyone in contact with EV technology, and especially distribution grids as a place for EV connectivity to the power source. In [1] are presented an overview of charging infrastructure for the road vehicles using data from EU parliament researches. The study [1] analyses the various challenges posed by implementations of EV systems and their technologies that enable the penetration of energy flow collection infrastructures within the EU.

These include existing technologies and standardization

Željko Hederić, Marinko Barukčić, Toni Varga are with Faculty of Electrical Engineering, Computer Science and Information Technology, J.J. Strossmayer University, Osijek, Croatia, E-mail: zeljko.hederic@ferit.hr,
marinko.barukcic@ferit.hr, toni.varga@ferit.hr,
Nenad Cvetkovic is with Faculty of Electronic Engineering, University of Niš, Niš, Serbia,
E-mail: nenad.cvetkovic@elfak.ni.ac.rs

issues that must accompany the accelerated growth of EV technology, measurement systems and business models, as well as connectivity to smart city systems with the implementation of renewables and Electricity Suitability technologies, all correlated with the public policies of individual EU member states.

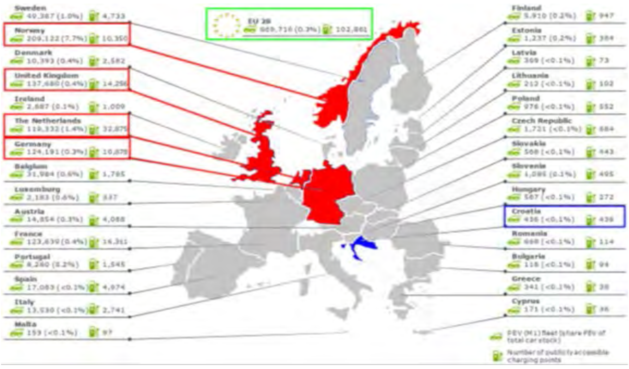


Fig. 1. The representation of EV distribution in EU with red marked area with 90% of total EV and blue are marked Croatia with less than 0,1% of total EV [1]

When we talk about electric vehicles and their charging points (influenced by marketing) we usually imagine them as shown in Fig. 2.

Fig. 2. The representation of marketing of top EV technologies, beautiful EV sports car by Rimac and fancy city small car on charging station (source: The Financial Times web newspaper)

Problem is that few are mention problems shown on Fig. 3. We should now that EV are not toys; they demand a huge density of energy in new storage compartment with not so tested technology of connection, control and use. In [1] are mentioned that 85% of the Fire departments are not familiar to procedures for safe extinguishing the fire of a burning electric vehicle containing an electrical outlet and possible high-current circuits.

Fig. 3. The firefighters need to train new ways and approaches to dealing with battery-powered vehicle fires
(source: The Sun web newspaper)

If we apply the rules of analogy thinking, then consider linking familiar things like fossil fuel vehicles to new EV problems. Why not have a fuel pump in every garage? What are the standards for a gas station? What can go wrong when connecting an EV to a charging station? How to respond preventively and predictively, and how to design them? Why was it not important to emphasize protection? Is it installed in an electric vehicle charging station?

The components of the electric vehicle power drive, as well as the charging station, are shown in Fig. 4. The EV drive itself is realized by an electric motor that controls the frequency converter; the DC bus system connects the subsystems inside the vehicle; the board battery charge input system and the battery itself with the battery management system. Then, different forms of EV charger stations that connect EVs to the distribution grid are shown.

Charging time (as well as battery life) are related to the charger's characteristics. The battery charger must be efficient and reliable, with high power density, low cost and low power consumption.

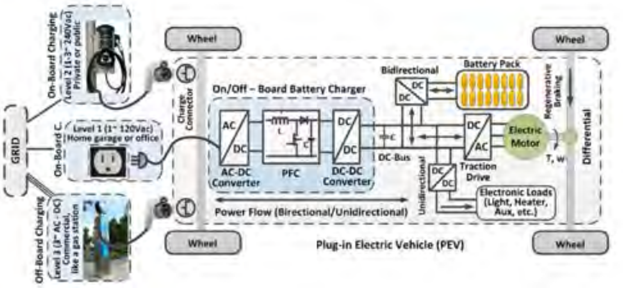


Fig. 4. On/off board charging system and power levels for EV different connections on charger [2]

The EV and charger subsystem shown in Fig. 4 is highly represented in the literature but it actually hides the error. What is missing in the picture? Why was it not important to emphasize protection? Is it irrelevant or is it integrated into the charger itself? Does safety devices by default need to be in the charging station? These are questions that are making a lot of confusion today.

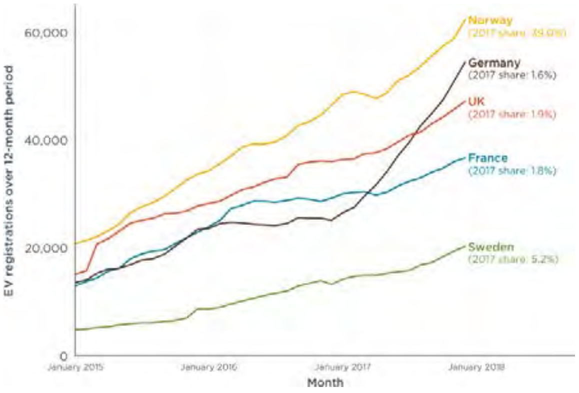


Fig. 5. Electric passenger car registrations numbers per years in

some EU countries who are leaders in EV technology implementation [1]

As standards are not changed quickly enough, the problem of accountability is emerging in many countries. If the standard is not set correctly, then wrong projects, which are not legally limited, are allowed to arise. Fig. 5 present rise of EV registration in the EU for the last 5 years what means developed EV technology and studies [1] show that EU countries standards are practically nonchanged in that period.

A. Battery pack design and modelling problems

The battery is an essential part of electric vehicles (EVs), so its design significantly affects the vehicle's drive characteristics, but also the vehicle's parameters as a dynamic load when charging vehicles at the charging station. In addition to the battery, in EV are several electronic converters that control the way they use the switch technology or the system that generates unwanted harmonics.

The method of connecting cells to the total package depends on a large number of design requirements (availability of energy, temperature protection, ...). Most often, they are connected in parallel blocks to provide capacity, and then the blocks are connected in series to the battery to ensure an adequate voltage level. What is often forgotten is that the battery does not store electricity. The battery is a reservoir of chemicals that produce electricity and store energy by a chemical reaction [3,4].

Because of this, the battery is a system that is continuously switched on, regardless of the total battery circuit being open inside the battery cells and the battery blocks, equalization currents are present. The Battery Management System (BMS) is designed to constantly monitor and optimize such a complex battery system, which has major problems with dynamic mode. Then cells are irregularly emptied, asymmetries are created and this can potentially break down.

A poorly defined material fatigue condition will cause the BMS, as a multiplicated switch system, to start to make irregular switching states and thus become a source of unwanted harmonics. This is especially true when charging quickly by connecting the vehicle to a charging station, in situations where the battery pack is below 50% of capacity level (this is by far the most common case when EVs are being charged to a charging station).

The battery pack design, as well as the BMS design, should be aligned. However, this is only fulfilled in initial states, and any intensive use promotes fatigue of the material and creates conditions for defective states. An example of this is the potential equalization currents inside a battery block, which in the long run leads to irreversible electrolyte degradation (battery potential).

Essentially, the stored chemicals within the cells are not capable of responding normally to electricity generation, by disintegrating and/or dissolving the structure of the active material bringing the cell to life [2]. There is numerous literature in which authors explain the problems that lead to battery failure. [5,6].

The fatigue of the materials that the electrodes are made of results in various side effects of Solid Electrolyte Interphase (SEI) [7], collectively referred to as ageing effects on the negative electrode and positive electrode (Fig. 6).

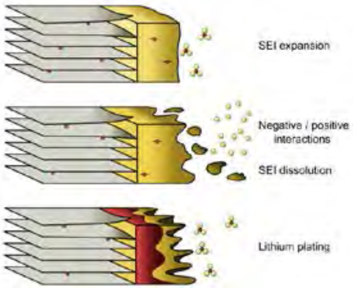


Fig. 6. Illustration of the ageing effect on battery negative electrode [7]

As can be seen in the block diagram (Fig. 7), numerous internal mechanisms of the battery and the environment affect the health of the battery. Most of them cause thermal decay of the chemical process, but these are already extreme situations that modern BMS systems are relatively successful at solving.

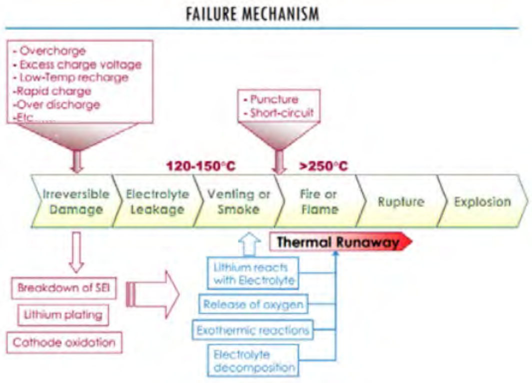


Fig. 7. Block diagram of failures event in Li-Ion batteries [1]

The problem arises when BMS fails as a control of the chemical process. As usual, man and negligence, that is, unprofessional work, play a major role.

Today, we have many vehicles that have come from workshops, garages and that have been successfully certified by the owners, although they have not passed even the most basic tests performed in the factories of the largest manufacturers. However, such vehicles also have access to public EV charger stations, with all the problems they can cause.

In [9], a modelling procedure is presented that gives rise to entirely new elements in battery circuits that are not provided by conventional designs. That is why most BMS systems today have deviations in EV battery management, that is, they have the wrong model by which they try to control the system of cells and blocks inside the battery, which then leads to uncontrolled switching states - the cause of unwanted harmonics. If we pay attention to the battery block (shown in Fig 8a), we can see that in the structure there is a current loop, which is exposed by its cross-section to magnetic fields generated in dynamic modes of operation. [6, 8, 9].

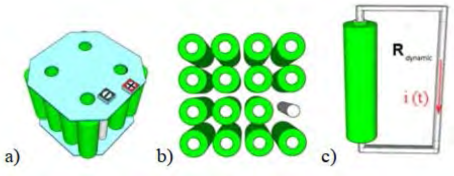


Fig. 8. Battery block design example and lumped model:
a) Battery block built by 15 parallel cells and interconnection bar. The contacts are on the same side on the upper plate,
b) Cross-section of the block with shown interconnection bar,
c) Lumped model of the electric circuit consists of one cell and interconnection bar [5]

The load, presented as a dynamic resistor, is a function of time (EV drive demand for energy from the battery as a function of time). Finally, the loop surface used to calculate the induced voltage source as a function of cell current (for the given geometry presented in Fig. 8.c) corresponds to the yellow marked area on Fig. 9.

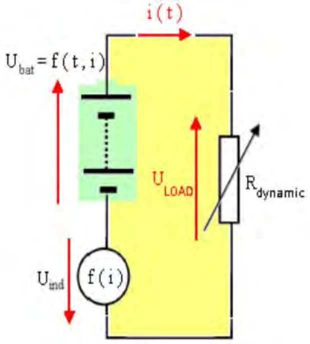


Fig. 9. Lumped model of an electric circuit for one battery cell as part of the battery parallel block [9]

B. The impact of harmonics on the distribution grid

To study what happens when we connect EVs to the distribution grid, it is necessary to define a well-known consumer model [10], for which there are well-developed and verified models as shown in Fig. 10.

Assuming that all consumers are within the IEC standard, no undesirable harmonic or THD levels should occur in the distribution grid. By entering EV charger in such topology and assuming that the charger itself meets the criteria of IEC standards by measurements and simulations, it is confirmed that distortions over the prescribed levels still occur in such altered topology [10].

Modelling of harmonic distortion, especially current waveforms in phases comparing with to without EV connected states, is presented in [10]. By plugging the EV into the charging station (location red in Fig. 10), only one phase current remains in the limits, while in the remaining phases and in the neutral conductor the current level rises as much as 6 times, and with it a significant increase in undesirable current harmonics.

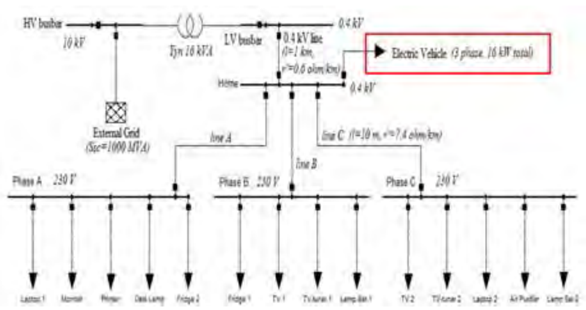


Fig. 10. Scheme of the composed model in DigSilent Power Factory for harmonic levels in the distribution grid simulation [10]

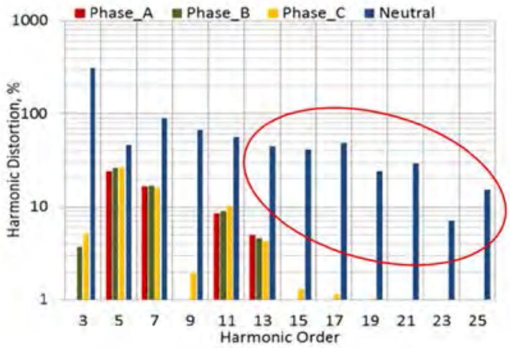


Fig. 11. Current harmonic distortion with EV connected in the distribution grid through charger [10]

In Fig. 11 showing current harmonic distortions with EV, it is evident that asymmetry and distortion occur, which consequently represents suitable conditions for resonant current harmonics, which are extremely dangerous for the highly sensitive and present ubiquitous computer, communication, monitoring technique.

C. BMS malfunction during EV charging

In the previous sections, we presented the problems of connecting EVs through charger units to the distribution grid, modelling batteries as part of an EV system. Another issue to consider is the battery management electronics system, which consists of numerous blocks. Battery management system (BMS) is to manage the output, charging and discharging and provide notifications on the status of the battery pack. They also provide critical safeguards to protect the batteries from damage. In all previous cases we present problems that occur with BMS normal function. What happens when a BMS malfunction occurs? Is it happen often?

A small 0.2ms decay or impulse jump causes 3-5kHz harmonics to "pass" through the transducer because they fit as a subharmonic of the switching frequency, and if no filters pass into the grid and look for a resonant pair.

Normal overcurrent protection does not respond. And now we have a wandering impulse in an unsecured distribution grid. Its power is enough to burn the Client–server systems power supply or air conditioning control.

This failure can result from a faulty battery pack model that cumulatively alters the charging process by the time the BMS does not recognize the "wrong" state relative to the estimated state. Then the BMS tends to reprogram and, due to the speed of operation of the switching part of the BMS, causes the appearance of an unwanted pulse that can be extended beyond the EV.

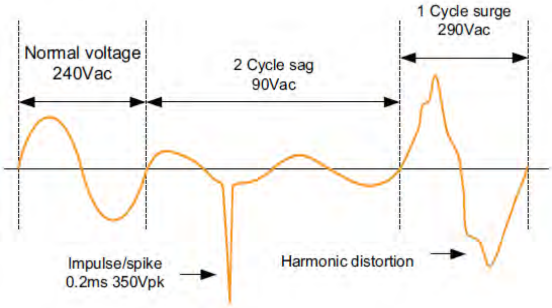


Fig. 12. Impulse present in the output from BMS during charging EV as an malfunction problem [11]

A further problem that arises is the case that there are multiple EV charger sites at the charger station at the same time [11]. When EVs of different manufacturers come on the distribution grid at the same time, we have a situation where undesirable harmonics of a much wider spectrum appear as in Fig. 12.

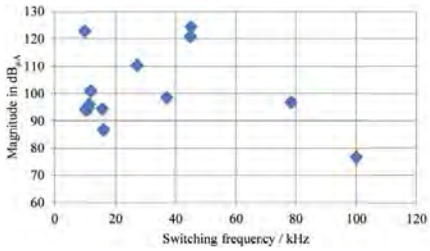


Fig. 13. Switching frequencies and magnitudes of different EV connected to EV charging group at same time [11]

Particularly prominent are distortions at frequencies below 2kHz (harmonics) and above 2kHz to 150kHz (supra-harmonics). Although the harmonic emission of EVs is strongly dependent on the grid voltage distortion, the supra-harmonic emission at the switching frequency of the inverter and its multiples is mainly determined by the design of the EV filter circuits.

The results of the harmonic and subharmonic distortion analysis of 19 different EVs connected via charging stations to the same distribution grid bus are presented in [11]. Note that different charge stations have approximately similar harmonic distributions but different harmonic spectra because they have different switching frequencies (Fig. 13). These are all kHz and are not multiples of the main harmonic of the distribution grid. The result is a partial resonance with a resonant pair within the distribution grid, a frequent supply of sensitive electronics that do not have the standard protection against such an attack.

What the authors point out in [12] is the fact that the classic standard safety devices are observing the distribution grid with substandard sampling, which does not seem to be a problematic harmonic image. Fig. 14 shows one such case in a 5EV chargers in the distribution grid. Significant differences exist in the harmonic performance between different EV types, which is majorly a result of different designs by the manufacturers.

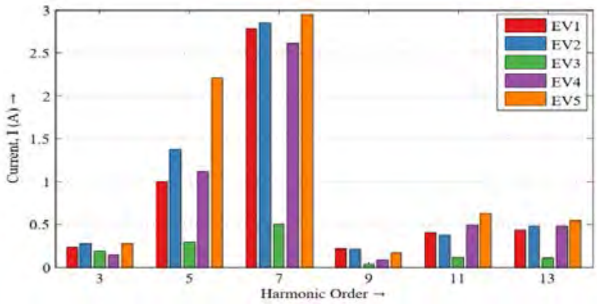


Fig. 14. The spectrum of current harmonic of the five EVs connected instantaneously on central charging infrastructure [13]

It is pointed out in [13] that in addition to the harmonic current emission, it is important to study the effect of the EV impedance on the distribution grid harmonic impedance. This is determined by the distribution grid side filter, which behaves often capacitively as it is required to reduce the high-frequency EV emission. Particularly if many EVs are connected to the same connection point, the risk of parallel resonances at low harmonics can be increased [13].

III. IMPACTS OF HARMONIC DISTORTION AND PROTECTION DEVICES

Current harmonics mainly affect the power distribution system from the bus to the inverter:

- Additional losses in wires and cables
- Additional heating of distribution grid transformers
- switch failures
- Third harmonics increase the zero component of the current and the voltage

Voltage harmonics affect mainly other devices connected to the bus (secondary of the distribution grid transformer):

- malfunctioning of telecommunication systems, computers, audio video equipment, monitors,

electronic control equipment. Etc.
- Resonance with capacitors for power factor correction
- Executive engine failures shortening the service life

How to reduce harmonic distortion? We need to be able to calculate - evaluate the harmonic distortion of the distribution grid and the impact of individual actors. This aspect is well reflected in practice by the standards defined for Structured Cabling and proper and high-quality grounding. Many IEC standards that have been defined for industrial plants where filter installation technology, especially external active filters, is normally present.

Very often we forget when observing the spectrum of currents that the time response is not a mean value but a "comb" [14] (Fig. 15).

The increase of harmonics at switching on and off of individual actors in a distribution grid that has combined connected consumers, an injected energy from a solar power plant and EV charger units can lead to the activation of protective devices (if installed) or to the swing of the distribution grid voltage.

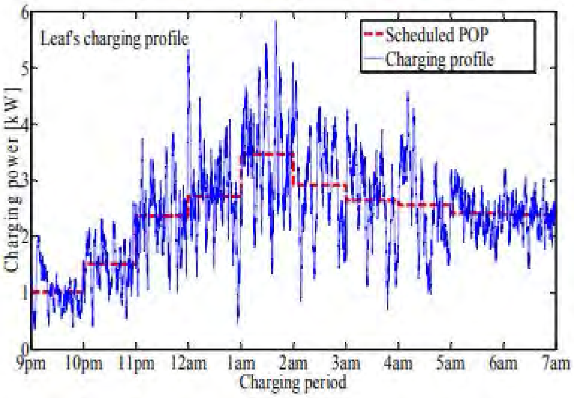


Fig. 15. Nissan leaf charging profile – the real measurement of power toward predicted profile [14]

There is a great danger of resonant harmonics - the harmonics of a particular EV charge can be in the rank of a single consumer opposite in phase - the resonance will capacitively burden the consumer and lead to possible adverse effects.

Current harmonics mainly affect the power distribution system from the bus to the inverter [15]:
- additional losses in wires and cables
- additional heating of distribution grid transformers
- switch failures (Fig. 16)
- third harmonics increase the zero component of current and voltage

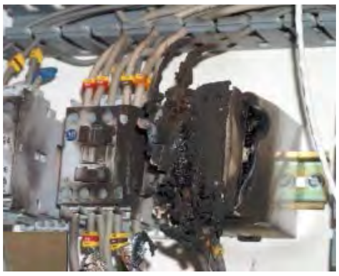


Fig. 16. Failure in switch box caused by current harmonics

Many authors recommend the installation of Active Power Filter (APF) filters that monitor only the 5th harmonic, and based on it evaluate all others, and therefore adjust the operation of the active filter. This is an expensive solution involving the existence of a shared bus, which has never been fulfilled in smaller rural areas and suburbs [15].

IV. CONCLUSION

The paper presenta a survey of harmonic distortion impact on an distribution grid as a result of EV charging, with respect to the design of the EV subsystems and its correlation to EV chargers. The given overview aims to highlight the impact of EV charger stations on power quality considering harmonic occurrences caused by the EV chargers operation. Based on the literature overview and real life data, the conclusion is that there is a need for better regulation of the EV charging station operation trough the standardization. The need for the standardized design of EV charges stations is detected based on the given overview.

Different situation (EV charging connection topology) can occur that can provide impressed harmonics into a distribution grid that mainly influence to unprotected and week loads in the same subcircuit (distribution grid connect on the distribution grid transformer secondary bus). A feeder is mostly likely to have a random distribution of EV chargers along with it. Therefore, for protection devices installation it must be carefully observed change of the power flow that ultimately rise current harmonics produced mainly from high power EV chargers.

ACKNOWLEDGEMENT

This work is supported by Josip Juraj Strossmayer University of Osijek scientific-research funding based on the cluster analysis and strategic interests of the region (Project UNIOS ZUP2018-110: Establishment of an environment for testing the electric vehicle drive subsystems).

REFERENCES

[1] Matthias Spöttle et al, “Research for TRAN Committee - Charging infrastructure for electric road vehicles”, *European Parliament research committees*, Brussels,

2018

[2] Yilmaz, Metin and P. T. Krein. “Review of Battery Charger Topologies, Charging Power Levels, and Infrastructure for Plug-In Electric and Hybrid Vehicles.” *IEEE Transactions on Power Electronics* 28 (2013): 2151-2169.

[3] MIT Electric Vehicle Team: “A guide to understanding battery specifications”, http://web.mit.edu/, 2008

[4] Steve McMullen: “Battery basic”,http://american-solarchallenge.org/ , 2015.

[5] Vetter, J., et al, “Ageing mechanisms in lithium-ion batteries”. *Journal of power sources*, 147(1-2),269-281. 2005

[6] F. Joho, P. Novak, M.E. Spahr, “Safety aspects of graphite negative electrode materials for lithium-ion batteries”, *J. Electrochem. Soc*. 149, A1020-A1024, (2002)

[7] P. Arora, R.E. White, M. Doyle, „Capacity Fade Mechanisms and Side Reactions in Lithium-Ion Batteries“, *Journal of the Electrochemical Society* 145 (10) 3647-3667, 1998

[8] F.A. LeBel, S. Wilke, B. Schweitzer, M.A. Roux, S. AlHallaj, J.P.F. Trovao: “A lithium-ion battery electrothermal model of parallellized cells”, *Vehicular Technology Conference*, 2016 IEEE 84th, 1-6

[9] T. Benšić, Ž. Hederić, M. Barukčić, M. Hadžiselimović, N. Cvetković, D. Krstić, “Battery Pack Design Problems – Influence of the Transverse Magnetic Field on Internal Battery Resistance”, *Safety engineering*, 7 (2017), 2; 49-53

[10]Niitsoo, Jaan & Kilter, Jako & Palu, Ivo & Taklaja, Paul & Kütt, Lauri. (2013). “Harmonic levels of domestic and electrical vehicle loads in residential distribution grids”. *IEEE AFRICON Conference*. 1-5. 10.1109/AFRCON.2013.6757800.

[11]Muller, S., Jan Meyer, Peter Schegner and Sasa Z. Djokic. “Harmonic Modeling of Electric Vehicle Chargers in Frequency Domain.” *Physics* (2015).

[12]A. Malano, S. Muller, J. Meyer, S. Bachmann, “Harmonic interaction of electric vehicle chargers in a central charging infrastructure”. *17th Int. Conf. on Harmonics and Quality of Power* 367-372 (2016).

[13]L. Kütt, E. Saarijärvi, M. Lehtonen, H. Molder, “A review of the harmonic and unbalance effects in electrical distribution grids due to EV charging”. 556-561 (2013), 10.1109/*EEEIC*.2013.6549577.

[14]Zhou, Niancheng & Wang, Jiajia & Wang, Qianggang & Wei, Nengqiao. (2014). “Measurement-Based Harmonic Modeling of an Electric Vehicle Charging Station Using a Three-Phase Uncontrolled Rectifier”. *IEEE Transactions on Smart Grid*. 6. 1-1, 10.1109/TSG.2014.2374675.

[15]Zhou, Niancheng, Jiajia Wang, Qianggang Wang, Nengqiao Wei and Xiaoxuan Lou. “Capacity Calculation of Shunt Active Power Filters for Electric Vehicle Charging Stations Based on Harmonic Parameter Estimation and Analytical Modeling.” (2014).

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Automatization of Digital Predistortion and Crest Factor Reduction

Borisav Jovanović, Srđan Milenković

Abstract – Power amplifier (PA) linearization, achieved by utilization of digital predistortion (DPD) technique, improves PA energy efficiency and reduces running cost of telecommunication equipment. In modern modulation schemes, the reduction of the peak to average power ratio of transmitted waveforms is necessary operation which ensures PA linearity. To achieve this, the Peak Windowing and DPD are implemented in Software Defined Radio (SDR) base station. The results are presented for Long-Term Evolution (LTE) waveforms and 10W modulated output PAs.

Keywords - Crest factor reduction, Peak to Average Power Ratio, Peak Windowing method.

I. INTRODUCTION

Digital predistortion (DPD), based on complex valued memory polynomials (MP), is established as an efficient method for radio frequency (RF) power amplifier (PA) linearization [1, 2]. However, modern modulation schemes facilitate high peak-to-average power ratio, which results in intercarrier interference, high out-of-band emission and bit error rate performance degradation [1]. Therefore, RF PA power has to be backed-off reducing the PA energy efficiency. To ensure that signals at PA input stay within linear region of the PA transfer function, the solution is in dealing with signals which have reduced PAPR. In this case it is possible to increase signal average power without the risk of PA operation in non-linear region [2]. This solution significantly cuts the running cost of wireless infrastructure.

Crest Factor Reduction (CFR) techniques are used for PAPR reduction [3, 4]. Our method for CFR is based on Peak Windowing (PW) [5]. The combination of DPD and PW PAPR reduction methods is implemented on Software Defined Radio (SDR) board and the results are given in the paper.

II. METHODS FOR PA LINEARIZATION

A. Crest Factor Measurement

Crest Factor (CF) a signal $x(n)$ is defined as the ratio between the amplitudes related to the largest $x(n)_{\max}$ and the mean values $x(n)_{\mathrm{rms}}$ of a signal:

$$CF = \frac{\|x(n)\|_{\max}}{x(n)_{rms}} \tag{1}$$

In literature one more parameter can be found - the Peak to Average Power Ratio (PAPR), which is the squared value of CF:

$$PAPR_{dB} = 10\log_{10}\frac{\|x(n)\|^2_{\max}}{x^2_{rms}} \tag{2}$$

The PA operates in the linear region whenever PAPR value of a waveform present at PA input is reduced. The one of the goals when implementing some CFR technique is to minimize in-band and out-of-band distortions. Unfortunately, signal distortion cannot be completely avoided. To quantify the performance of CF reduction operations, the distortion is measured by Error Vector Magnitude (EVM) for in-band and Adjacent Channel Power Ratio (ACPR) for out-of-band signal distortion. [2]

B. PA linearization by DPD

DPD computes the inverse function of the PA transfer characteristic. Utilization of DPD provides PA linearity and high energy efficiency. [1]

We have implemented the DPD solution which models the behavior of PA using complex-valued memory polynomials (MP). The MP model takes into account both PA nonlinearity and memory effects and represents the simplified structure of the Volterra series [2].

The predistorter is positioned in the base-band (BB), before the digital-to-analogue (DA) conversion and BB to RF frequency up-conversion are done. The input signal $x(n)$ is processed by DPD block which produces the output signal $y(n)$ according the following equation:

$$y(n) = \sum_{i=0}^{N}\sum_{j=0}^{M} w_{ij} \cdot x(n-i) \cdot e(n-i)^j \tag{3}$$

where w_{ij} are complex valued model coefficients:

$$w_{ij} = a_{ij} + \mathbf{j}b_{ij} \tag{4}$$

Borisav Jovanović and Srđan Milenković are with the Faculty of Electronic Engineering Niš, University of Niš, Aleksandra Medvedeva Street 14, E-mail: borisav.jovanovic@elfak.ni.ac.rs

$$e(n) = x_I(n)^2 + x_Q(n)^2 \tag{5}$$

The parameters N and M, used in Eq. (3), represent memory length and nonlinearity order respectively. The signal *e(n)* is the envelop of the input signal *x(n)*.

The DPD coefficients are found in iterative training process which uses recursive least square (RLS) method and indirect learning architecture. The detailed description of DPD method and its implementation in SDR hardware is described in detail in [6].

C. Peak Windowing Method

The Hard Clipping (HC) is conventional CFR approach to constrain signal peaks which exceed selected threshold. The operation of HC is described by Eqs. (6) and (7):

$$y(n) = c(n)x(n)\,, \tag{6}$$

where *x(n)* is the input signal, *y(n)* is the signal obtained after clipping operation is performed. The signal *c(n)* is the clipping function:

$$c(n) = \begin{cases} 1, |x(n)| \le A \\ \dfrac{A}{|x(n)|}, |x(n)| > A \end{cases}, \tag{7}$$

where the parameter A is clipping threshold. *x(n)* and *y(n)* are complex signals consisting of quadrature signal components. The HC operation has disadvantage in introducing sharp edges in output signal which increases out-of-band distortion.

In PW method the large signal peaks of input signal are multiplied with a windowing function to smooth the sharp signal edges at clipping points, thus minimizing in-band and out-of-band signal distortion. This not only improves the ACPR of the resulting signal but also limits the peaks to stay under the threshold.

The PW operation replaces the clipping coefficients *c(n)* with new ones *b(n):*

$$b(n) = 1 - \sum_{k=-\infty}^{k=\infty} (1 - c(k))q(n-k)\,, \tag{8}$$

where *q(n)* is a common symmetric window function, for which implementation the Kaiser, Hamming or Hann windowing functions can be used [3, 7].

The calculation of clipping coefficients *b(n)* is implemented by PW FIR filter. Two conditions have to be met. First to ensure that the value of *y(n)* is less than the threshold A, the condition given by Eq. (9) must be met:

$$b(n) \le c(n) \tag{9}$$

Besides, to minimize in-band signal distortion the last inequality must be near the equality as much as possible. [3] This implies that narrow window lengths should be used for PW FIR filter implementation.

If clipping rate is too large or wide window length is selected, the adjacent windows in *b(n)* overlap more often, reducing *b(n)*, which yields to more attenuation of the CFR output signal *y(n)*.

The detailed description of PW operation and its implementation in SDR hardware is described in detail in [8].

III. IMPLEMENTATION RESULTS

A. Hardware implementation

The SDR board [9] utilizes two transceiver LMS7002M [10] ICs for frequency conversion between base band (BB) and radio frequencies (RF) and an Altera Cyclone V FPGA, which are used for implementation of DPD and CFR digital blocks. The board is connected to CPU core through high-speed PCIe interface. For the development or demo, test waveform can be uploaded and played from WFM RAM Block, implemented using Altera Cyclone V FPGA resources. In real applications, the CPU Core performs BB digital modem functions, LTE for example. Through PCIe interface the real waveform is fed directly to the CFR block input [10].

The SDR board implements 2x2 Multiple-Input and Multiple-Output (MIMO). It has two transmit TX and two receive RX channels. One spare RX chain is used as DPD monitoring path. In each TX channel, beside DPD and CFR blocks, a low-pass FIR filter is used. The CFR output is filtered by low- pass FIR filter, reducing out of band spectrum regrowth.

PC/GUI implements graphical display for demo and debugging purposes. GUI is capable to show important signals at CFR and DPD block input and outputs in FFT (frequency), time and constellation (I vs. Q) domains.

The board uses LMS7002M on-chip DACs/ADCs [10]. On chip data converters are 12-bit devices. The CFR digital block is implemented as 40-tap FIR filter and operates at 30.72 MS/s sample rate. The CFR block has provision for changing FIR filter order in the range from 1 to 40. Using the same interface, the clipping threshold can be changed up to the signal amplitude maximum level.

The DPD operates at sample rate of 61.44 MS/s. In DPD implementation, the memory length N=4 and nonlinearity order M=2 are selected.

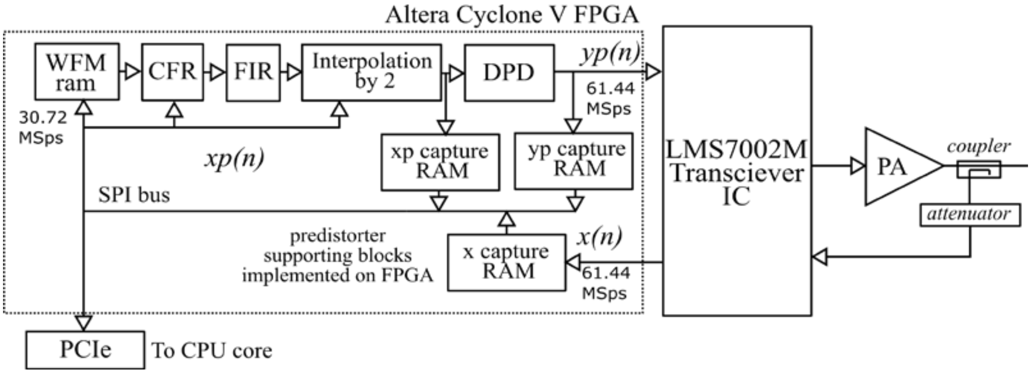


Fig. 1. The CFR and DPD implementation based on SDR transceiver board

The SDR board, together with two 40 dBm Band13 10W power amplifiers have been incorporated in a Personal Computer (PC) based 2x2MIMO base station (BS). To provide power supply to PAs, additional DC/DC converters are incorporated in the BS case. The DC/DC converters get 12V from PC power supply unit and provide 28V, necessary for PA operation.

B. Software implementation

The DC/DC converters and PAs are embedded into base station and these blocks have special input enable pins which are used for external control. We implemented that the DC/DC converters and PAs are controlled directly by SDR board output pins and specially created software application, which is started from Linux command line prompt.

The application provides following commands:

- *startDCDC* and *stopDCDC* to start and stop the DC/DC, and
- *startPA* and *stopPA for* PA turn on and turn off.

Since two independent transmitting channels exist in BS, the commands take as argument constant values 1 or 2 selecting different transmitter channels and corresponding DC/DC and PA.

The same application provides software routines are used for full DPD and CFR control. They are described below.

CFR block is controlled by *setupCFR* command providing different CFR options and parameters. For channel selection there are two options: 1 or 2. The second argument of the *setupCFR* function is the bypass option. Namely, the CFR block can be bypassed by setting the bypass option to 1; otherwise, when CFR is used, this option should be set to zero value. The third argument is CFR FIR filter order, which is an integer value and can be chosen in range from 1 to 40; the last argument is the threshold, specified by positive real number which is less than value of 1.0.

The DPD operation is controlled by several commands which can be generally divided into following groups:

- commands for DPD parameter specification
- starting and stopping the DPD operation
- calibration and reset
- storing and reading the DPD parameters, as well as the DPD coefficients

The DPD parameters are specified by *setGainDPD*, *setLambdaDPD* and *setND_DPD* commands.

The *setGainDPD* is sets the digital gain *Gain* of DPD, adjusting the PA output power when DPD is in training process. The PA output power can be increased, decreased or set to the same level compared to the case when DPD is bypassed. The digital gain is positive real number which can be chosen in range from 0.33 to 6.0.

The command *setLambdaDPD* specifies the DPD “forgetting factor” λ which adjusts the speed of DPD algorithm convergence, e.g. how fast the PA becomes linearized. The λ is the real positive number close to one.

Command *setND_DPD* defines the DPD delay line length *ND*.

The command *calibrateDPD* automatically performs DPD parameter calibration. It calculates the DPD delay line length *ND* and gain value *Gain*. When DPD training process is started, the PA output signal power remains at the same level as in case when DPD is just bypassed. Without any loss in output power, the PA becomes linearized when DPD is calibrated and then, training process is started.

Command *resetDPD* is used for DPD coefficients reset operation. When DPD coefficients are reset, it is identical situation as when DPD is bypassed.

The commands *saveConfigDPD* and *loadConfigDPD* store into the memory and read from memory the last found DPD parameters *ND*, *Gain* and λ.

The operation of DPD training is started and stopped using *startDPD* and *stopDPD* commands respectively. It is strongly recommended that DPD is calibrated before it is started. When DPD is started, the DPD filter coefficients are continuously changed in the DPD training process.

Commands *loadCoeffDPD* and *saveCoeffDPD* store and read from memory the DPD filter coefficients. Using these commands the PA can be compensated when the DPD coefficients are just read from memory and loaded into FPGA, even when DPD training operation is not started by *startDPD*.

All software routines are written in C++.

C. Measurement Results, case of 10MHz LTE waveform and 10W PA

The CFR and DPD methods were implemented in SDR board FPGA IC to reduce the distortions of transmitted signals observer at PA output The SDR board has been incorporated into PC-based BS, together with two RF PAs having 10W modulated output power. The PAs have following characteristics: the bandwidth is 700-850 MHz; average output power at 1 dB compression point is 40 dBm at the frequency of 750 MHz.

The amount of out-of-band and in-band distortion is measured in terms of ACPR and EVM respectively. |Here, the results are provided for 10MHz LTE waveform. In the measurements different CFR parameters are considered. For example, the filter order is changed in following steps *L*=9, 19, 29 and 39. Also, different clipping thresholds are evaluated, staring from *Th*=1.0 down to *Th*=0.6. While these parameters are changed, the level of transmitted output power is kept at the same level as before linearization process is started.

The measured PAPR of original waveform, when DPD and CFR are bypassed, is PAPR= 10.3 dBm. The measured EVM in this case is EVM=1.26 %. When CFR and DPD are utilized, the EVM is worsened when threshold *Th* value is reduced, and also when the CFR filter order *L* is increased. For example, in case of *L*=39 and *Th*=0.6, the EVM =7.81%.

At starting point, when neither CFR nor DPD are utilized, the ACPR=-37.5 dBc and EVM=3.32%. The measured PA output power is Pout=39.7 dBm.

The distortion is removed by DPD and CFR. The ACPR is decreased from starting value of -49.6 dBc, obtained for threshold *Th*=1.0 to the minimum value equal to -50.8 dBc, obtained at *Th*=0.6. The power of PA output signal is maintained at the same level as at starting point, before any CFR and DPD data processing is done.

When PAPR is decreased by 2dB (from PAPR=10.3 dB down to PAPR=8.3 dB), for *L*=19 and *Th*=0.76, the EVM=2.42%, ACPR=-50.5dBc. Therefore, the ACPR is improved by 13 dBc and EVM by 0.8%.

IV. CONCLUSION

CFR and DPD algorithms have been implemented on SDR board and verified by measured results. When DPD is used without CFR, the modulated signal power at PA output needs to be backed-off for approximately 2 dB to enable DPD to remove distortions down to the system noise floor. When CFR is utilized, DPD is capable of cancelling any out-of-band distortion above system noise floor, without PA power back-off.

The DPD and CFR are controlled by specially created software application which is started through Linux command line prompt. The application provides full DPD and CFR control.

The main characteristics of our solution are low implementation complexity and good performance measured in terms of EVM and ACPR as key parameters.

REFERENCES

[1] Eun, C. and Powers, E. J., "*A New Volterra Predistorter Based on the Indirect Learning Architecture*", IEEE Trans. Signal Process., Vol. 45, No. 1, 1997, pp. 223–227

[2] Cavers, J. K. "*Amplifier Linearization Using a Digital Predistorter with Fast Adaptation and Low Memory Requirements*", IEEE Trans. Veh. Technol., Vol. 39, No. 4, 1990, pp. 374–382

[3] Lim, D.W., Heo, S.J., No J.S., "*An Overview of Peak-to-Average Power Ratio Reduction Schemes for OFDM Signals*", Journal of Communications and Networks, Vol. 11, No. 3, June 2009, pp. 229-239.

[4] Jiang, T., Wu, Y., "*An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM signals*", IEEE Trans. Broadcasting, Vol. 54, No. 2, June 2008, pp. 257-268.

[5] Mistry, H., "*Implementation of a Peak Windowing algorithm for Crest Factor Reduction in WCDMA*", Master of Engineering Thesis, Simon Fraser University, Canada, 2006

[6] Jovanović, B. Milenković, S, "*PA Linearization by Digital Predistortion and Peak to Average Power Ratio Reduction in Software Defined Radios*", Journal of Circuits Systems and Computers, World Scientific Publishing Co. Pte. Ltd., Singapore, accepted for publication on Oct 03, 2019, https://doi.org/10.1142/S0218126620501479

[7] Jovanović, B. Milenković, S, "*Peak Windowing for Peak to Average Power Reduction*", Proceedings of the 7th Small Systems Simulation Symposium, Nis, Serbia, 2018, pp. 33-36

[8] Jovanović, B. Milenković, S, "*The Peak Windowing For PAPR Reduction In Software Defined Radio Base Stations*", FACTA UNIVERSITATIS, Series: Electronic and Energetics, University of Niš, accepted for publication on Dec 11, 2019

[9] Limemicrosystems LimeSDR QPCIe (2019), https://wiki.myriadrf.org/LimeSDR-QPCIe

[10] Limemicrosystems LMS7002M (2019), https://limemicro.com/

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Modular Baseband Processing for mm-Wave and THz Communication

Goran Panić, Mohamed Eissa, Lukasz Lopacinski, Nebojša Maletić and Rolf Kraemer

Abstract - This paper discusses the challenges of baseband signal processing in wideband communication systems operating in millimeter-wave (mm-Wave) and terahertz (THz) frequency bands. The focus of the paper is the concept of modular baseband processing solution that is explained, implemented and demonstrated.

Keywords – modular baseband, mm-Wave, THz communications, signal processing

I. INTRODUCTION

The demand for higher data rates in modern wireless communication increases rapidly as the number of new users and applications constantly grows. As observed by Edholm's Law [1], wireless data rates will continue to increase exponentially, and by 2021, they should reach 100 Gbit/s. However, accomplishing such high data rates will be only possible either by dramatically improving the spectral efficiency in allocated mm-Wave spectrum or by utilizing large spectrums available in the THz frequency band [2]. Potential applications for high-speed wireless communication at mm-Wave and THz frequencies include kiosk downloading, high capacity links for wireless local area networks (WLANs), high throughput links for wireless personal area networks (WPANs), ultra high definition video streaming, augmented and virtual reality, wireless connection in data centers, fronthaul/backhaul wireless links for future 5G applications, and nano-cells [2,3].

Millimeter-wave spectrum is the band of spectrum between 30 GHz and 300 GHz. Since spectrum regulations and available bandwidth fairly limit conventional radio communication operating at 2.4 GHz and 5 GHz, the unlicensed spectrum between 24 GHz and 100 GHz is considered as potential target for realization of future long- and medium-range high-speed wireless communication links. The frequencies above 100 GHz are considered for high-speed, short-range indoor communication. Currently, the 7 GHz wide spectrum around 60 GHz is already allocated to mobile communications. The IEEE802.15.3c and IEEE802.11ad standards define 60 GHz communication for WLAN and WPAN short-range applications, respectively, allowing the data rates of up to 8 Gbit/s. A follow-up to the 802.11ad, the IEEE802.11ay standard, introduces channel bonding, higher modulation schemes and multiple antenna approach (4×4 MIMO), and promises aggregated data rates of up to 176 Gbit/s. However, recent commercial solutions for "ay" are still limited to the data rates not higher than 10 Gbit/s. The high complexity of hardware and the challenges in signal processing prevent reaching the ultimate goals of "ay" specification.

There are different definitions of the THz band in literature. Usually, the unlicensed spectrum between 275 GHz and 1 THz is referred to as the THz band. However, the spectrum between 100 and 300 GHz is also seen as a part of the THz band and is often referred to as a sub-THz band. The THz band makes itself an excellent candidate to answer the demands for even more bandwidth and extremely high data rates. Recently, the unlicensed D-band (110-170 GHz) received a lot of attention in the context of terabit wireless communication [4]. Radio waves in the THz-band are strongly affected by atmospheric attenuation [5]. Therefore, most applications consider line-of-sight (LOS) short-path indoor scenarios for THz communication. Nevertheless, by using high gain directive antennas, it is also possible to establish outdoor THz communication over a long distance [6].

The rest of the paper is organized as follows. Section II discusses the most important challenges in the realization of terahertz communication. Section III presents the concept of modular baseband signal processing solution for THz applications. Finally, Section IV presents the experimental validation of the proposed concept.

II. CHALLENGES IN THE THZ BAND

Some of the most important challenges related to THz communication include technology limits, antenna and transceiver design and challenges related to the realization of a communication system. Thereby, the system size, production costs, power budget and security aspects also play an important role. Many challenges specific to mm-Wave also apply to THz communications.

A. Device Technology

A straightforward choice regarding mm-Wave and sub-THz technology for cheap, highly-integrated, compact devices is CMOS (Complementary Metal-Oxide-Semiconductor). A standard silicon-only CMOS technology is already massively used in the design of

Goran Panić, Mohamed Eissa, Lukasz Lopacinski, Nebojša Maletić and Rolf Kraemer are with IHP – Leibniz-Institut für innovative Mikroelektronik, Im Technologiepark 25, 15236 Frankfurt (Oder), Germany, Email: {panic, eissa, lopacinski, maletic, kraemer}@ihp-microelectronics.com

60 GHz RF circuits [7]. Furthermore, CMOS is also receiving a lot of attention regarding the design of sub-THz circuits [8]. However, the losses due to parasitic effects and low output power limit the application of CMOS for integrated systems operating at extremely high frequencies. Recent advances in SiGe BiCMOS (Silicon-Germanium Bipolar CMOS) technology enable hetero-junction bipolar transistors (HBT) with cut-off frequencies (f_T) of up to 300 GHz and maximum frequency (f_{max}) of up to 500 GHz [9]. Having f_{max} significantly higher than its CMOS counterpart, the BiCMOS allows the design of ultra high-frequency components with much better performance and power figures. Also, the compatibility with CMOS makes BiCMOS a promising technology for the realization of future mm-Wave and THz communication systems. High-power applications in the THz band usually consider Gallium Nitride (GaN) or Indium Phosphide (InP) technology. Some other technologies are also considered for the generation and detection of THz waves such as photonic, Gallium Arsenide (GaAs), and graphene based. However, the wide usage of these technologies is restricted by high processing costs and immaturity of the technology process. Nevertheless, hybrid integration with CMOS or SiGe BiCMOS is also possible.

B. Transceiver Design

The realization of THz circuitry is closely related to the limitations posed by the target device technology. For standard silicon-based process technologies, the design of circuit components, such as oscillators and amplifiers, is a very challenging task. The oscillators for the THz band must deal with various issues such as low efficiency, low output power, and high phase noise [10]. Optical oscillators promise better noise figures and high output range, but they come at a high cost, having large size and limited output power. The power amplifier design in standard CMOS is mainly affected by low output power. The BiCMOS provides more gain, as presented in [11]. Another important element of a THz system is the antenna. For outdoor use, e.g. fixed backhaul radio links, a highly directive horn antenna with optional beam-collimating lenses could provide enough gain and enable long ranges [12]. More compact is a phased-array antenna with optional beam steering that could be efficiently used for both outdoor and indoor applications [13]. However, the outdoor beamforming is not easy to deploy, due to various environmental effects affecting the received signal. For a stable link, very high gain and excellent beam alignment are required. An integrated phased-array antenna with implemented beamforming is a suitable solution for consumer electronic devices with strict requirements for size and acceptable production cost.

C. Signal Processing

Wireless communication over THz channels is strongly affected by very high molecular absorption and spreading loss resulting in a very high and frequency-selective path loss for LOS links. Non-line-of-sight (NLOS) propagation is governed by high reflection loss depending on the shape of reflecting surface, its roughness and material features. A study on deterministic indoor LOS and NLOS channel models suggests that 1 Tbit/s of LOS and 100 Gbit/s of NLOS transmission over a distance of 1 meter is possible with the transmit power of 1 W [14]. In directed NLOS scenario, highly reflective mirrors in combination with steerable antennas can be used to avoid obstacles and minimize multi-path effects [15]. The measurements of an indoor broadband channel at 300 GHz are presented in [16]. The results showed a quite flat channel characteristic for symbol rates up to 10 GSymbols/s confirming potential applicability of the link for short-range indoor applications. Regarding outdoor applications, such as backhaul links relying on high gain highly-directive antennas, the multi-path effects are not of great concern. However, the effects related to atmospheric attenuation, scintillation and fog have to be taken into consideration.

Large available bandwidth allowing very high data rate is certainly a major advantage of the THz band. RF-circuits for THz transmitters and receivers have been already demonstrated. However, the processing of large bandwidth is an extremely challenging task. The issues due to a high level of noise and changes in channel properties make channel equalization and error correction very difficult. Therefore, baseband modems and data converters need to maintain high resolution for signal processing. They must also provide the required data throughput to higher communication layers. For example, the state-of-the-art successive-approximation-register (SAR) analog-to-digital converters (ADCs) in CMOS provide up to 2.4 GHz of Nyquist bandwidth and have the resolution of 3-5 effective numbers of bits, limiting the constellation to 16-QAM at most [8]. On the other hand, flash-based ADCs with sufficient bandwidth are power-hungry and not suitable for low power applications required in consumer devices. One viable option to deal with this problem is channel fragmentation and aggregation.

III. MODULAR BASEBAND PROCESSING

The concept of modular baseband processing relies on multi-level band fragmentation and link aggregation. The link aggregation and channel bonding are the concepts already applied in various IEEE standards and propriety solutions to add redundancy or increase the throughput. For IEEE802.11 WLAN, the impact of IEEE standards and regulatory restrictions make the use of channel bonding very difficult. This is partially due to dynamic frequency selection (DFS) requirements, a process by which the access point must detect the signature of existing government weather radar and other radio systems and vacate the channel for an hour. Also, increased baseband complexity, sensitivity to adjacent channel interference, frequency reuse issues and compatibility to legacy 802.11a/b/g products limit the application of channel bonding in IEEE802.11 WLAN systems operating at 2.4 or 5 GHz [17]. The channel bonding is also defined for the 60 GHz IEEE802.11 ad/ay allowing the bonding of up to 4 channels. The 60 GHz analog front-end (AFE) transceivers supporting channel bonding at the physical layer are presented in [18, 19]. However, no circuits for analog channel bonding have been presented yet.

The existing concepts of baseband processing with support to channel bonding are usually realized by adding more complexity to specific signal processing units, e.g. by increasing the FFT size, increasing the processing speed of de-multiplexer and demodulator, increasing the sampling rate and the bandwidth of data converters, etc. The idea of modular baseband processing is to employ multiple baseband cores running in parallel that are realized either as a module or as a single chip. In this concept, when transmitting, the big data stream generated by a host is split to multiple baseband streams processed by several baseband cores running in parallel. The generated analog signals at the outputs of baseband processors (usually I/Q outputs from the baseband DACs) get bonded in analog domain creating a wideband signal at an intermediate frequency (IF). The generated wideband signal gets up-converted to selected mm-Wave/THz carrier frequency in an analog front-end. At the receiver side, the received wideband signal gets down-converted to IF and split to multiple analog streams. The analog signals get converted to digital data at the ADC inputs of the baseband cores and processed in parallel. The processed digital data streams get combined at the interface to host and passed to the upper layers of the communication system. The concept of modular baseband processing using three baseband processors is illustrated in Fig. 1.

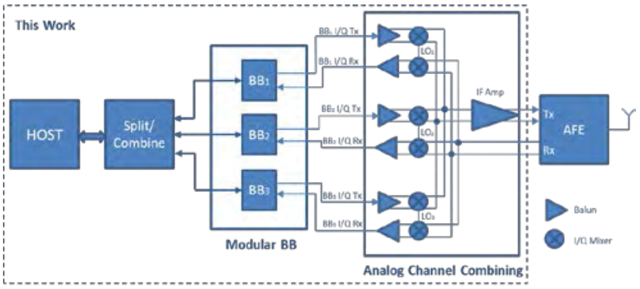


Fig. 1. A modular baseband processing solution.

There are several advantages of the proposed concept. Splitting the frequency band to moderately wide sub-bands allows the implementation of power-efficient data converters in the baseband cores. Furthermore, if the load of a network is not high so that the system is not required to operate at full capacity, the selected cores might be put into a sleep mode in order to save the power. From the access point view perspective, multiple cores running in parallel can serve single or multiple users. Assuming the baseband cores with an adaptive OFDM modulation, the sub-bands could be further channelized by assigning a group of subcarriers to single or multiple users. Finally, multi-level bandwidth fragmentation provides flexibility to apply various combinations of time, space and frequency multiplexing in the network layer. A potential drawback of the proposed concept is the complexity of the analog channel combining circuits that might require power-hungry filters in order to prevent the leakage between the channels. Also, the generation of IF frequencies (i.e. LO chain) might be a challenging task.

IV. EXPERIMENTAL VALIDATION

A communication link for validation of the modular baseband signal processing concept has been implemented and tested. The setup consists of transmitter (Tx) and receiver (Rx) stations, each consisting of a PC host, a split-and-combine unit, a modular baseband module and an analog channelizer block (see Fig. 2).

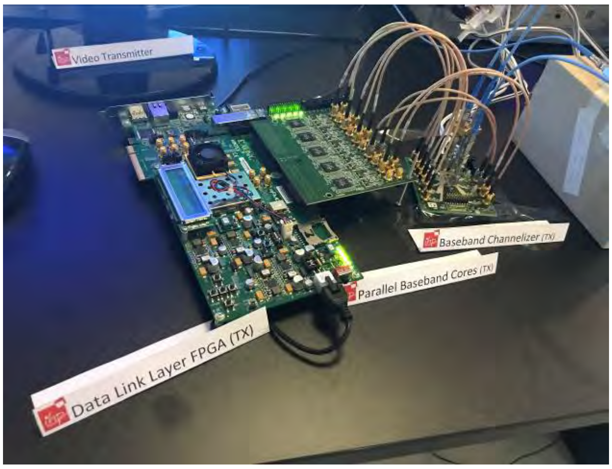


Fig. 2. Modular transmitter station.

A. Baseband Module

To validate the proposed concept, a multi-core baseband processor module has been developed, which implements five cores on a printed circuit board. This is shown in Fig. 3. The used baseband cores support OFDM-QPSK modulation with 528 MHz of bandwidth and up to 480 Mbit/s data throughput. The board interface to the host is realized by two FMC connectors, compatible with the FMC connectors on the Xilinx Kintex-7 FPGA board, which is used to implement the split-and-combine logic. The interface to the analog channel combiners is provided by SMA connectors as well as by compact Q-strip connectors. The power supply is realized either through FMC or externally.

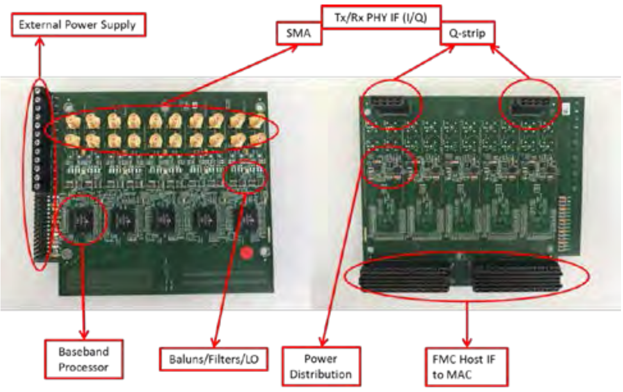


Fig. 3. 5-channel modular baseband processor module.

B. Analog Channelizers

The analog channelizer integrated circuits (Tx and Rx) were designed and fabricated in IHP's SiGe BiCMOS 130nm SG13G2 technology. The chips are mounted and bonded on module boards containing SMA interfaces to baseband channels (see Fig. 4). The module also implements a wideband, high frequency, differential output connector to mm-Wave or THz AFE and an input connector for the frequency source. The board also provides power supply connectors. The channelizers support the bonding of 5 channels with 2.5 GHz bandwidth per channel. More information on channelizer circuits can be found in [20].

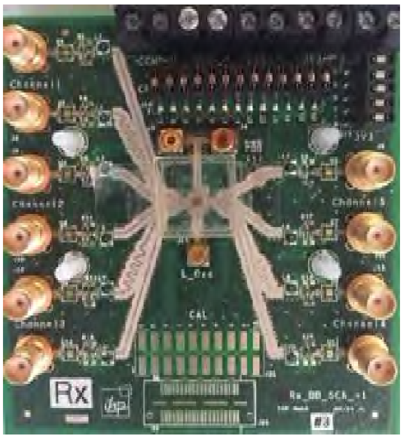


Fig. 4. Rx channel debonding module.

C. Split-and-Combine Module

The host interface is realized over an FPGA split-and-combine module that sends and receives packets from a PC host running the data link and application software. The interface between the PC host and the FPGA board is realized by a 10 Gbit/s optical Ethernet link. The split-and-combine module receives jumbo packets from the host and generates data streams to the cores of the modular baseband. It also receives the data streams from the baseband cores and assembles the jumbo packets that are sent to the host. Two FMC connectors realize the interface to the baseband module. Some other functionalities of the split-and-combine module include packet scheduling and packet enumeration. The initialization of the baseband module is also done by a program sequence running on the microcontroller within an FPGA device.

D. Communication Link

The final demonstration employs a back-to-back unidirectional wired link between two stations, as depicted in Fig. 5. The video stream signal is generated by the transmitter PC and sent to the receiver station, which displays the received stream. Both stations implement the software for packet processing and control of the Ethernet communication. An uncompressed 720p video stream could be transmitted and received in real-time over all channels without visible errors in transmission.

Fig. 5. Demonstration of communication link with modular baseband signal processing.

V. CONCLUSION

The concept of modular baseband processing has been proposed and validated in this paper. The modular approach, based on multi-level band splitting, parallel processing and analog channel combining, is shown to be an effective solution for the realization of communication links in wideband mm-Wave and THz spectrum.

ACKNOWLEDGEMENT

This work has been funded by the German Federal Ministry of Education and Research within the research project ForMikro-6GKom (project number 16ES1107).

REFERENCES

[1] Cherry, S., “Edholm’s Law of Bandwidth,” *IEEE Spectrum*, Vol. 41, No. 7, pp. 58–60, 2004.

[2] Kürner, T., Priebe, S., “Towards THz Communications - Status in Research, Standardization and Regulation,” *Journal of Infrared, Millimeter, and Terahertz Waves*, Vol. 35, No. 1, pp. 53-62, 2014.

[3] Nagatsuma, T., “THz communication systems,” *Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1-90, IEEE, 2017.

[4] Frecassetti, M. et al., “D-Band Transport Solution to 5G and Beyond 5G Cellular Networks,” *European Conference on Networks and Communications (EuCNC 2019)*, Valencia, Spain, 2019.

[5] Grigat, M. et al., “Propagation Aspects of Terahertz Outdoor Fixed Wireless Links,” *IEEE 802.15 Document 15-13-0175-01-0thz*, Atlanta, Orlando, 2013.

[6] Federici, J., Moeller, L., “Review of terahertz and subterahertz wireless communications,” *Journal of Applied Physics*, Vol. 107, No. 11, pp. 6, 2017.

[7] Razavi, B., “Design of millimeter-wave CMOS radios: A tutorial,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 56, No. 1, pp. 4-16, 2009.

[8] Tang, A., Chahat, N., Decrossas, E., “CMOS THz communication links for wireless applications: Where do they fit into mobile access and fixed access?,” *39th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz)*, pp. 1-2, Tucson, 2014.

[9] Rucker, H., Heinemann, B., Fox, A., “Half-terahertz SiGe BiCMOS technology”, *IEEE 12th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, pp. 133–136, IEEE, Jan. 2012.

[10] Khatibi, H., Afshari, E., “Towards efficient high power mm-wave and terahertz sources in silicon: One decade of progress,” *IEEE 17th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, pp. 4-8, IEEE, 2017.

[11] Eissa, M., Kissinger, D., “A 13.5 dBm Fully Integrated 200-to-255GHz Power Amplifier with a 4-Way Power Combiner in SiGe:C BiCMOS”, *IEEE Solid-State Circuits Conference (ISSCC)*, IEEE, San Francisco, 2019.

[12] Rebeiz, G. M., Rutledge, D. B., “Integrated horn antennas for millimeter-wave applications,” *Annales des télécommunications, Springer-Verlag,* Vol. 47, No. 1-2, pp. 38-48, 1992.

[13] Hur, S., Kim, T., Love, D. J., Krogmeier, J. V., Thomas, T. A., Ghosh, A., “Millimeter Wave Beamforming for Wireless Backhaul and Access in Small Cell Networks,” *IEEE Trans. Communications*, Vol. 61, No. 10, pp. 4391-4403, 2013.

[14] Moldovan, A., Ruder, M. A., Akyildiz, I. F., Gerstacker, W. H., “LOS and NLOS channel modeling for terahertz wireless communication with scattered rays,” *IEEE Globecom Workshops (GC Wkshps)*, pp. 388-392, IEEE, 2014.

[15] Piesiewicz, R., Kleine-Ostmann, T., Krumbholz, N., Mittleman, D., Koch, M., Schoebei, J., Kurner, T., „Short-range ultra-broadband terahertz communications: Concepts and perspectives,” *IEEE Antennas and Propagation Magazine*, Vol. 49, No. 6, pp. 24-39, IEEE, 2007.

[16] Priebe, S., Jastrow, C., Jacob, M., Kleine-Ostmann, T., Schrader, T., Kurner, T., “Channel and propagation measurements at 300 GHz,” *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 5, pp. 1688-1698, IEEE, 2011.

[17] Texas Instruments, “WLAN channel bonding: Causing greater problems than it solves,” *White Paper*, 2003.

[18] Okada, K., Minami, R., Tsukui, Y., Kawai, S., Seo, Y., Sato, S., Musa, A., “A 64-QAM 60GHz CMOS transceiver with 4-channel bonding,” *IEEE Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pp. 346-347, IEEE, 2014.

[19] Wu, R., Kawai, S., Seo, Y., Fajri, N., Kimura, K., Sato, S., Liu, B., “A 42Gb/s 60GHz CMOS transceiver for IEEE 802.11 ay,” *IEEE Solid-State Circuits Conference (ISSCC)*, pp. 248-249, IEEE, 2016.

[20] Eissa, M., Malignaggi, A., Panic, G., Lopacinski, L., Kraemer, R. and Kissinger, D., “Modular Wideband 1-15 GHz Transmitter Channelizer for High Data Rate Communication”, *11th Global Symposium on Millimeter Waves (GSMM)*, 2018.

LimeRFE – A Software Definable RF Front-End Module for SDR Platforms

Branko Bukvić, Mihajlo Božović, Zako Arsović, Dušan Grujić, Milan Savić

Abstract – This paper presents a software-definable RF front end module for SDR platforms. The Lime RF Front End (LimeRFE [1]) is an open hardware front end module with appropriate amplification and filtering [2], as well as supporting circuitry to augment the SDR platforms, providing a complete solution that addresses real life applications ranging from HAM (radio amateur) radio to standards-compliant cellular network implementations. Measurement result examples are also presented.

Keywords – Software Defined Radio, HAM Radio, Software-Definable RF Front-End.

I. INTRODUCTION

Radio frequency (RF) front-end provides the essential functionality for any radio system, by providing the amplification and filtering which are necessary in real-life applications. There can't be a completely software defined radio system without the software configurable, flexible RF front-end. LimeRFE provides exactly this functionality, a software-definable RF front-end module for SDR platforms. It supports most of radio amateurs bands (HF band, 6 m, 4 m, 2 m, 1.25 m, 70 cm, 33 cm, 23 cm, 13 cm and 9 cm bands), several cellular bands (Band 1, 2, 3, 7 and 38) and, additionally, the whole frequency spectrum from 1 MHz up to 4 GHz is covered with 2 wideband modules installed on the board (1 – 1000 and 1000 – 4000 MHz bands). The LimeRFE board comprises Software Defined Radio concept [3]. It relays on versatile hardware which is capable to operate under the several different standards, which means it can support different frequency bands, operating powers, modulation schemes, etc.

Section II gives a brief introduction of the LimeRFE board. Basic hardware schematics and software controlling tools, together with some implementation details, are explained in this section. Measurement result examples are presented in Section III, including 4G signal results. Also small signal gain (Gain), noise figure (NF), intermodulation product of 3rd order (IP3) and characterization of 1dB compression point (P1dB) were performed for the receiver and transmitter paths of the LimeRFE. Finally, a conclusion was given in Section IV.

II. DESCRIPTION OF THE LIMERFE BOARD

A. Hardware

A picture of the LimeRFE board is given in Fig. 1. The board dimensions are 125 x 162.5 mm.

Fig. 1. Picture of the LimeRFE board

Table I gives all frequency bands supported by the LimeRFE.

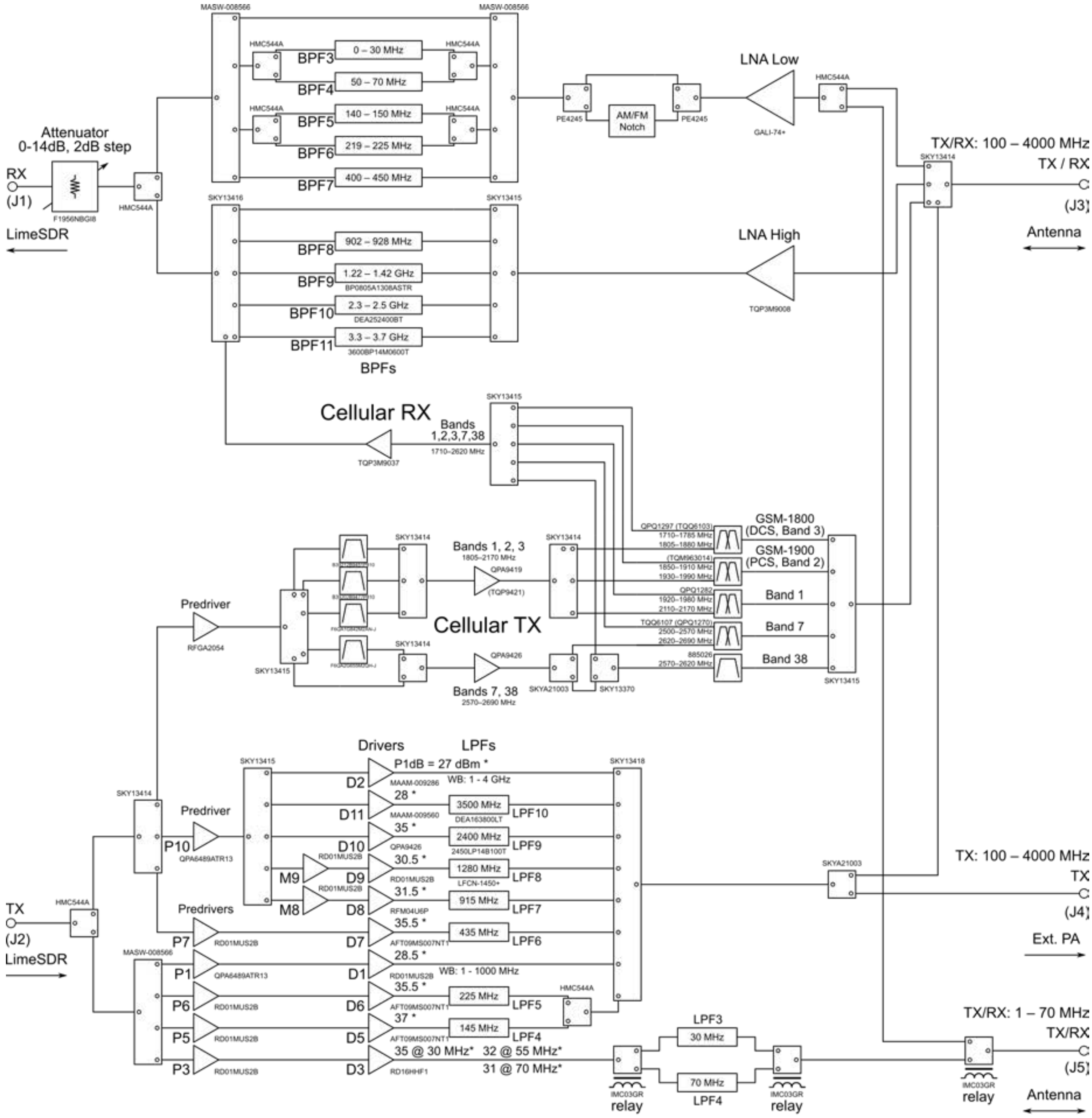
TABLE I
LIMERFE FREQUENCY BANDS

HAM		**Cellular**	
Band	*Freq.* [MHz]	*Band*	*Frequency* [MHz]
HF	1.8 – 30	Band 1*	2110 – 2170 / 1920 – 1980
6 & 4 m	50 – 70	Band 2*	1930 – 1990 / 1850 – 1910
2 m	144 – 146	Band 3*	1805 – 1880 / 1710 – 1785
1.25 m	220 – 225	Band 7*	2620 – 2690 / 2500 – 2570
70 cm	430 – 440	Band 38	2570 – 2620
33 cm	902 – 928	**Wideband**	
23 cm	1240 – 1325	*Band*	*Frequency* [MHz]
13 cm	2300 – 2450	WB 1000	1 – 1000
9 cm	3300 – 3500	WB 4000	1000 – 4000

* Freq. for cellular bands 1, 2, 3 and 7 are in format Downlink / Uplink

Block diagram of the LimeRFE RF part is shown in Fig. 2. HAM and Wideband transmitters are grouped in the bottom, cellular transceivers are located in the centre, while the upper part presents the receiver paths.

Authors are with Lime Microsystems d.o.o, Bulevar Mihaila Pupina 10A/I, 11070 Belgrade, Serbia. E-mail: b.bukvic@limemicro.com



* Note: Output 1 dB compression values are preliminary assesment only. Could be changed. Given values are at the drivers' outputs. Output values will be lower.

Fig. 2. Block diagram of the LimeRFE RF part

PCB stackup is given in Fig. 3, bellow. Top and bottom substrates are Rogers 4350B, with thickness of 500 µm (20 mil). Rogers 4350B was preferred over standard FR4 due to better control of thickness and substrate properties, as well as lower dielectric loss at higher frequencies. Metalization thickness was not of great importance for this design, so common thickness of 18 µm (0.5 oz) was used. All RF lines are realized as grounded coplanar waveguides. Between top and bottom layers, there are three 200 µm thick RF4 substrates, enabling total of 6 layers. Top and bottom two layers are used mainly for RF lines, while the inner two layers are used for digital signal routing and supply distribution. All vias are implemented as filled through-hole. The total board thickness is 1.6 mm.

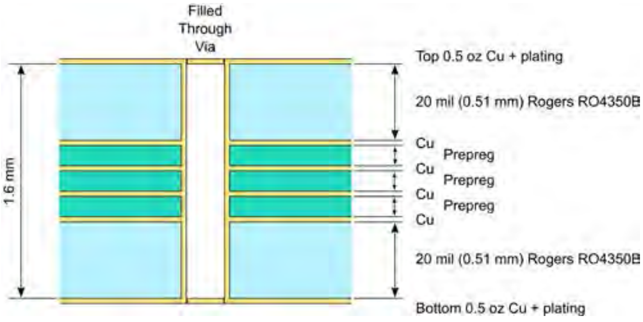


Fig. 3. Stackup of the LimeRFE board

Components are placed on both sides of the board, while the large heatsink occupies the most of the bottom side.

LimeRFE board includes the SWR meter subsystem (Fig. 4). Forward and reflected signals are to be provided from the external directional coupler. There is an option to measure power at the output of the power amplifiers in the cellular bands using on-board couplers.

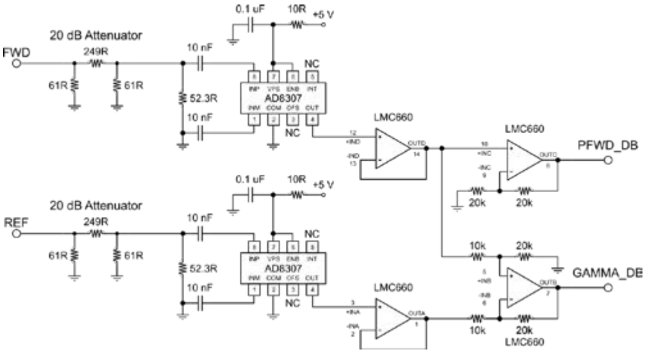


Fig. 4. LimeRFE SWR subsystem

Complete board was designed so that standard off-the-shelf components are used. However, the performance in HF bands, as well as in 6 and 4 m bands, can be further improved by using add-on balun board [4] shown in Fig. 5.

Fig. 5. Picture of LimeRFE add-on balun board

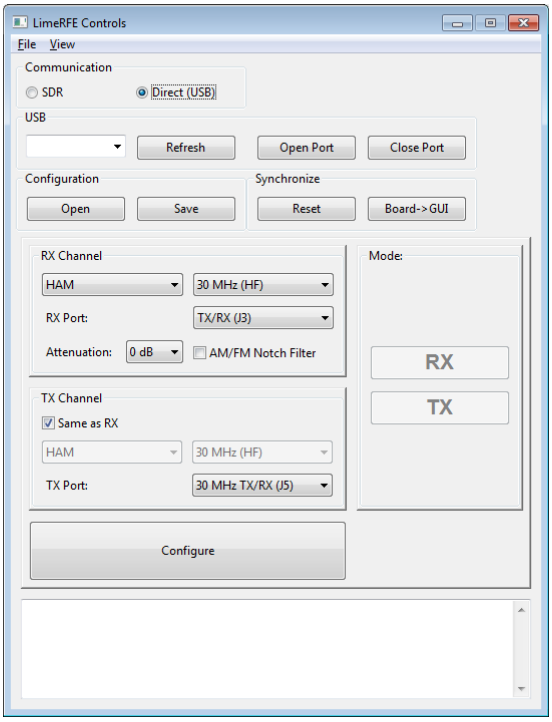
Next hardware block consists of digital circuits for controlling RF parts (select desired frequency band, turn on/off particular power amplifier and LNA, select appropriate filter, SMA connector, etc.). LimeRFE utilizes Arduino Nano – compatible microcontroller (ATmega328), and is programmable via Arduino IDE [5]. MCU (by data, clock and reset signals) drives 11–bytes shift register which stores the control bits. Control of the shift register requires only few of the MCU IO (input/output) pins, while others are used for time-critical controls. Voltage level shifters, USB communication circuit are also part of this hardware block.

Another part of the LimeRFE, denoted as GPIO subsystem, provides level shifters and relay drivers intended for communication and control of other parts of the complete radio system (e.g. external power amplifiers) by the SDR platform, making the use of an additional board with this functionality [6] obsolete.

Bias subsystem comprises several DC–DC convertors and voltage regulator units providing required supplies (2.5 V, 3.3 V, 5 V, 7.5 V and 12 V).

B. Controlling Software Tools

LimeRFE is supported by the LimeSuite [7], an open-source multi-platform software for control of the Lime Microsystems' SDR platforms. Library with API functions for control automation is also provided.



Communication: ○ SDR ● Direct (USB)

RX Channel: HAM, 30 MHz (HF)

RX Port: TX/RX (J3)

Attenuation: 0 dB ☐ AM/FM Notch Filter

TX Channel: ☑ Same as RX; HAM, 30 MHz (HF)

TX Port: 30 MHz TX/RX (J5)

Fig. 6. LimeRFE controls GUI module

Further details of the LimeRFE are available in [8].

III. MEASUREMENT RESULTS EXAMPLE

Measurements were performed on the test-bench shown in Fig. 7. The test-bench consists of Agilent E8267D signal generator, Agilent E4440A spectrum analyser, Rigol DP832 DC supply unit, Agilent 346A noise source, and a PC. Measurements were automated using Python scripts.

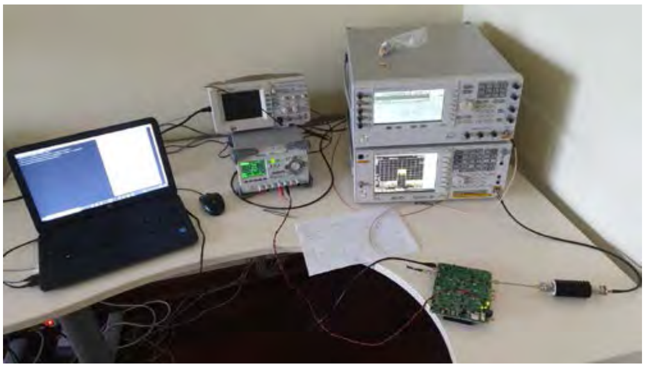


Fig. 7. Picture of the test-bench with the LimeRFE as device under test (DUT)

Small signal gain measurements were performed on receiver and transmitter paths of each of the bands of the LimeRFE board. Measurement results for 430 – 440 MHz (70 cm) HAM band receiver are given in Fig. 8.

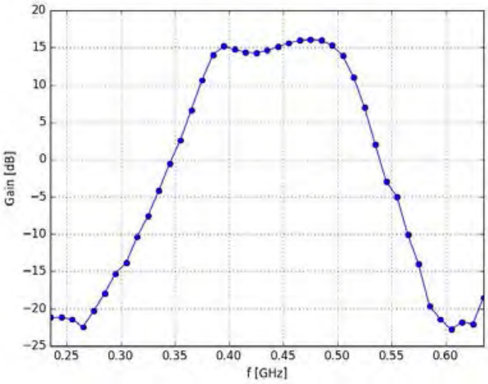


Fig. 8. Gain measurements for receiver in HAM 430 – 440 MHz (70 cm) band.

NF measurements were performed for receiver paths in each of the bands. For these measurements, additional Agilent 346A noise source was used as noise reference. Measurement results for receiver in HAM 900 – 928 MHz (33 cm) band are given in Fig. 9.

Noise figure measurements for each frequency point were performed twice, in order to estimate measurement accuracy.

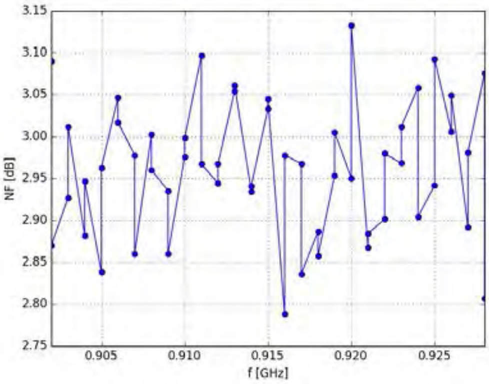


Fig. 9. NF measurements for receiver in HAM 902 – 928 MHz (33 cm) band.

IP3 measurements were performed for receiver paths in each of the bands. Measurement results for receiver in Cellular Band 7 are given in Fig. 10.

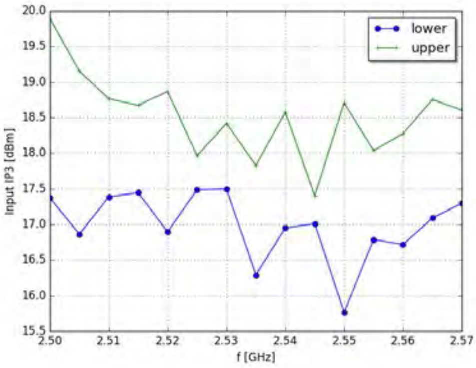


Fig. 10. IP3 measurements for receiver in Cellular Band 7.

P1dB measurements were performed on transmitter paths of each of the bands. Fig. 11 presents Gain vs input power, while Fig. 12 gives output power vs input power for HAM 2 m band. In Fig. 13 measurement results of output power at 1 dB compression point and saturation power vs frequency are presented.

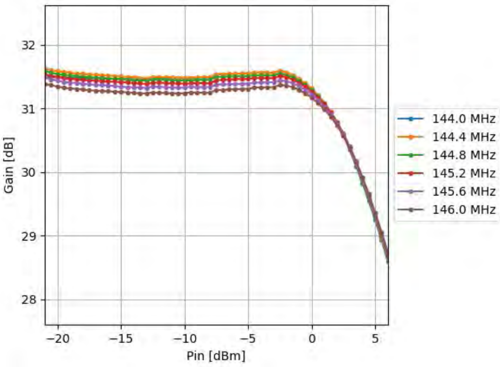


Fig. 11. Gain vs Pin measurements for different frequencies, for transmitter in HAM 145 MHz (2 m) band.

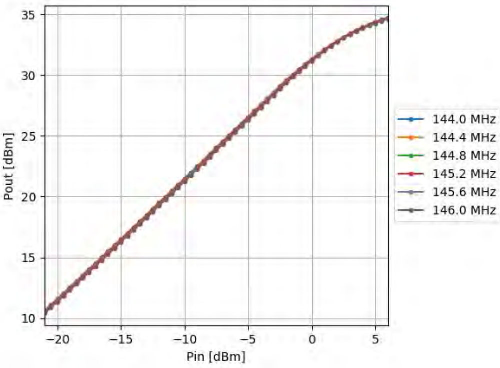


Fig. 12. Pout vs Pin measurement results for different frequencies, for transmitter in HAM 145 MHz (2 m) band.

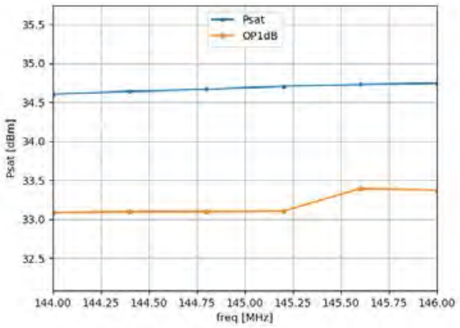
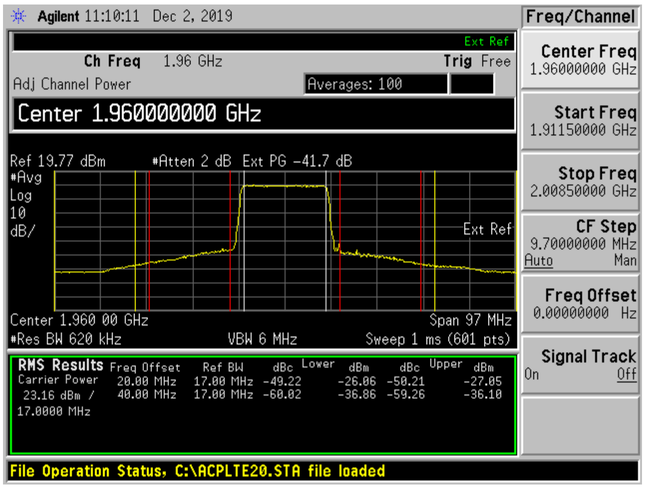


Fig. 13. Output power at 1 dB compression point and saturation power vs frequency measurement results for transmitter in HAM 145 MHz (2 m) band.

Measurements with LTE signals were performed for each of the cellular bands. Adjacent Channel Power Ratio (ACPR) measurements for Cellular Band 2 are shown in Fig. 14. Test signal LTE TM3.1 with 20 MHz bandwidth was used. It can be noted that for the modulated output power of 23 dBm ACPR is better than -49 dBc.


```
Agilent 11:10:11   Dec 2, 2019
                                                      Ext Ref
        Ch Freq    1.96 GHz                        Trig  Free
Adj Channel Power                    Averages: 100
Center 1.960000000 GHz

Ref 19.77 dBm       #Atten 2 dB  Ext PG -41.7 dB
#Avg
Log
10
dB/                                               Ext Ref

Center 1.960 00 GHz                              Span 97 MHz
#Res BW 620 kHz             VBW 6 MHz     Sweep 1 ms (601 pts)

RMS Results Freq Offset  Ref BW   dBc  Lower  dBm    dBc  Upper  dBm
Carrier Power  20.00 MHz 17.00 MHz -49.22     -26.06 -50.21      -27.05
 23.16 dBm /   40.00 MHz 17.00 MHz -60.02     -36.86 -59.26      -36.10
17.0000 MHz

File Operation Status, C:\ACPLTE20.STA file loaded
```

Fig. 14. Cellular Band 2 ACPR measurement results

IV. CONCLUSION

This paper presented the LimeRFE board – a software-definable RF front-end module for SDR platforms. Since it follows the software defined radio paradigm, it is easily reconfigured for operation in desired frequency band. This board can be connected to any of the Lime microsystems' SDR family platforms, as well as with any SDR platform in general, providing a full radio solution for numerous applications, from HAM radio to standards-compliant cellular networks.

Combination of a general purpose computing platform, SDR capable transceiver, and open-source software-definable RF front-end will provide open, compact, and affordable all-round radio platform. With proper software support from the community it will provide unprecedented benefits. By combining LimeNET-Micro [9] with LimeRFE such a platform is readily available (Fig. 15).

Fig. 15. LimeNET Micro & LimeRFE – all-round, open, compact and affordable radio platform

REFERENCES

[1] Lime Microsystems, "LimeRFE – A Software-Definable RF Front End Module for LimeSDR Platforms", available at: https://www.crowdsupply.com/lime-micro/limerfe.

[2] Lazović, S., Litovski, V., "Osnovi elektronike [Fundamentals of Electronics]", *Čuperak plavi*, Niš, 1996. [In Serbian]

[3] F. M. Ghannouchi, "Power amplifier and transmitter architectures for software defined radio systems", in *IEEE Circuits and Systems Magazine*, vol. 10, no. 4, pp. 56 – 63, Fourthquarter 2010.

[4] H. Arnold – DL2EWN, "Breitbandige QRP-Linearendstufe mit HF-Leistungs-MOSFET", *Box73 Amateurfunkservice GmbH* 2008. [In German]. Available at: http://www.dj7th.de/download/10W%20PA.pdf

[5] "Arduino IDE", available at: https://www.arduino.cc/en/main/software

[6] Lime Microsystems, "LimeSDR GPIO Board", available at: https://wiki.myriadrf.org/LimeSDR_GPIO_Board

[7] Lime Microsystems, "LimeSuite – Driver and GUI for LMS7002M-based SDR platforms", source code available at: https://github.com/myriadrf/LimeSuite

[8] Lime Microsystems, "LimeRFE Quick Starter Manual", 2019. (URL address is to be defined)

[9] Lime Microsystems, "LimeNET Micro", available at: https://wiki.myriadrf.org/LimeNET_Micro.

Cogita by Vtool – A new approach to simulation debugging

Hagai Arbel

Abstract – This paper describes Cogita – a simulation debugging platform developed by Vtool. It describes a real-life case study that was conducted by one of the customers in order to measure Cogita's efficiency vs. traditional debugging methods.

Keywords – Simulation, Debugging, log-files, Data Analysis.

I. ABOUT VTOOL'S COGITA

Vtool has produced a new, next-generation debug solution named Cogita. [1]

Cogita takes an abstract, visual approach to debug that solves many of the issues of failure analysis and debug. It incorporates state-of-the-art ergonomic, visualization technology, makes use of psychological studies into problem-solving and large scale data analysis, and applies modern Machine Learning (ML) algorithms in order to examine large data sets in a cognitive manner.

Cogita has been proven to accelerate the debug process anywhere from 3X to 10X compared to the alternatives, log files and wave forms debugging or interactive software debugging.

Cogita can be applied to large-scale block verification, particularly tuned to operate with complex UVM testbenches. It also brings unique debug capabilities in tracking down intricate corner cases from SoC verification runs.

A. Improvements to the debug process

Cogita applies the following improvements to the debug process:

- Reduces the chance of taking an incorrect assumption for granted.
- Supporting the process of asking the right questions.
- Validating assumptions and answering questions faster.
- Revealing possible paths that are otherwise easily overlooked.

B. Key benefits

The key benefits of Cogita are:

- Accelerates debug time, up to an order of magnitude, for a broad range of complex bug types. Given that debug represents 25% of the entire development time of a semiconductor, this represents a huge resource-saving and time-to-market advantage.
- Improves design and verification quality over and above coverage assessment through visibility into verification scenarios, allowing a clear understanding of convoluted design code for easy team communication and cooperation.
- Extends debug for large-scale systems on hardware-software co-verification. [2]

II. BACKGROUND AND PROBLEM DEFINITION

Company-X develops complex ASICs and FPGA that facilitates massive traffic from CPUs to memory. The main module of the system (the DUT) and the UVM testbench are illustrated in Fig. 1.

Company-X verification team assessed from previous projects that such subsystem verification will take 10 months for 4 engineers.

Below we explain the main functionality of the DUT and testbench and the challenges in debug that Cogita, later on, was able to solve. The DUT, in essence, accepts READ/WRITE layered packets via the YBUS OB UVC. When these packets are aimed at the AXI4 interface on the right-hand side, they are called DIRECT. The DIRECT datapath is the simpler one but it has many complications on its own:

- Each read or write packet contains many fields that are later on translated to AXI4 bursts. The AXI4 UVC Slave responds to the packets and the DUT processes the response and sends out a YBUS response to the YBUS IB UVC.

Hagai Arbel is with Vtool DOO, Milutina Milankovica 1G, Novi Beograd. E-mail: hagai@thevtool.com.

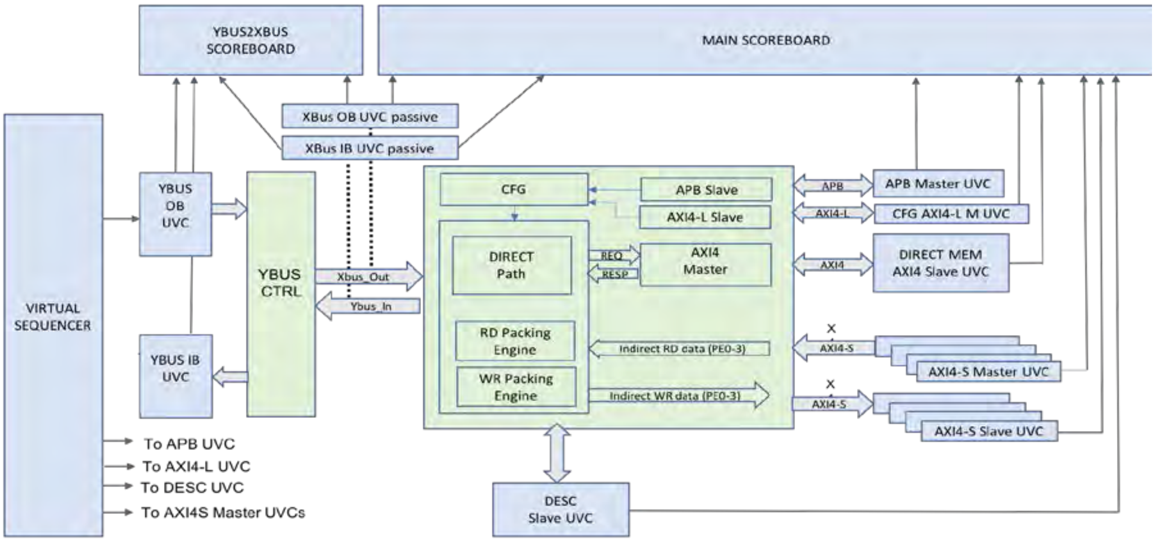


Fig. 1. The DUT and testbench

- There is a complex logic that transforms YBUS packets to AXI4 bursts. This also depends on the DUT configuration and Lookup Tables that are injected dynamically via the APB and the AXI4-L UVCs.
- The whole process is executed interleaved for 16 channels.

In parallel to the DIRECT path (and interleaved over the YBUS OB and IB interfaces), the INDIRECT datapath operated. This path is even more complex:

- In the WRITE INDIRECT path, the data that arrives in from YBUS OB UVC is layered to L4 packets, L2 packets, and YBUS Requests. Descriptors that are read from the DESC UVC, are telling the DUT how to pack the data and send it over the AXI-Streaming four interfaces.
- Similar mechanisms are used on the read side.

There is another datapath from the APB to the YBUS IB and back.

A fully random test in this testbench has a matrix of processed data that is almost impossible to track:

- Six different datapaths
- Three layers of packets and complex transform mechanism to/from the memory system.
- 16 parallel and interleaved channels.

Waveforms are almost useless in order to grasp what is happening in the test and in the testbench because they only show the DUT's signals. Without Cogita, the only way to analyze the sequences and the scoreboard is the log file.

In order to try and debug these complex tests, the team added a lot of UVM messages. But even then, they ended up with log files of 20 million lines or more, 2-5G of data.

Every failing test triggered many questions and challenges:

- What happened in this test? Which channels were active, which datapaths, what was the configuration at any moment?
- Was the layered data generated from the YBUS UVC legal? Did the descriptors interact well with the data?
- How to correlate the events in the waveforms to the log file?
- How to be able to track and analyze the data from the log, while still holding the bigger picture of the test scenario?

The team reported that a complex debugging session can take hours and sometimes days when you include the communication with the RTL design team. In some cases, a failure switched hand between design and verification engineers for a week until the root-cause (either testbench issue or RTL bug) was found.

III. CASE STUDY

In order to test the efficiency of Cogita solving the above problem, the team conducted a case study.

A. *Saved player configuration*

First, they created a player configuration in Cogita that will show in one image the "story" of the test. Before the actual debugging of each failing test, this image was loaded in order to understand what happened.

Below are several examples of screenshots (Figs. 2 and 3) and the conclusion the team made.

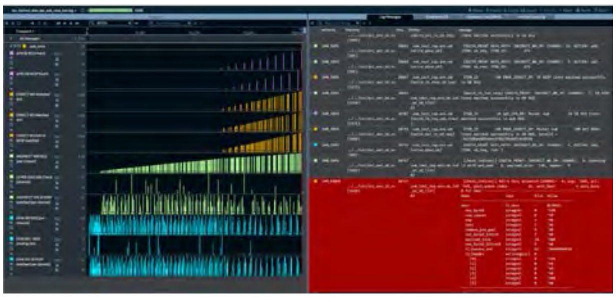


Fig. 2. Saved player configuration

- The test runs for 20,747 ns and then fails with 13 errors, the first one is INDIRECT WR datapath, channel 0, data integrity error.
- The APB datapath (purple players) is active both on request and response sides and has only 10 packets to that point.
- The DIRECT WR and RD paths (orange players) are active on request, AXI and response interfaces and have 108 and 109 items respectively.
- The INDIRECT WR path (green players) has 376 items till the failures, seems that all channels till 13 are active, AXI-S packets are checked properly till the failure point.
- The INDIRECT RD path (blue players) seems to operate well.

Since the failure is in the INDIRECT WR path, channel 0, the next player configuration (also saved by the team) provides a more relevant picture of the failure.

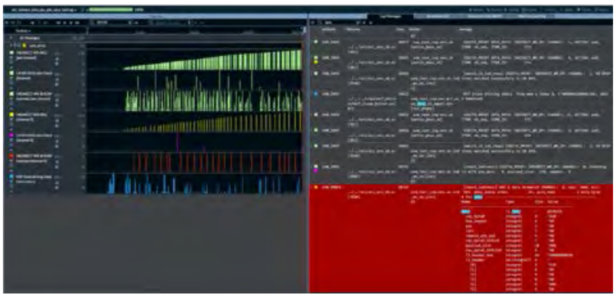


Fig. 3. Indirect WR path players

Now we can see channel 0 over the yellow player and that the failure happens in the third AXI-S check (the purple player).

B. Debugging a failure with Cogita

In this example, after a few good transactions, errors begin to appear, reporting data mismatch on the AXI port. This could mean wrong sampling by the AXI monitor, mismatch of expected vs. collected data in the scoreboard and actually, many other things. We do not know at this point.

```
UVM_ERROR ../../src/acc_env_sb.sv(1540) @ 1563 ns:
uvm_test_top.env.sb.ind_wr_sb_list[#5] [match_ind_wr_axis]
AXI-S data mismatch in byte 0 CHANNEL:  5, exp:
'h8043353c28db739f9a048e9f879c5fcd, act:
'h78efee8e652372f58403525e1363046f
```

<u>Step 1: Open the log file in Cogita</u>

Loading the log file into Cogita shows the basic view in Cogita (Fig. 4).

- Left side - A timeline with bars representing messages, the higher the bar, the more messages there are at this time. The red cursor is placed at the first error by default.
- Right side – Messages from the log, with the scroller placed in the time point the cursor is at.

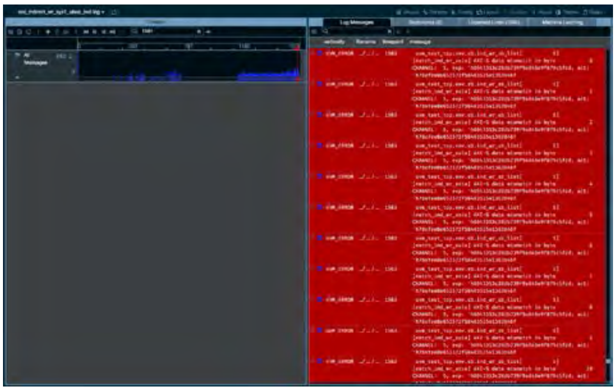


Fig. 4. Open the log in Cogita

<u>Step 2 – Validating correct sampling by the monitor</u>

Question: Is the reported “actual” value really the output of the DUT or is there a sampling problem with the monitor?

Answering with Cogita is easy because the engineer does not have to switch between viewing the log file in an editor and the waveform. Cogita can read data from multiple sources, waveform database included (vcd, trn, vpd, fsdb) and to jointly present them over a single timeline. Fig. 5 clearly shows the error message, the relevant signals of the AXI-S I/F.

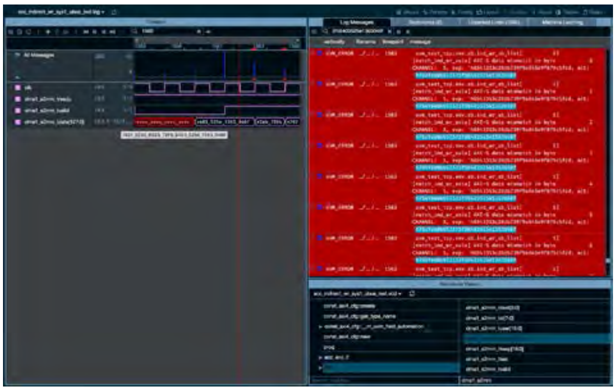


Fig. 5. Looking at waveforms

Answer: Yes, Is the reported "actual" value really the output of the DUT.

Step 3 – Looking for the actual and expected data at the input

Question1: Was the expected data sent over the input to the DUT and when?

Question2: Was the actual data sent over the input to the DUT and when?

Answering each of the above questions is extremely confusing without Cogita. One would have to search the log, for two different values, open the waveform viewer, look again, extracting the time point, etc. Not that it is impossible; It is time-consuming and more importantly, prone to the kind of mistake that send one back to the debug starting point.

Fig. 6. Looking at "expected" and "actual" players

With Cogita, we simply create two new "Players", as shown in Fig. 6.

Answer: Both expected and actual data were sent over the input. Actual (Fig. 6, Brown player) first at 1215 ns and Expected (Fig. 6, Green player) at 1267 ns.

Step 4 – Analyzing the data channels

So now, we know that both expected and actual data from the error were sent into the DUT.

Question: To which channel each data belongs?

To answer this question with Cogita, we set the Height of the bars to represent values of interest from within the message.

As shown in Fig. 7, each player's Height represents the channel number. It is clearly visible that expected data (starts with h'8043) was sent over the input with channel 5, while the actual data (starts with h'78ef) has channel number 4. Since the UVM_ERROR at the end has a height of 5 (it checks channel 5), we conclude that the DUT had switched the data from channel 4 onto the output of channel 5.

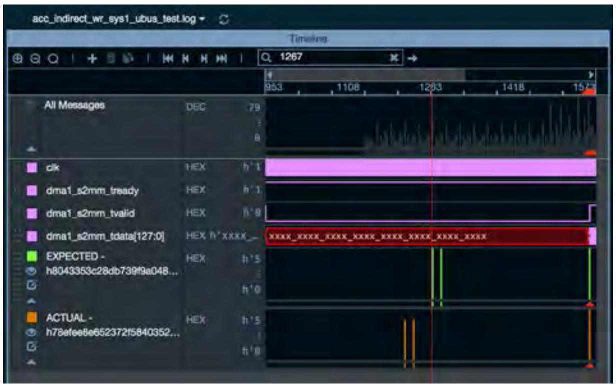


Fig. 7. Looking at Channels 4 and 5

Answer: Data h'78ef belongs to channel 4 but sent out over channel 5. Data h'8043 belongs to channel 5 and is lost. This is an RTL bug.

C. Debug process summary

In this case study, the debug process with Cogita included three steps.

Cogita helped the process in several ways:

1. Each answer leads to the next meaningful question.
2. Finding the answers (or validating the assumptions) is much faster.
3. The chances of taking a wrong turn, restarting the debug process all over, is almost impossible using Cogita.

IV. COMPARISON

This chapter shows a comparison between Cogita and two traditional debugging methods, both use Synopsys debugging tools. The first is the DVE waveform viewer and the second is Verdi, step-by-step debug.

The team had created the tables I and tables II-IV in order to assess Cogita's efficiency. Cogita accelerated the debug process in many aspects. When it comes to junior engineers and complex systems, Cogita has an added value of helping in understanding the system and the test scenarios.

TABLE I
PROJECT PROPERTIES

Category	Project attribute
Company	Company X
Design type	Packet Manager
Team	4 verification engineers, 3 designers
Device Under Test (RTL) size	25,000 RTL lines
Test Bench (System Verilog) size	100,000 System Verilog UVM lines

TABLE II
COMPARISON BETWEEN COGITA AND TRADITIONAL DEBUGGING TOOLS - DVE

Category	DVE Waveforms + Log Files (Synopsys)
Tool ramp-up	None - We assume everybody knows how to use them
Ability to understand immediately what the test generally did	Impossible
Time for the test case debug	Assessing over one hour
Debugging speed factor for senior engineer - On average	1
Debugging speed factor for junior engineer - On average	1
Helping communication within the team	Limited - attaching a waveform snapshot to the bug description

TABLE III
COMPARISON BETWEEN COGITA AND TRADITIONAL DEBUGGING TOOLS - VERDI

Category	Verdi step-by-step (Synopsys)
Tool ramp-up	We already knew this but it took the team a few months to learn
Ability to understand immediately what the test generally did	Impossible
Time for the test case debug	Assessing we would not use Verdi - It would not be efficient for this failure
Debugging speed factor for senior engineer - On average	1.5X
Debugging speed factor for junior engineer - On average	1.5X
Helping communication within the team	Limited - attaching a waveform snapshot to the bug description

TABLE IV
COMPARISON BETWEEN COGITA AND TRADITIONAL DEBUGGING TOOLS - COGITA

Category	Cogita (Vtool)
Tool ramp-up	Two weeks for the basic level. One month for very high efficiency
Ability to understand immediately what the test generally did	Very easy
Time for the test case debug	10 minutes
Debugging speed factor for senior engineer - On average	5X
Debugging speed factor for junior engineer - On average	10X
Helping communication within the team	Helps a lot - We used Cogita in bug reports and communication

IV. CONCLUSION

This case steady demonstrates Cogita's efficiency in finding bugs during pre-silicon ASIC simulations. By applying new approach to data analysis, it helps verification team finding bugs faster than ever before.

This entire debug session took 10 minutes of work. The engineer who conducted this had only a few weeks of experience using Cogita. The engineer assessed that it would take at least five times to debug this error without Cogita.

REFERENCES

[1] Anna M. Ravitzki, Uri Feigin, Hagai Arbel, "*Between the Dialog and the Algorithm or Innovative Technological Narratives Leveraging the Idea of Authenticity in a Human Being. Visualizing log files enables intuitive comprehension of complex test scenarios.*" DVCON, Munich, 2017.

[2] Anna M. Ravitzki, Uri Feigin, Hagai Arbel, *"The Big Data Revolution - Beautiful Servant or Dangerous Monster?*" DVCON, Shanghai, 2018.

OTA-C Filter Synthesis Based on Existing LC Solutions

Vančo B. Litovski, Jelena Milojković, and Miljana Milić

Abstract: The problem of synthesis of high frequency analog integrated CMOS circuits is visited. Theory is developed and implementation demonstrated for synthesis of analog OTA-C (or GM-C) filters based on existing low-pass LC solutions. Instructive example is given produced by the *RM* software for filter design. This report may be considered as a case study supported by a rather complex design example.

Keywords – Filter design, OTA-C filters, Gm-C filters, CMOS integrated filters, High frequency filters, Cascade circuit synthesis.

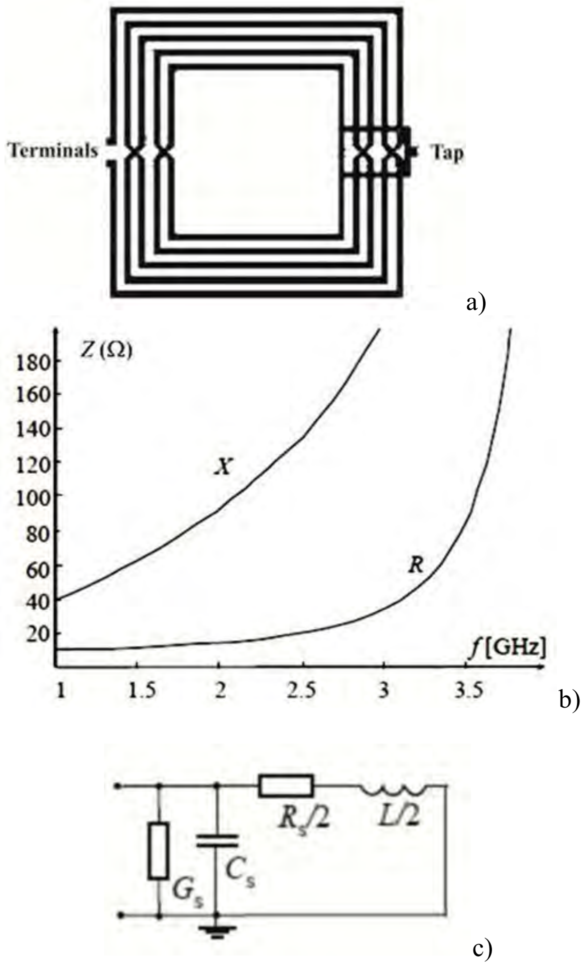
I. INTRODUCTION

One of the problems encountered in high frequency analog integrated filter production is the area needed to produce an inductor. It is realized in a form of a flat spiral line the inductance of which is limited not only by the area but by its huge parasitic capacitance. One such inductor realized in CMOS was reported in [1].

The most important electrical parameters of an integrated inductor are the inductance (L), its resistance (R), its parasitic capacitance and its Q-factor (Q) as a secondary parameter. The layout of one inductor of this kind is depicted in Fig. 1a. One may see that the wires are twisted to reduce the parasitic capacitance. This inductor is specific in the sense that it has a tap terminal allowing specific uses. Fig. 1b depicts the dependence of the reactance and the resistance of the integrated inductor on the signal frequency which is an additional problem when designing filters with this kind of components. The numerical values of the equivalent circuit depicted in Fig. 1c, at f=2.43 GHz, are i=6 nH, R_S=9.3 Ω, C_S=220 fF and G_S=0.2 mS. The value of the Q-factor is Q=8.33.

To avoid such a component, attempts were made to find an active circuit that simulates the inductance good enough to be implemented in high frequency CMOS integrated circuits.

Vančo Litovski is independent researcher. email: vanco.litovski@elfak.ni.ac.rs
Jelena Milojković is with the Innovation Centre of Advanced Technologies, Niš, Serbia. email: jelena.milojkovic@intt.rs
Miljana Milić is with the Faculty of Electronic Engineering, University of Niš, Serbia, email: Miljana.Milic@elfak.ni.ac.rs



a)

b)

c)

Fig. 1. Planar integrated inductor. a) layout, b) Frequency dependence of the impedance, and c) Equivalent circuit of the half of the inductor

There are many solutions offering a simulated inductor using active components and capacitors. These are based on the concept of NIC (negative impedance converter) which will be not elaborated further here as a general circuit theoretical issue. The main difference in all such circuits if high frequencies are planned to be used is related to the active component. Namely, the CMOS operational amplifier (OA) performance is not good enough for these purposes so that the OTA (operational transconductance amplifier) took over. Based on that solutions were found for the simulated inductances [2][3] of which we will later on elaborate the one based on gyrators [4] and described in [5].

The OTA itself is not perfect, too. Its main characteristics are the frequency dependence of its transconductance and output capacitance. Ideally, one would like to have a perfect OTA which means a component with zero valued output capacitance and frequency independent controllable transconductance. There is no such perfect component, however, despite the fact that improvements are reported almost on daily basis [6]. As an example Fig. 2 depicts the frequency dependence of the transconductance of an OTA [7] obtained by simulation while Fig. 3 depicts the corresponding output impedance. As can be seen high frequencies are reached (cut-off frequency is claimed to be 567 MHz) the transconductance being much more frequency independent than the output impedance. Unfortunately neither the transconductance nor the output impedances is given in absolute values so one is not capable to extract further conclusions.

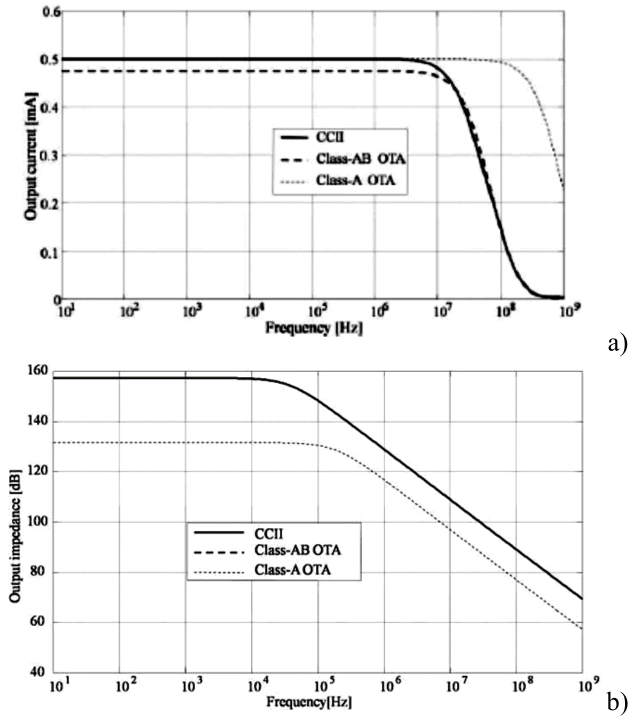


Fig. 2. a) Transconductance and b) Output impedance (modulus) as a function of frequency [7] (Courtesy of the authors)

The fundamental idea of implementation of simulated inductance is based on the availability of LC cascaded circuits which are synthesized by some other filter synthesis software system or even extracted from an existing catalogue such as [8]. In that way the inductors are substituted by an equivalent circuit containing OTAs and a capacitor and the rest of the filter elements (capacitors) remain the same. That, of course, is a very attractive method and even designer with extremely limited knowledge of filter design can produce successful designs.

In this paper we will review very briefly the equivalent circuit to an inductor and demonstrate how the synthesis is functioning. Equivalent circuits to the corresponding LC cells realizing proper transmission zeros will be introduced. The method described here will be tested using ideal OTAs (infinite output impedance) to check for the synthesis process and not for the properties of the practical realization. The example given is rather complicated confirming the effectiveness of the synthesis method. It was produced by the *ℛℳ* software for filter design [9][10][11].

II. SIMULATING INDUCTORS BY GYRATORS

The fundamental building block which will be used to create the simulated inductor is the gyrator as depicted in Fig. 3.

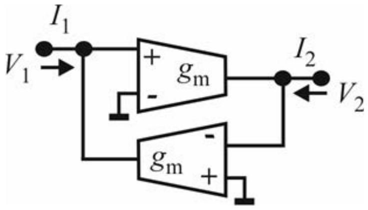


Fig. 3. Gyrator realized by a pair of OTAs

For this circuit the nodal equations are:

$$\begin{aligned} I_1 &= g_m \cdot V_2 \\ I_2 &= -g_m \cdot V_1 \end{aligned} \tag{1}$$

When loaded by an impedance Z_L, the output voltage will be:

$$V_2 = -Z_L \cdot I_2 \tag{2}$$

After substitution in (1) one gets

$$I_1 = g_m^2 \cdot Z_L \cdot V_1 \tag{3}$$

For $Z_L = 1/(j\omega C)$, one has

$$I_1 / V_1 = \frac{g_m^2}{j\omega C} = \frac{1}{j\omega L_e} \tag{4}$$

where

$$L_e = C / g_m^2 \tag{5}$$

A gyrator loaded by capacitor will behave as an inductor. Since the capacitor has one terminal grounded the resulting simulated inductor will be grounded, too.

It is up to the designer to decide whether to use a fixed value for the transconductance or for the capacitor in order to create the desired value of the inductance. Namely, the transconductance may be voltage controlled while the voltage digitally created which means that, in general, finer granulation may be achieved. The capacitance is programmed by connecting and disconnecting incremental capacitances already available on the chip. From the functionality point of view, however, there may be an additional consideration. Namely, one would prefer smaller transconductances since in that case the output impedance is expected to be larger which makes the component nearer to the perfect OTA. If the transconductance is chosen to be variable this opportunity will be not available.

A. Simulating a floating inductor

The schematic depicted in Fig. 4 represents a connection of two gyrators and a capacitor to produce a floating inductor. To show that we will write the nodal equations for the circuit as

$$\begin{aligned} I_1 &= g_{\mathrm{m}} \cdot V_C \\ I_2 &= -g_{\mathrm{m}} \cdot V_C \\ \mathrm{j}\omega C \cdot V_{\mathrm{C}} - g_{\mathrm{m}} V_1 + g_{\mathrm{m}} V_2 &= 0 \end{aligned} \tag{6}$$

After eliminating VC from the third equation and having in mind $I_1 = -I_2$, one gets

$$Z_{\mathrm{L}} = \frac{V_1 - V_2}{I_1} = \frac{\mathrm{j}\omega C}{g_{\mathrm{m}}^2} \tag{7}$$

This means that the circuit of Fig. 4 behaves as a floating inductor of inductance

$$L_{\mathrm{e}} = C / g_{\mathrm{m}}^2 \tag{8}$$

Note, at ω=0, from (6) one gets $V_1 = V_2$ which corresponds to real inductor. The voltage V_{C} is undefined and so are the currents I_1 and I_2. The last two will be defined by the outer circuit as is the case with the inductor.

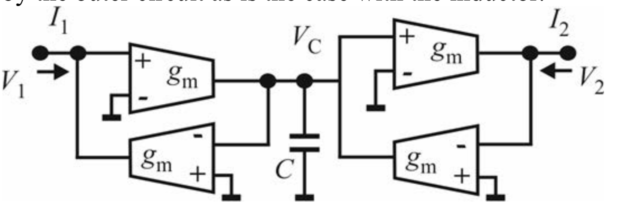


Fig. 4. Floating inductor realized by two gyrators and a capacitor

B. Simulating an ideal transformer

To produce a simulated transformer based on gyrators one may use the circuit of Fig. 5.

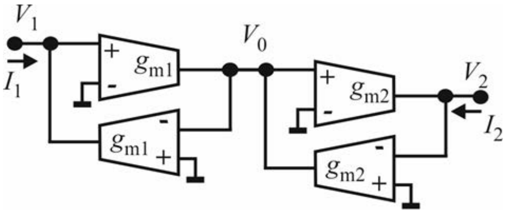


Fig. 5. Transformer simulated by two gyrators

Following are the nodal equations of this circuit

$$\begin{aligned} I_1 &= g_{\mathrm{m1}} \cdot V_0 \\ I_2 &= -g_{\mathrm{m2}} \cdot V_0 \\ -g_{\mathrm{m1}} V_1 + g_{\mathrm{m2}} V_2 &= 0 \end{aligned} \tag{9}$$

The transformer's equation is now

$$\frac{V_2}{V_1} = \frac{g_{\mathrm{m1}}}{g_{\mathrm{m2}}} \tag{10}$$

Note the transformer depicted here has one terminal of both input and output, grounded. That is acceptable for the implementation of the cell realizing complex transmission zero (D-section) depicted in Fig. 14.6b.

Negative "turn-ratio" may be obtained by inverting (interchanging the input terminals of) both transconductance amplifiers in one of the gyrators.

III. GM-C CIRCUIT SYNTHESIS

The circuit synthesis of this kind of filters is straightforward. One is to synthesize first an LC filter using a conventional synthesis procedure using a conventional synthesis program e.g. *cascade_LP* of the *RM* software. Of course, one may use data from catalogues [8] too. The next step is to substitute the inductors and, if necessary, transformers with their models using OTAs. This stem is performed by the *GM_LC* program of the *RM* software. Here we demonstrate the equivalent circuits. A limited set is given to save space. Nevertheless, this set is satisfactory for most physical realizations especially when low-pass circuits are sought.

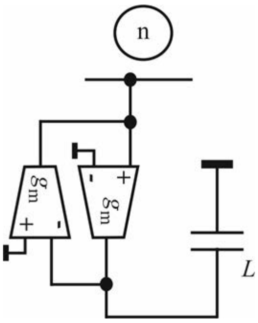


Fig. 6. A grounded simulated inductor

Fig. 6 depicts the equivalent circuit to the grounded inductor. As can be seen from now on the transconductance is considered a constant while the value of the capacitance is evaluated from (5) to be

$$C = L \cdot g_m^2 \tag{11}$$

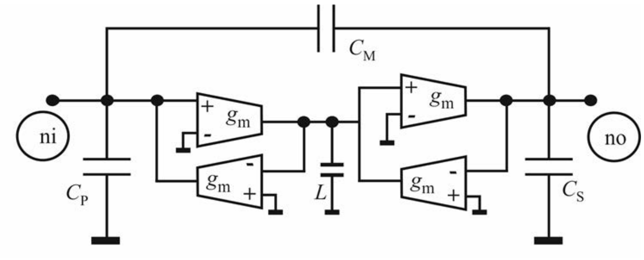


Fig. 7. A cell realizing a transmission zero at the ω-axis

At ω=0 the equivalence is failing since no current flow towards the ground is possible. In the case of synthesis of Gm-C filters based on LC prototypes this problem is usually mitigated by the fact that the inductor is either connected in series with a capacitor or there are two capacitors (to the left and to the right) which disconnect the inductor from DC signals. This will be demonstrated later on by the cell realizing a complex transmission zero without a transformer.

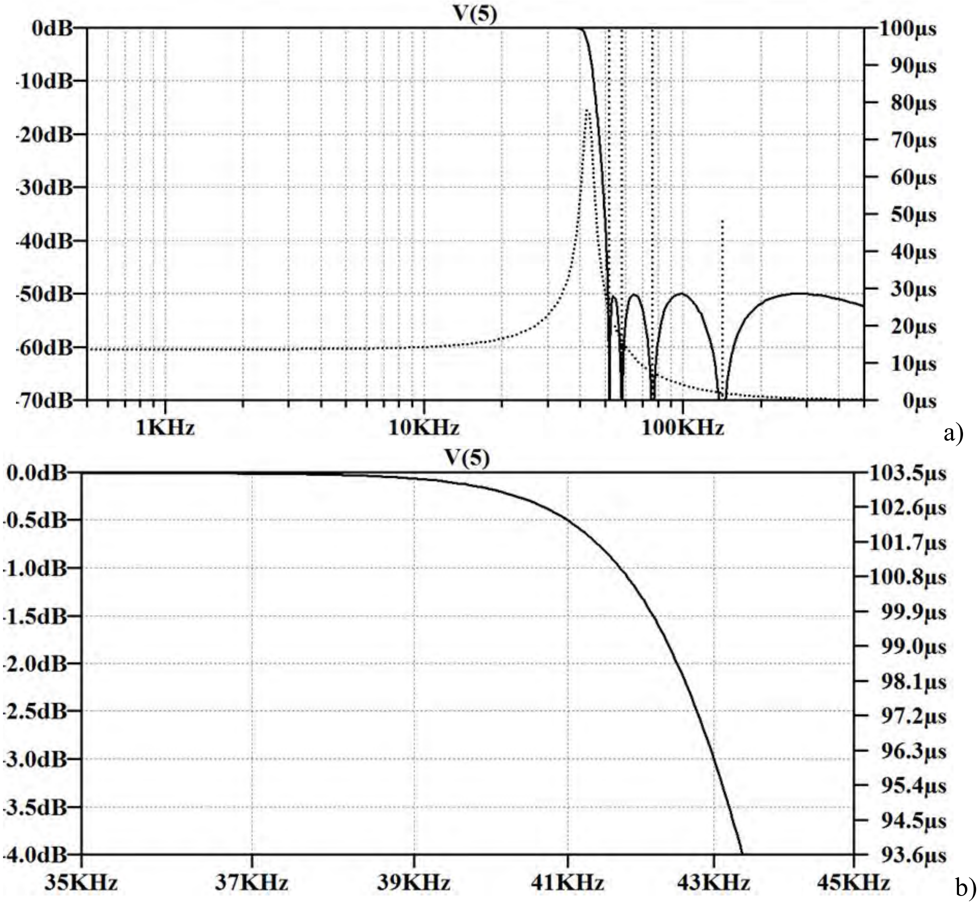


Fig. 8. SPICE simulation results for the example. a) The overall gain and group delay and b) the passband gain at the band edge

This equation may be used as a reference when choosing the value of the transconductance. Namely, a small transconductance e.g. g_m=10^{-6} S, would produce extremely small capacitances. If, for example L=100 µH, one would produce C=0.1 fF which is fairly small value and is in the range of the parasitic capacitances in any CMOS technology. In the opposite case, when large gm is chosen, the resulting capacitance may become very large. For example, if g_m=10^{-1} S, and L=100 µH, one gets C=10 nF. It seems that for this inductance a value of g_m=10^{-3} S, would be preferable. The question is, however, which is the output resistance of such an OTA. If satisfactory, the goal is reached. If not, one must go for a compromise.

Fig. 7 depicts the equivalent circuit to the one realizing a pair of transmission zeros at the ω-axis.

IV. DESIGN EXAMPLE

As an example a 9th order LSM [9] filter exhibiting $a_{\max}$=3 dB attenuation in the passband will be used. The stopband attenuation was set to $a_{\min}$=50 dB. The cut-off frequency was set to 43 kHz and, as can be seen, g_m= 10^{-6} S was used. The normalized poles an zeros of the transfer function of the LSM filter are given in Table 1. Fig. 8. depicts the SPICE simulation results.

TABLE I. NORMALIZED ZEROS AND POLES OF THE EXAMPLE LSM FILTER

Zeros	
Real part	Imaginary part
0.000000000e+000	±1.207678823e+000
0.000000000e+000	±1.347839032e+000
0.000000000e+000	±1.774177189e+000
0.000000000e+000	±3.290123935e+000

Poles	
Real part	Imaginary part
-7.170315101e-002	±9.991239141e-001
-2.565027588e-001	±9.818358904e-001
-1.451251999e+000	±0.000000000e+000
-5.858625045e-001	±9.061210530e-001
-1.112461676e+000	±6.430061028e-001

To show the schematic of the resulting GM-C filter excepts from the .html report file produced by the GM_LC program are given below.

RM Welcome to The Electronic Filter Design Software

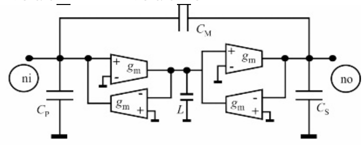
EXTRACTION OF THE CELLS

\---

k=1 (th) ZERO AT THE IMAGINARY AXIS=3.2901239e+000

\---

The capacitance (denoted L) within the simulated inductor is L*gm*gm=5.2452831e-013
Cp=2.1382177e-009 Cm=2.4127365e-010 Cs=-2.1680919e-010
Node_i=1 Node_o=2

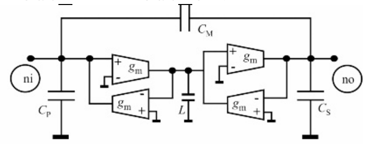


\---

k=2 (th) ZERO AT THE IMAGINARY AXIS=1.7741772e+000

\---

The capacitance (denoted L) within the simulated inductor is L*gm*gm=5.8743338e-013
Cp=6.2896543e-009 Cm=7.4088491e-010 Cs=-6.6280975e-010
Node_i=2 Node_o=3

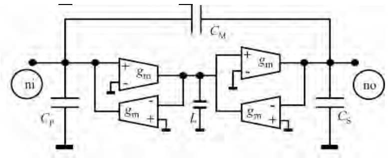


\---

k=3 (th) ZERO AT THE IMAGINARY AXIS=1.3478390e+000

\---

The capacitance (denoted L) within the simulated inductor is L*gm*gm=2.2310528e-013
Cp=6.8013276e-009 Cm=3.3800058e-009 Cs=-2.2579092e-009
Node_i=3 Node_o=4

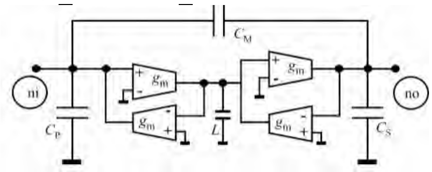


\---

k=4 (th) ZERO AT THE IMAGINARY AXIS=1.2076788e+000

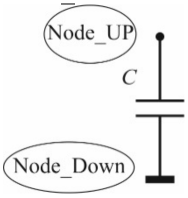
\---

The capacitance (denoted L) within the simulated inductor is L*gm*gm=2.6055821e-014
Cp=-4.4624834e-009 Cm=3.6049205e-008 Cs=5.0929306e-009
Node_i=4 Node_o=5



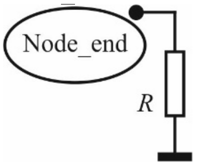
\--

k=5 (th) ZERO AT INFINITY: parallel capacitance C=9.3862427e-010
Node_UP=5 Node_Down=0



\-----------------Residual load impedance-------------------

Residual is resistor (Implemented in the SPICE description) and R=1.0000000e+003
Node_end=5



```
========================================
Here ends the synthesis process
```

V. CONCLUSION

A case study of synthesis of an low-pass GM-C filter was reported. It was shown that if an LC prototype is available (taken from a catalog e.g. [8] or produced by a dedicated software e.g. [9]) one may easily synthesize a proper GM-C equivalent. The example given here is related to a low-pass filter but, in general, no limitations exist as to where is the location of the passband of the filter. It is our experience [9] that transmission zeros of any kind may be realized by production of GM-C cells equivalent to the ones realized as LCM (M stands for use of transformers). It is our intention to further study the properties of the obtained solution in order to learn on the influence of the imperfections of the OTAs to the properties of the filters.

ACKNOWLEDGEMENT

This research was partly funded by the Serbian Ministry of Education, Science, and Technological Development within the project “Advanced technologies for measurement, control, and communication on the electric grid” under the cipher TR32004.

REFERENCES

[1] Haobijam G, Palathinkal RP (2014) Design and analysis of spiral inductors. Springer India, New Delhi

[2] Tsividis YP, Voorman JO (1993) Integrated continuous-time filters: Principles, Design, and Applications. IEEE, Piscataway, NJ

[3] Lawanwisut S, Siripruchyanun M (2012) An electronically controllable active-only current-mode floating inductance simulator. Proc. of the 35th Int. Conf. on Telecommunications and Signal Processing (TSP), Prague, Czech Republic: 386 -389

[4] Tellegen BDH (1948) The gyrator, a new electric network element. Philips Research Rep., 3:81-101

[5] Uzunov IS (2008) Theoretical model of ungrounded inductance realized with two gyrators. IEEE Trans. on Circuits and Systems-II: Express Briefs 55(10):981-985

[6] Sanchez-Sinencio E, Silva-Martinez J (2000) CMOS transconductance amplifiers, architectures and active filters: a tutorial. IEE Proc. Circuits Devices Syst. 147(1):3- 12

[7] Santos MM, Bertemes-Filho P, Vincence VC (2012) CMOS transconductance amplifier types for low power electrical impedance spectroscopy. XXIII Congresso Brasileiro em Engenharia Biomédica – XXIII CBEB:1382-1386.

[8] Zverev AI (2005) Handbook of Filter Synthesis. Wiley-Interscience, New York.

[9] Litovski, V. B., “Electronic Filters, Theory, Numerical receipts, and Design practice based on the RM software”, Springer, 2019.

[10]https://www.youtube.com/channel/UCF_Ipw_YD2gwrRpJDUJJULw/playlists?view_as=subscriber, last visited Oct. 2019.

[11]http://leda.elfak.ni.ac.rs/projects/RM%20software/RM %20software.htm, last visited Oct. 2019.

Parallel synthesis of active RC filters revisited

Dejan Mirković, Jelena Milojković and Vančo Litovski

Abstract - In early attempt to design parallel active RC filters which went almost forgotten within the filter design community was an inspiration to develop a complete procedure which seems to be absolutely simple and leads to a schematic feasible for practical implementation. After analysis of most alternative solutions (synthesis based on simulated inductances; state-variable filters; cascade solution) we are describing the complete algorithm for synthesis of parallel active RC filters. To verify we synthesized two higher order band-pass filters one as cascade and the other as a parallel realization. Study of the results revealed that, apart of the price, the parallel solution is by far in most electrically based aspects, preferable.

Keywords - Analog circuits, Active filters, Circuit synthesis, RLC circuits, Operational amplifiers.

I. INTRODUCTION

Active RC filters implemented in different technologies such as discrete PCBs, hybrid and monolithic integrated circuits are around almost for half a century now and the feeling [1], [2] is spread that this issue is settled and the main alternative to convert the transfer function into a schematic is cascade synthesis. Accordingly, a wide variety of cells realizing biquad functions were developed intended to be cascaded to produce a higher order function [3]. Unfortunately, this process of conversion of a transfer function into a cascade leads a long way encompassing several serious algorithms to be implemented and leaving ambiguities which are still not resolved satisfactorily. There are some less frequently used alternatives such as state-space realization [5] which is stemming from the non-factored representation of the transfer function, and a realization based on simulated inductances [6] using an intermediary step – passive LC synthesis. The last two methods are not frequently implemented for higher order transfer function which will be addressed later on.

While parallel implementation is frequent in design of digital IIR and switched capacitor filters, to our knowledge there was a single systematic attempt to implement parallel realization of active RC circuits [7]. In [7], an effort was made to produce a realization implemented in hybrid technology with limitations on the capacitance values more economical than the cascade on the expense of reduced sensitivity. To achieve that the author created a specific (multiple input and multiple output) cells with reduced number of capacitors which we find the main reason for this method not to be further investigated.

In these proceedings we will develop a straightforward procedure for parallel synthesis based on repetitive use of one type of widely known and frequently used second order cell [4]. To our knowledge this is the first complete description of the procedure. To come to a base for comparison with a cascade solution, high order band-pass filters will be synthesized by both techniques and their properties evaluated. At least basically, similar comparison was made for switched capacitor filters in [8] but on a much smaller scale. Also, in [8] the comparison is made for filters operating in the lower audio band so avoiding to consider the influence of the imperfections of the operational amplifiers.

In the sequel we will first give a glimpse to the synthesis procedures based on cascade, state-variable, and passive LC realizations. Then we will fully develop the method for synthesis of parallel active RC circuits. Finally, a case study will be presented enabling comparison of the cascade and parallel procedure.

II. CASCADE SYNTHESIS OF ACTIVE FILTERS

In this section the synthesis of active RC filters in a form of a cascade of second order cells (in the case of odd-order filters one additional first order cell is needed) will be discussed.

The transfer function is usually expressed in the following form,

$$H(s) = A_0 \frac{\prod_{i=1}^{m}\left(s - z_i\right)}{\prod_{i=1}^{n}\left(s - p_i\right)} \qquad (1)$$

where n is the order of the filter, m is the number of finite transmission zeros, s is the complex angular frequency, $\mathbf{z} = \{z_1, z_2, \ldots, z_m\}$ is the vector of finite transmission zeros, $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$ is the vector of poles, and A_0 is a constant defining the gain of the filter at the central frequency of its passband (ω_0).

The creation of a cascade realization corresponding to a given transfer function faces many challenges the main of which will be listed below.

The transfer function (1) may be created in a large

Dejan Mirković, is with the Dep. of Electronics, Uni. of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, Niš, Serbia, E-mail:dejan.mirkovic@elfak.ni.ac.rs.

Jelena Milojković, is with the ICNT, Bul. N. Tesle 40, Niš, Serbia E-mail: jelena.milojkovic@icnt.rs.

Vančo Litovski was with the Dep. of Electronics, Uni. of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, Niš, Serbia, E-mail: vanco.litovski@elfak.ni.ac.rs. Retired.

number of variants depending of the pairing poles and zeros into biquads and depending on the ordering of biquads so obtained. According to one study, for example, in the case of n=m=8, there are 18 possible combination to create biquads while for n=m=12, one may create 1350 combinations. Of course, in the similar way rises the number of filter structures due to the ordering of the biquads in the cascade. To choose among all combinations a procedure is to be implemented enabling pairing in order to get optimal biquads and ordering of the biquads in order to get optimum from linearity and noise point of view. That may improve the final solution from noise, linearity, and range of element values i.e. total silicon area point of view [9]. The following rules are advised: Pairing the transfer function poles having highest imaginary part with the attenuation poles having minimal frequency; High-Q sections should be in the middle; All-pass sections should be near the input; Last stages should be high-pass or band-pass to avoid output dc offset.

After that, depending on the properties of a cell such as: Type of the cell-function (low-; band-; high-; all-pass or notch); Order of the cell-function (first or second); Sign of the gain (inverting or non-inverting); Type of the transmission zeros (at the origin; at infinity; pair on the imaginary axis; complex pair in the right-half plane; real single in the right-half plane; mixed in several combinations); Value of the Q-factor of the pole (low or high); a choice is made as to which circuit (cell type) should be the most appropriate for realization [1]. Note we came to the number of 22 of different cell types which need a proper circuit realization. The very cell structure and the element value calculation are usually based on the literature e.g. [4].

Everything settled, for a prescribed overall gain G_0, an algorithm is to be implemented to correctly distribute the gain to the cascaded cells and recalculate the element values.

One should have in mind that the resolution between high and low Q of a pole is in essence arbitrary and, what is more problematic, abrupt. It looks as if the boundary between high and low Q is depending on the cell itself and not on a universal criterion. Here comes the absence of solutions for the "medium Q" which seems to be the most frequent one.

Having all this in mind one may conclude that synthesis of active RC circuits of higher order in a form of a cascade of second order cells is complex and challenging a task, and being pessimistic, one may say a never finished task.

III. SYNTHESIS BASED ON SIMULATED INDUCTANCES AND THE STATE VARIABLE APPROACH

Having available large amount of already synthesized passive LC filters of many kinds in the literature e.g. [2] an approach was developed to create active RC and Gm-C filters by different circuit transformations mostly based on circuits simulating the inductance [6], [10].

Catalogues, however, despite the fact being very useful, are not covering many important solutions and limit the designer to solutions which may not be the most appropriate for the problem to be handled. For example, second-order Chebyshev type II and Elliptic filters never have asymptotic slope larger than 12 dB/oct. In addition, equi-ripple approximation of constant group delay and phase correctors are nowhere to be found catalogued. Accordingly, we think that in many cases the catalogues are not sufficient. If so, one goes for synthesis of its own transfer function and creates LC circuit that realizes it. Such an example function will be used throughout this paper [11]. It is based on LSM filters [12] but extended with complex and zeros on the axis of real frequencies.

LC filters are usually synthesized as a cascade of cells realizing one transmission zero each [13]. The process is based on first creation of the input impedance of the filter which needs the solution of the Feldtkeller equation (of order $2n$) giving the reflection zeros needed. After extraction of a cell the residual input impedance is calculated. That involves subtraction and division of polynomials. In the case of higher order filters higher order polynomials produced by the Feldtkeller equation are to be solved. This creates errors at the very start of the synthesis. Later on, the error produced by division and subtraction related to a cell is accumulated making the element values of the last cells highly unreliable even if "long double" arithmetic is implemented. For that reason, we find the technique based on simulated inductances (not to mention the transformers frequently needed in LC synthesis) and the Gm-C technique not convenient for synthesis of higher order filters.

State variable filters are found to be very simple to synthesize provided the transfer function is given in a form of quotient of polynomials using summation notation. Analysis made in [14], however, claims that the gain-bandwidth product of the op-amps and the bandwidth of the OTAs must be much larger than the desired frequency of operation to ensure stability. For that reason, we find the state-variable solution not feasible for higher order filters.

IV. PARALLEL SYNTHESIS OF ACTIVE FILTERS

In this section a procedure will be proposed for synthesis of active RC circuits in a form of parallel network. After proper transformation of the transfer function into partial fractions, choice of cells realizing the cell-functions of the fractions will be offered and procedures of calculating element values will be given.

A. Decomposition of the Transfer Function

The continuous time transfer function may be represented in a form of sum of partial fractions as follows [2],

$$H_a(s) = A_0 \cdot H(s) = A_0 \cdot \frac{\prod_{i=1}^{m}(s-z_i)}{\prod_{i=1}^{n}(s-p_i)} =$$

$$A_0 \cdot \sum_{i=1}^{n} \frac{r_i}{(s-p_i)} = \begin{cases} \sum_{i=1}^{n/2} H_e(s),\ n-\text{even} \\ H_o(s) + \sum_{i=1}^{\lfloor n/2 \rfloor} H_e(s),\ n-\text{odd.} \end{cases} \tag{2a}$$

Note (2a) is valid for $n>m$ only. Index e is used to denote a second order fraction constructed by a complex pair of poles while o denotes a first order fraction constructed by a simple real pole. Complex residue in the pole p_i is denoted with r_i. We will denote $p_i=\sigma_i+j\cdot\omega_i$, $z_i=\alpha_i+j\cdot\beta_i$ and $r_i=\mu_i+j\cdot\xi_i$.

In the case $n=m$ polynomial long division must be carried out first (as explained in [7]) which leads to,

$$H_a(s) = A_0 \cdot [1 + H(s)]. \tag{2b}$$

The structure of the summing circuit is depicted in Fig. 1.

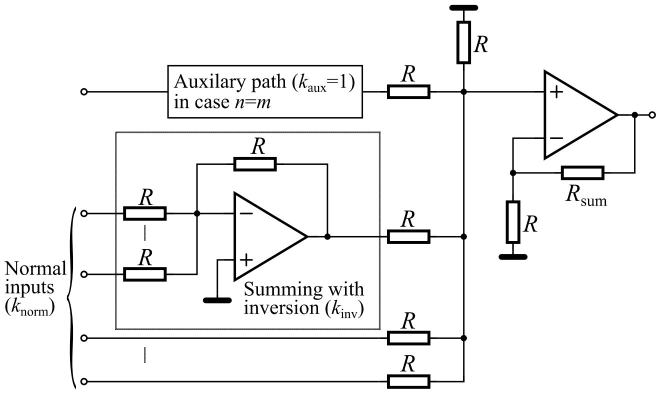


Fig. 1. Realization of the summing subsystem

As can be seen an auxiliary unity gain path is allowed for filters having n=m.

The summands in (2a) are given by

$$H_e(s) = G_i \frac{s + b_{0,i}}{s^2 + a_{1,i}s + a_{0,i}} \tag{3a}$$

with

$$G_i = 2 \cdot A_0 \cdot \alpha_i, \tag{3b}$$

$$b_{0,i} = \begin{cases} -\left(\alpha_i + \dfrac{\mu_i \cdot \alpha_i}{\mu_i}\right), \xi_i \cdot \omega_i > 0 \\ \left(\dfrac{\xi_i \cdot \omega_i}{\mu_i} - \alpha_i\right), \xi_i \cdot \omega_i < 0 \end{cases}, \tag{3c}$$

$$a_{1,i} = -2\,\mathrm{re}\{p_i\},\ a_{i0} = |p_i|^2,\ H_o(s) = G_o \frac{1}{s + a_o}, \tag{3d}$$

with $G_o = A_0 \cdot r_o$, and $a_o = -p_o$. In the above "re" stands for "real part" while "im" for "imaginary part".

The residues needed for the realization of the above computations are obtained (for the case of simple poles) as follows,

$$r_i = \lim_{s \to p_i} \left\{ \left[(s-p_i) \cdot H(s) \right] \right\} = \left[(s-p_i) \cdot H(s) \right]_{|s=p_i} \tag{4}$$

Accordingly (3a) and (3f) may be rewritten as.

$$H_e(s) = \frac{G_i \cdot s + G_i \cdot b_{0,i}}{s^2 + a_{1,i}s + a_{0,i}} \tag{5a}$$

with

$$G_i = 2 \cdot A_0 \cdot \mu_i, \tag{5b}$$

$$G_i \cdot b_{0,i} = \begin{cases} -2 \cdot A_0 \cdot (\mu_i\sigma_i + \xi_i \cdot \omega_i), \xi_i \cdot \omega_i > 0 \\ 2 \cdot A_0 \cdot (\xi_i \cdot \omega_i - \sigma_i \cdot \mu_i), \xi_i \cdot \omega_i < 0 \end{cases}, \tag{5c}$$

$$a_{1,i} = -2 \cdot \sigma_i,\ a_{i0} = |p_i|^2 = \sigma_i^2 + \omega_i^2 \tag{5d}$$

and

$$H_o(s) = G_o \frac{1}{s + a_o} \tag{5e}$$

with $G_o = A_0 \cdot r_o = A_0 \cdot \mu_0$ and $a_o = -p_o = -\sigma_0$.

The developments expressed so far are (apart of the notation) equal to the ones used in [7]. The difference and, accordingly, the novelty we are introducing, is in the use of standard and universally accepted circuits (cells) which are realizing (5a) and (5e) in the place of “multiple entry” cells used in [7]. Since two types of cell transfer functions are in view, only two types of circuit cells will be involved. Note that the second order cell has one zero at infinity and another on the real axis of the frequency plane being not restricted to any part of the real axis.

B. Second order cell design

According to the literature there are several concepts for creation of second order cells mainly intended to be used in cascade synthesis of active RC filters. The specifics of parallel synthesis may be seen from (5a) which represent a second order cell with a zero at the real axis of the complex frequency plane.

There were several cells already proposed at the time that are qualified for implementation in a case of transfer function containing a zero at the real axis. For the sake of brevity, we will discuss only the one known as Tow-Thomas (TT) (Fig.2) [10] biquad having low sensitivities to parasitics.

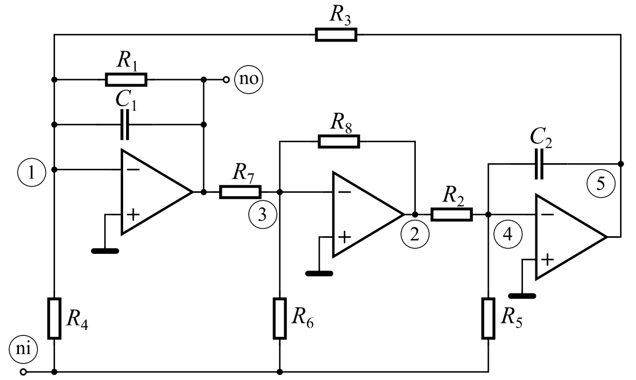


Fig. 2. Tow-Thomas second order filter cell.

To get the design equations we first simplify the notation. On may find easily by analogy that (5a), for a given cell, may be rewritten as

$$H_e(s) = g\frac{s+q}{s^2+as+b} \tag{6a}$$

Now, after circuit analysis one gets

$$\{a,b\} = \left\{1/(C_1R_1), R_8/(R_2R_3R_7C_1C_1)\right\}, \tag{6b}$$

$$\{q,g\} = \left\{\frac{R_4(R_5R_8-R_2R_6)}{R_2R_3R_5R_6C_2}, -1/(C_1R_4)\right\}. \tag{6c}$$

To get the element values we firstly adopt $C_1=C_2=C$. Then, we adopt $R_2=R_3=R_7=R$. With that set, one may calculate

$$\{R_1,R_4,R_8\} = \left\{1/(aC), 1/(|g|C), bC^2R^3\right\} \tag{7a}$$

and for

$$\{R_5,R_6\} = \left\{R, R_8/\left[R(qRC/R_4+1/R_5)\right]\right\}, \text{for } q>0, \tag{7b}$$

$$\{R_5,R_6\} = \left\{1/\left[R_8/(R\cdot R_6)-qRC/R_4\right], R\right\}, \text{for } q<0, \tag{7c}$$

$$\{R_5,R_6\} = \{R, R_5R_8/R_2\}, \text{for } q=0. \tag{7d}$$

Note, in the above case (g<0) both left and right half-plane zeros at the real axis are possible i.e. q>0 and q<0 is allowed. To keep this property, when the gain is positive we introduce inverters for every such cell.

C. First Order Cell

The first order cell is a simple inverter-integrator circuit as depicted in Fig. 3.

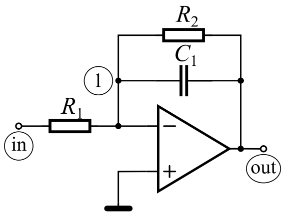


Fig. 3. Inverting-integrator first order filter cell.

To accommodate to the sign of the first order fraction we suggest using an additional inverter.

The transfer function may be expressed in the form

$$H_\mathrm{o} = g/(s+a) \tag{8a}$$

where by analogy we have

$$C_1 = C,\ R_1 = 1/(|g|C_1),\ R_2 = 1/(aC_1). \tag{8b}$$

Of course, the value of C_1, it may be chosen to be equal to the capacitances used within the TT cells.

V. A CASE STUDY

The example used for comparisons will be a wideband 16th order band-pass filter with central frequency at 10kHz and with relative bandwidth of 65%. We used an 8th order selective LSM low-pass filter having its group delay corrected by transmission zeros in the right-half plane [11] as a prototype. The final version was obtained after low-pass to band-pass transformation. The poles and zeros of the filter to be synthesized are given in Table I.

Two filters were created one as parallel and another as cascade using the $\mathcal{RM}$ software for filter design [16]. The parallel version used 8 TT cells while the cascade used 6 TT cells, one high-pass high Q, and one low-pass high-Q cell [3]. Table II depicts a comparison in number of circuit elements used. As expected the parallel solution needs more resistors and operational amplifiers while less capacitors.

As can be seen from Table III the parallel solution will be realized with a span of the resistors more than three times smaller than in the cascade case, which may be decisive when integrated resistors are used. In addition, in the parallel solution all capacitances may be equal. That makes the parallel solution extremely convenient for programmable filters [17].

TABLE I
ZEROS AND POLES OF THE EXAMPLE TRANSFER FUNCTION

Zeros	Poles
$\pm j1.5219870$	$-0.036474871732 \pm j1.360312888830$
$\pm j0.6570359$	$-0.019697178268 \pm j0.734596838828$
$\pm j1.7125020$	$-0.119413574934 \pm j1.302119716210$
$\pm j0.5839408$	$-0.069841675066 \pm j0.761573566207$
$0.1659470 \pm j0.8802626$	$-0.171449892048 \pm j1.164512716530$
$0.2068132 \pm j1.0970360$	$-0.123747207952 \pm j0.840509116528$
	$-0.173697990673 \pm j1.053741406880$
	$-0.152294259327 \pm j0.923895356878$

TABLE II
ZEROS AND POLES OF THE EXAMPLE TRANSFER FUNCTION

Type	No. Resistors	No. Capacitors	No. Opamps
Cascade	60	19	20
Parallel	64+14	16	26

TABLE III
ZEROS AND POLES OF THE EXAMPLE TRANSFER FUNCTION

Type	$\frac{R_{max}}{R_{min}}$	$\frac{C_{max}}{C_{min}}$	THD^{a} $[\%]$	v_n^{b} $[V/\sqrt{Hz}]$	V_{offset}^{a} $[V]$
Cascade	3132	1.39	0.17	0.43	0.3
Parallel	1042	1	0.23	1.17	0.036

[a] $V_{out}=1V$; [b] $f=10kHz$

The rest of the comparison will be made based on simulations. The results depicted in Table III are related to the use of the THS4211 opamp, having GBW of 140 MHz implemented with ±10V supply. Its gain-bandwidth product is advertised to be 140MHz. Its model was downloaded from [18].

Against expectations the parallel solution exhibits slightly larger distortions even though the "measurement' took place at relatively small amplitude of the output signal. Opposite stands for the offset at the output, the one produced by the cascade solution being significant. This happened despite the fact that the last cell in the cascade is a high-pass filters as advised. The noise performance is according expectations.

Finally, the amplitude characteristics (as depicted in Fig. 4) will be discussed. Three traces are shown: one for the case when infinite gain operational amplifiers were used (named ideal); one for the parallel; and one for the cascade solution. As for the passband, as can be seen from Fig. 4a, the cascade solution exhibits serious deterioration of the passband characteristic. We believe that this is a consequence of inadequacy of one or more cells from the point of view of the value of the Q of the critical pole.

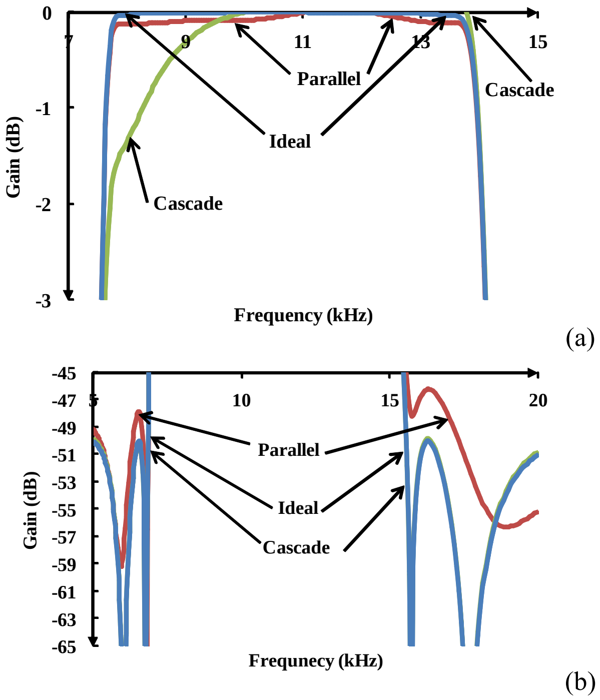


Fig. 4. The amplitude characteristics of the solutions using infinite gain amplifiers (ideal) and parallel and cascade solutions using the THS4211 operational amplifier.

Either the threshold (between high and low Q) was wrongly set or one needs a cell which may be stated as one with medium Q.

On the other side, Fig. 6b claims that the stopband response of the parallel solution is highly distorted. We think that this is a consequence of the smaller value of the gain of the op-amps than needed which makes the small numbers manipulated in the summing (and subtracting) amplifier slightly erroneous which in turn gives rise to the arithmetic error Better amplifier and passive components with small tolerances would be needed to overcome this disadvantage.

To verify the claims related to the value of the gain of the opamps we repeated both designs with a new opamp. We used LTC6268-10 having gain-bandwidth 4GHz. Its Spice model was taken from [19]. The new simulation results are depicted in Fig. 5. The responses of the cascade and the parallel solutions almost overlap. That means that better operational amplifiers are needed for both schematics.

While in the case of the parallel solution we think that the “arithmetic” was improved, for the case of the cascade solution we think that the larger value of the gain improves the pole Q sensitivity so making the choice of high- or low-Q cell less important.

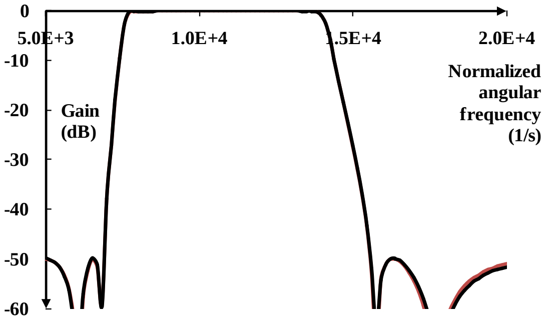


Fig. 5. The amplitude characteristics of the solutions using parallel and cascade solutions using the LTC6268-10 operational amplifier.

VI. CONCLUSION

In these proceedings we tried to draw the attention of the filter design community to an alternative to the most frequently used cascade synthesis. Special attention was paid to higher order filters which are encountered when selective band-pass and band-stop solutions are sought. We showed that the parallel implementation is, apart from the price, equally feasible (if not preferable) as the cascade. Its main advantage is related to the design algorithm which is straightforward and easy to program. In addition, it is always convenient for design of programmable filters since all capacitors may have the same value. Finally, significantly lower spread of the resistor values was observed in the case of parallel solution which may lead to significant savings in the occupied chip area.

ACKNOWLEDGEMENT

This work was partially funded by Serbian Ministry of Education, Science and Technological Development under contract No. TR32004.

REFERENCES

[1] P. A. Mohan, VLSI Analog Filters: Active RC, OTA-C, and SC. Springer Science & Business Media, 2013.

[2] A. B. Williams and F. J. Taylor, Electronic filter design handbook, 2006, no. Sirsi) i9780071471718.

[3] R. P. Sallen and E. L. Key, “A practical method of designing rc active filters,” IRE Transactions on Circuit Theory, vol. 2, no. 1, pp. 74–85, 1955.

[4] P. Fleischer and J. Tow, “Design formulas for biquad active filters using three operational amplifiers,” proceedings of the IEEE, vol. 61, no. 5, pp. 662–663, 1973.

[5] S. Pandit, S. Kar, C. Mandal, and A. Patra, “High level synthesis of higher order continuous time state variable filters with minimum sensitivity and hardware count,” in Proceedings of the conference on Design, automation and test in Europe: Proceedings. European Design and Automation Association, 2006, pp. 1–2.

[6] M. Tan and R. Schaumann, “Simulating general-parameter lc-ladder filters for monolithic realizations with only transconductance elements and grounded capacitors,” IEEE Transactions on circuits and systems, vol. 36, no. 2, pp. 299–307, 1989.

[7] P. Moran, “A low-cost parallel implementation for active filters,” IEEE Journal on Electronic Circuits and Systems, vol. 2, no. 1, pp. 21–25, 1978.

[8] I. Sahawaneh, A. Sakla, and E. El-Masry, “Comparison between cascade and parallel switched capacitor structures,” in Circuits and Systems, 1989., Proceedings of the 32nd Midwest Symposium on. IEEE, 1989, pp. 973–976.

[9] C.-F. Chiou and R. Schaumann, “Refined procedure for optimizing signal-to-noise ratio in cascade active-rc filters,” in IEE Proceedings G-Electronic Circuits and Systems, vol. 128, no. 4. IET, 1981, pp. 189–191

[10] T.-Y. Lo and C.-C. F. Hung, 1V CMOS Gm-C filters: design and applications. Springer Science & Business Media, 2009.

[11] V. Litovski, “Synthesis of monotonic passband sharp cutoff filters with constant group delay response,” IEEE Transactions on Circuits and Systems, vol. 26, no. 8, pp. 597–602, 1979.

[12] B. Rakovich and V. Litovski, “Least-squares monotonic lowpass filters with sharp cutoff,” Electronics Letters, vol. 9, no. 4, pp. 75–76, 1973.

[13] J. Scanlan and J. Rhodes, “Unified theory of cascade synthesis,” in Proceedings of the Institution of Electrical Engineers, vol. 117, no. 4. IET, 1970, pp. 665–670.

[14] T. Bakken and J. Choma, “Stability of a continuous-time state variable filter with op-amp and ota-c integrators,” in VLSI, 1998. Proceedings of the 8th Great Lakes Symposium on. IEEE, 1998, pp. 77–82.

[15] J. Friend, C. Harris, and D. Hilberman, “Star: An active biquadratic filter section,” IEEE Transactions on Circuits and Systems, vol. 22, no. 2, pp. 115–121, 1975.

[16] V. Litovski. (2018) The rm software for filter design (a set of ten presentations). YouTube video. Accessed Sep. 03, 2018. [Online]. Available: https://www.youtube.com/watch?v=QLxfq_kqR_E

[17] M. Banu and Y. Tsividis, “An elliptic continuous-time cmos filter with on-chip automatic tuning,” IEEE Journal of Solid-State Circuits, vol. 20, no. 6, pp. 1114–1121, Dec 1985.

[18] Texas instruments (ti). [Online]. Available: http://www.ti.com/product/THS4211/toolssoftware

[19] Analog devices (ad). [Online]. Available: http://www.analog.com/en/design-center

Performance Evaluation of a Time-of-Arrival Based Indoor Localization System

Vladica Sark, Nebojsa Maletic, Jesús Gutiérrez and Eckhard Grass

Abstract – The appearance of global navigation satellite systems (GNSS) has boosted the development of many location based services and applications. Additionally, abundancy of WiFi access points has brought these services and applications indoors, where GNSSs are not available. In indoor environments, WiFi is mainly used for coarse localization, since no other, widely available, solution exists. In order to increase the localization precision in indoor environments, many time-of-flight (ToF) solutions are emerging. These solutions are mainly working in the 2.4/5 GHz or 60 GHz industrial-scientific-medical (ISM) band or they use impulse-radio (IR) ultra-wideband (UWB) technology. In this paper, we evaluate the performance of a time-of-arrival (ToA) based localization, which is using ToF. The performance of the ToA localization is evaluated for systems working in the 2.4/5 GHz ISM band. The channel bandwidths, which are evaluated, are typical for these bands. The performed simulations show that localization precision better than 1 meter can be achieved. Only line-of-sight (LOS) scenarios were considered.

Keywords – time-of-flight, time-of-arrival, localization, ISM, line-of-sight, location based services

I. INTRODUCTION

The appearance of global navigation satellite systems (GNSS) has opened huge possibilities for many location based applications and services. Initially, the GNSSs, or more specifically global positioning system (GPS), was reserved mainly for the military applications as well as marine navigation of civil vessels, mainly due to the absence of other high precision systems and the high price of the GPS equipment. With the newest advances in technology, GPS (also other GNSSs) receivers have become more affordable, smaller and power efficient. This has opened new possibilities for integrating them in car navigations units and, nowadays, in every smartphone. The GPS (and also other GNSSs) has an impressive advantage due to its global availability, previously not offered by other localization systems. Nevertheless, this global availability has a few disadvantages, which hardly would be resolved in the future. These disadvantages originate from the limited transmit power of the GPS transmitters as well as the huge distance between the transmitter, being a low Earth orbit satellite, and the receiver. The received power in an outdoor environment is about -130 dBm [1], which introduces a huge challenge for the GPS receivers. Additionally, this power goes down by additional 10-30 dB indoors [2] making it impossible for decoding even by the most sensitive GPS receivers. The latest advancements in GPS receiver design have shown that sensitivities down to -160 dBm can be achieved [2], making them usable in some indoor environments. Unfortunately, this is very limited and not possible in multi-story buildings, made of concrete and steel, where the additional attenuation is even higher than 30 dB. Therefore, many islands with no GPS coverage exist and severely limit many location-based services and applications. Additionally, in urban canyons (streets surrounded by tall buildings) only a limited visibility of GPS satellites is feasible. This strongly limits the positioning precision of the GPS system reducing its usability.

In order to solve the main coverage problem of the GPS (and GNSS) systems, many indoor localization systems start to appear [3]. Many of them develop proprietary hardware, while most of the solutions try to use the already available infrastructure, i.e. available WiFi access points, in order to benefit from cheap and easy deployment.

One of the first commercially available indoor and outdoor localization system is LOCATA [4]. This system is mainly used for outdoor applications, where the GNSS performance is poor or not available at all. It can be also used indoors for indoor navigation in factories etc.

One of the mostly used approach nowadays is localization using WiFi. Many different WiFi localization solutions are already available. Some of the approaches use a database with positions of the available WiFi access points. These methods can perform coarse localization, with precision not better than a few meters, but the main advantage is that they are straightforward to implement and available in almost all urban environments where WiFi is deployed. There are already a few databases with WiFi access points and their corresponding locations available [5].

In order to improve the precision of the WiFi localization, an approach called WiFi fingerprinting is extensively being investigated [6]. With this approach, a map of received signal strength (RSS) from different WiFi access points is constructed. This mapping is a painstaking process because many measurements must be performed. The maps are stored in a database. When a user demands to obtain its position with, for example, his smartphone, the RSS value for each visible WiFi access point is measured.

Vladica Sark, Nebojsa Maletic and Jesús Gutiérrez are with the IHP – Leibniz-Institut für innovative Mikroelektronik, Frankfurt (Oder), Germany
Eckhard Grass is with the IHP – Leibniz-Institut für innovative Mikroelektronik, Frankfurt (Oder), Germany and Humboldt-University of Berlin, Berlin, Germany
E-mail: {sark, maletic, teran, grass}@ihp-microelectronics.com

The obtained RSS values are queried in the database and the position with the highest likelihood is found. This process can be quite complex and computationally intensive. Nevertheless, with this approach, already available hardware is used and no infrastructure deployment is necessary. Localization precision of less than 1 meter can be achieved [3]. Main disadvantages include creating the fingerprint database, which can be a long process, as well as the complexity of the position estimation algorithm, which in some cases must be performed on a high performance server computer, i.e. in the cloud.

Another group of localization systems uses proprietary hardware in order to perform localization. The simplest approach is the RSS based ranging which is used to estimate the distance between two wireless nodes. Knowing the distance between the anchor nodes, i.e. nodes with known positions, and the mobile node, i.e. with unknown position, one can perform trilateration and find its position. This is a very simple approach, since almost every radio receiver can estimate the RSS value. The main disadvantage is that due to the multipath propagation, the free space path loss formula is not valid anymore, since the path loss exponent is lower or larger than 2 [7]. It is hard to estimate this coefficient correctly since it depends on the environment. Despite the fact that this approach leads to extremely imprecise location estimates, it is still being attractive due to its simplicity. It can be used in very simple systems, like for example wireless sensor nodes, which usually have very limited hardware and radio resources available.

In order to achieve better precision and accuracy, usually methods based on ToF of the radio signal are used. These methods are used to estimate the propagation time of a RF signal between two wireless nodes. Having the propagation time and multiplying it with the speed of light, the distance between the nodes can be estimated. A few different approaches for measuring the ToF are possible. One of the simplest is to perform a so-called two way ranging (TWR) [8]. With this approach, the ToF, i.e. distance, between two wireless nodes is estimated and later used to estimate node's position, using trilateration. The main advantage is that no precise synchronization between the nodes is required. The main issue is the large number of wireless transmissions needed for location estimation in a wireless network where large number of nodes demand position estimation. In order to avoid large number of transmission, a so-called time of arrival (ToA) based approach is usually used. In this approach, multiple anchor nodes are synchronized and transmit timestamped frames. Nodes that require estimation of their position receive these frames and estimate its own position. The receiving nodes must not transmit any frames back to the anchor nodes in order to estimate their own position. The main complexity of this approach is in the synchronization of the anchor nodes. Nevertheless, it allows unlimited number of users to estimate their positions. However, in all of the ToF based localization methods, it is important for the RF signals to travel the line-of-sight (LOS) path. Otherwise, a positive bias would be added to the range, or position, estimate if the non-line-of-sight (NLOS) paths are used for distance estimation, since these are the reflected paths and are longer than the LOS paths.

In this paper, we evaluate the position estimation error of a ToA based indoor positioning system. We use a simulation approach, since finding a closed form expression for the error is significantly complex.

In addition, worth mentioning is that large number of non-RF based localization approaches are being investigated. Some of them used cameras, other use inertial navigation, some use ultrasound etc. They would not be addressed in this paper.

The rest of the paper is organized as following: in Section II we describe the method for position estimation in indoor environment, in Section III the simulation scenario is presented, Section IV gives the obtained results and Section V is concluding this paper.

II. ToA LOCALIZATION

A. Position estimation using ideal range measurements

In a ToA based localization, multiple synchronized anchor nodes are transmitting timestamped frames in a previously defined sequence. A scenario with four anchor nodes and a single mobile node is shown in Fig 1.

The anchor nodes are synchronized and share a common time base. The synchronization can be performed using wireless or wired IEEE 1588 [9] protocol or other synchronization method. It can be assumed that the nodes transmit sequentially starting at t_{TX1} for the anchor node 1, t_{TX2} for anchor node 2 and t_{TXi} for anchor node i. The transmitted frames are received at the wireless node with the unknown position (ST) at times t_{RX1}, t_{RX2} and t_{RXi} correspondingly. The distances from the anchor nodes to the mobile node can be calculated as

$$r_i = c \cdot (t_{RXi} - t_{TXi} + t_{OS}) \tag{1}$$

where r_i is the distance between the anchor node and the mobile node, t_{TXi} and t_{RXi} are the transmission and reception times and t_{OS} is the offset of the real time clocks between the mobile node and the anchor nodes. It is worth mentioning that the anchor nodes and the mobile node have different time base, since they are not synchronized, and the time difference between them is exactly t_{OS}.

In this paper, without loss of generality, we are describing a 2-dimensional case, but the process is the same for a 3-dimensional case.

The distances between the anchor nodes and the mobile node can be calculated as

$$\begin{aligned}\sqrt{(x-x_i)^2+(y-y_i)^2} = r_i = \\ = c \cdot (t_{RXi} - t_{TXi} + +t_{OS})\end{aligned} \tag{2}$$

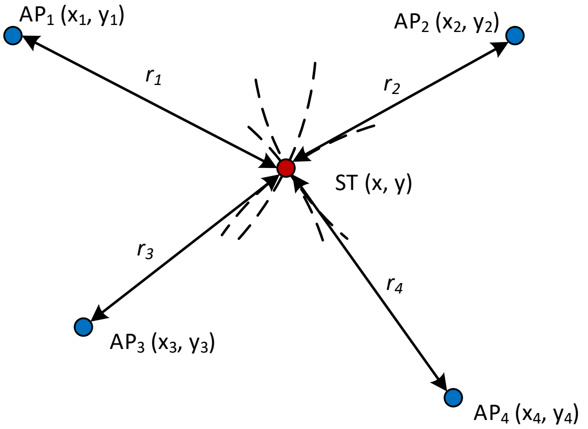


Fig. 1. Position estimation of a mobile node and four anchor nodes using trilateration

Eq. (2) is a quadratic equation valid for each anchor node. There are three unknowns in a 2-dimensional space. The unknowns are the position coordinates x and y and the time offset t_{OS}. In order to find the position coordinates and eventually the time offset, the system of equations consisted of Eq. (2) for each anchor node should be solved. Therefore, for a single solution of the system of equations, four equations are needed. Nevertheless, even when a fully determined system of equations is available, its solution can be challenging, if it is not linear, as in this case. In order to solve this system, the nonlinear equations can be linearized and the system can be solved [1].

B. Position estimation using noisy range measurements

The approach discussed is the ideal case. In real systems, additional noise would be present and the circles described by the radii r_1, r_2, r_3, and r_4 in Fig. 1, would not intersect in a single point. Therefore, we need to find a position that minimizes the sum of squared differences i.e.

$$(\hat{x},\hat{y}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{i=1}^{n}[r_i^2 - (x-x_i)^2 - (y-y_i)^2]^2 \quad (3)$$

where the $(\hat{x},\hat{y})$ is the estimated position. This is a minimization problem, which can be solved using different approaches as Gauss-Newton or Levenberg–Marquardt algorithms [10]. With this approach, even more than the minimum necessary anchor points can be used in order to improve the precision of the noisy position estimate. Anyway, if less than four access points are available for two-dimensional case, additional constraints for the possible positions can be introduced in order to have a fully determined system. Otherwise, the algorithms would not be able to converge.

C. ToA estimation

As shown in Eq. (2), time of arrival of the transmitted frames, t_{RXi}, should be estimated for each frame arriving from the access points. The precise estimate of the ToA would lead to precise position estimate. Just as an example, if the receiver has a sample rate of 200 Msps, each 2 successive samples would be spaced 5 ns apart. If the ToA is estimated with a resolution of 5 ns, this would lead to a distance estimation resolution of 1.5 m, because the speed of light is c = 300 000 *km/s*. Therefore, additional measures must be taken in order to improve the ToA estimation resolution.

The ToA at the receiver can be estimated using a few different approaches. The simplest one is to detect the increase in the received power over a given threshold and to use this time as a ToA. This method is rather simple, but sensitive to noise in low SNR scenarios.

The second approach is to transmit a pseudo-noise (PN) sequence with a strong autocorrelation peak. When this sequence is received at the receiver, it is correlated with the same sequence, locally generated. This would generate a peak, which position corresponds to the ToA. Nevertheless, this would only solve the issue with the noise in low SNR scenarios, but would not improve the ToA estimation resolution. Therefore, additional interpolation over the cross-correlation peak is performed. A quadratic interpolation is usually sufficient.

In Fig. 2 a cross-correlation peak of a received PN sequence is shown. The time at the peak maximum represents the ToA. The red points are the samples of the peak. Additional interpolation can lead to a subsample resolution of the ToA estimation.

III. SIMULATION SCENARIO AND LOCALIZATION PRECISION EVALUATION

In order to evaluate the localization precision, a two dimensional indoor scenario was assumed. The simulation scenario uses a square area of 10 × 10 meters, with four anchor nodes positioned in the corners of the area at positions (0, 0), (0, 10), (10, 0) and (10, 10) meters.

The mobile node was positioned at positions (x_i, y_i) where $x_i = 1, 2, .. 9$ and $y_i = 1, 2, .. 9$ meters. Dipole antennas with an antenna gain of 1.76 dB were assumed. It was assumed that the propagation loss is equal to the free space path loss (FSPL), since for some indoor environments the propagation loss can be lower or higher compared to the FSPL, due to constructive and/or destructive combining of the multipath propagation components. The noise figure (NF) of the low noise amplifier for the 2.4/5 GHz can go down to 1.2 dB, which is used in this simulation. A channel bandwidth of 20 MHz was used and a transmit power of 10 dBm. The maximal transmit power in these ISM bands is limited to 20 dBm.

For each position of the mobile node, a location estimation was performed multiple times and the location error was calculated.

As can be seen in Fig. 3, the red dots represent the true positions of the nodes and the blue dots represent the estimated positions. The black dots are the anchor nodes.

The estimated positions are concentrated around the true positions.

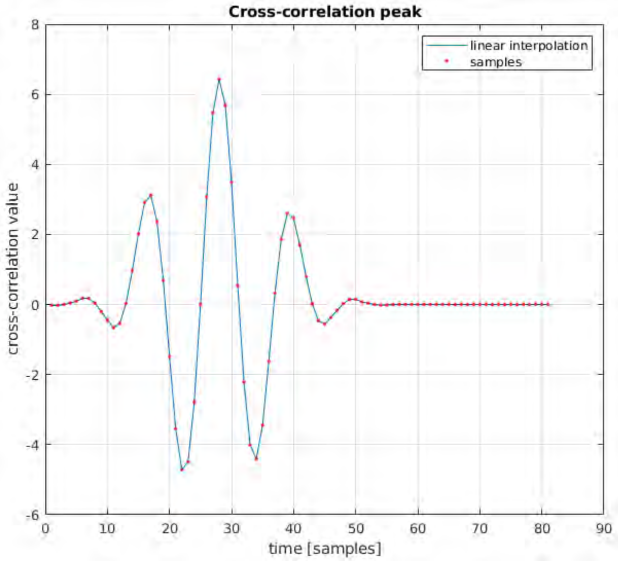


Fig. 2 – Cross-correlation peak of a received PN sequence

The positioning error empirical cumulative distribution function (CDF) was also calculated using the simulation data. It is shown in Fig. 4. As can be noticed, 90 percent of the positioning errors are under 1 meter. The root-mean-square (RMS) error for this use case was calculated to be 0.8 meters.

This error is comparable or even better to the systems reported in [3]. Nevertheless, these systems were tested in realistic scenarios where LOS and NLOS conditions are expected as well as strong multipath propagation. Having a narrow bandwidth, would cause significant issues when multipath propagation components should be resolved.

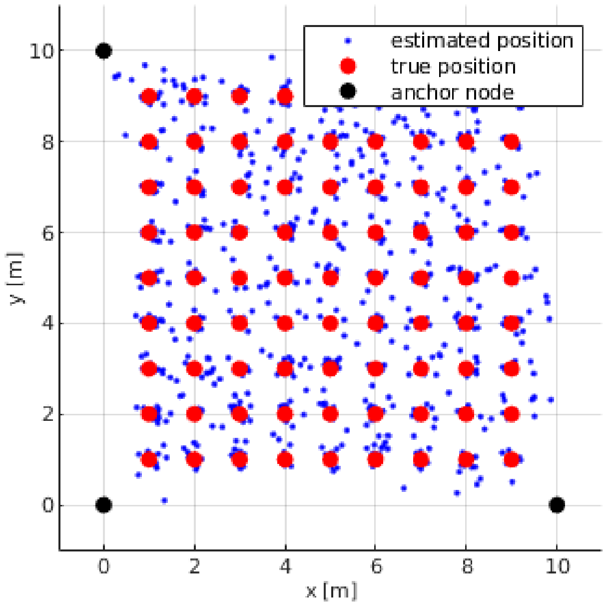


Fig. 3 – Estimated and true position of the mobile node

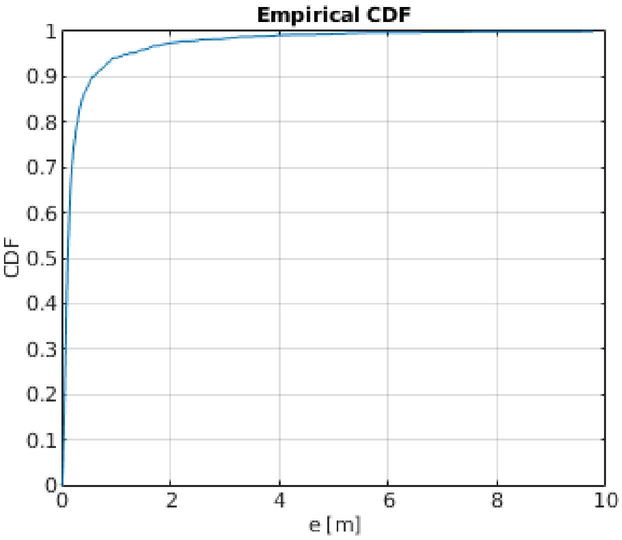


Fig. 4 – Empirical CDF of the positioning error

Therefore, the obtained results are expected to be slightly worse in realistic scenarios. In order to improve the precision longer PN sequences can be transmitted in order to obtain better coding gain and larger bandwidths can be used to improve the multipath resolution.

IV. CONCLUSION

In this paper, we evaluated the precision of a localization system which uses ToA approach. We use additive white Gaussian noise (AWGN) channel to evaluate the noise performance of this approach. The system was evaluated using simulation since obtaining a closed form solution is not straightforward due to the nonlinearity of the system of equations used. It was shown that with parameters typical for a WiFi, working in 2.4/5 GHz ISM band, positioning precision of better than 1 meter is possible.

ACKNOWLEDGEMENT

This work has received funding from the European Union's Horizon 2020 research and innovation programme, WORTECS project, under grant agreement No 761329.

REFERENCES

[1] Tsui, J. B.-Y., "Fundamentals of global positioning system receivers: a software approach", Vol. 173, *John Wiley & Sons*, 2005.

[2] Ozsoy, K., Bozkurt, A., Tekin, I., "Indoor positioning based on global positioning system signals", *Microwave and Optical Technology Letters*, May 2013.

[3] Lymberopoulos, D., Liu, J., et. al. "The microsoft indoor localization competition: Experiences and lessons learned", *IEEE Signal Processing Magazine,* Sep. 2017.

[4] Barnes, J., Rizos, C., Wang, J., Small, D., Voigt, G., Gambale, N., "Locata: the positioning technology of the

future", *InProceedings of the 6thInternational Symposium on Satellite Navigation Technology Including Mobile Positioning & Location Services*, Melbourne, Australia, July 2003.

[5] https://wigle.net/

[6] Basri, C., El Khadimi, A., "Survey on indoor localization system and recent advances of WIFI fingerprinting technique", *5th International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 253-259, Sep., 2016

[7] Kaddouri, S., El Hajj, M., Zaharia, G., El Zein, G., "Indoor Path Loss Measurements and Modeling in an Open-Space Office at 2.4 GHz and 5.8 GHz in the Presence of People", *29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1-7, Sep, 2018.

[8] Sark, V., Maletic, N., Ehrig, M., Gutiérrez, J., Grass, E., "Achieving Millimeter Precision Distance Estimation using Two-Way Ranging in the 60 GHz Band", *European Conference on Networks and Communications (EuCNC)*, pp. 310-314, Jun., 2019.

[9] Cooklev, T., Eidson, J.C., Pakdaman, A., "An implementation of IEEE 1588 over IEEE 802.11 b for synchronization of wireless local area network nodes", *IEEE Transactions on Instrumentation and Measurement*, Sep., 2007.

[10] Croeze, A., Pittman, L., Reynolds, W., "Nonlinear least-squares problems with the Gauss-Newton and Levenberg-Marquardt methods", *University of Mississippi*, Oxford, UK, Jun., 2012.

A Graph Perspective on Logic Circuits: Faults, Robustness & Eigenspectrum

Miljan Petrović

Abstract – Graphs can be used to model logic circuits with vertices denoting nodes of the electrical network and edges representing connections between nodes through a logic gate. Consequently, linear algebra, perturbation and spectral matrix theory provide a framework for exploration of concepts such as circuit robustness, and fault detection and identification. This paper, through a set of benchmark circuits, illustrates how the aforementioned mathematical and engineering concepts are intertwined. In particular, through means of simulation, it was revealed that some of ISCAS'85 circuits exhibit higher robustness to open-circuit faults, but higher ability to locate (fatal) short-circuit faults.

Keywords – Faults, Graph modelling, Circuit Robustness, Spectral Analysis.

I. INTRODUCTION

Design of logic circuits (and electrical networks in general) represents a great challenge from the perspective of error-tolerance (robustness), fault detection and testability. Employed design techniques are often highly dependable on the nature of the desired circuit function [1] and require to consider different types of faults simultaneously. This hinders comparison between circuits and measurement/approximation of their robustness. Particularly, in testing of combinational circuits, obstacles emerge such as delay and timing skew effects [2] rendering the definition of (near)-minimal test set a hard problem. Some sophisticated techniques rely on probabilistic approaches to error estimation [3].

Furthermore, use of graph algebraic theory in the analysis and design of electric circuits is a rich discipline that has provided concepts such as the connection between Laplacian matrix eigenvalues and effective resistance [4], and Kron reduction of graphs as a way to analyse an electrical network through a generalization of $\mathrm{Y}-\Delta$ transformation [5]. However, perception of circuits' robustness and testing has not been often viewed upon from a graph perspective. Hence, this paper provides an illustration of how some of the related features of combinational logic circuits reflect in (spectral) graph theory.

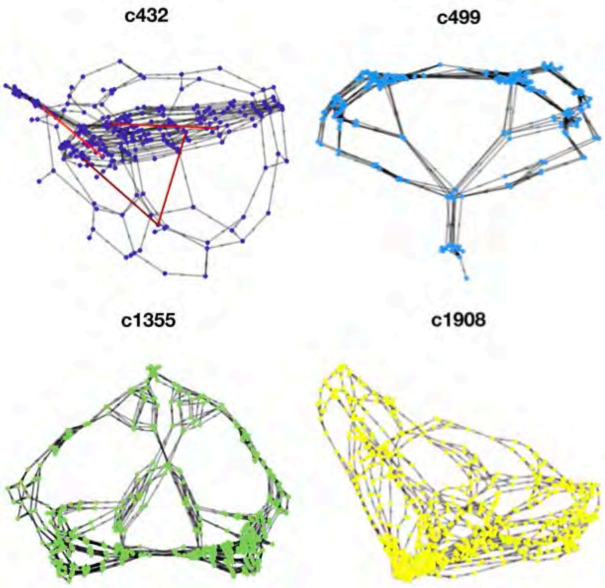


Fig. 1. Graph representation of the benchmark circuits

On several ISCAS'85 benchmark circuits it is shown that the density of graph eigenvalues expresses a circuit's sensitivity to (open/short circuit) faults, *i.e.* that higher distance between eigenvalues makes the circuit more robust. Moreover, the emergence of a dominant eigenvector in a faulty circuit was recognized as potential marker for detection of a fault and its localization.

The main section of the paper consists of three subsections: in the first, the description of benchmark circuits is provided; further, spectral and statistical approach for evaluation of circuits' robustness is presented with the main results reported; and then, a discussion is given on the relationship between the spectral results, and circuits' robustness and identifiability of faults. The conclusion contains a summary of advantages of the graph perspective on circuits' robustness, and the main limitations of the approach in the current form.

II. GRAPH SPECTRAL ANALYSIS OF CIRCUITS

A. Data

The analysis reported in this paper was performed on combinational ISCAS'85 benchmark circuits of different sizes, specifically, c432 (27-channel interrupt controller), c499, c1355 (32-bit single-error-correcting circuits), and c1908 (16-bit single-error-correcting/double-error detecting circuit) [6]. These circuits were collected in the form of

Miljan Petrović is with the École Polytechnique Fédérale de Lausanne, Institute of Bioengineering, and University of Geneva, Department of Radiology and Medical Informatics, Campus Biotech, Chemin des Mines 9, Geneva, Switzerland, E-mail: miljan.petrovic@epfl.ch.

Verilog descriptions. In order to perform graph-based analysis on the data, Verilog netlists needed to be transformed into appropriate graph representations. This was done with java-based toolbox *Verilog2GEXF*, which was developed as a free and universal visualization tool [7]. For each netlist, the tool exports a graph with n nodes and m edges that can be described as a square adjacency matrix $\mathbf{A}$ of size n with exactly m nonzero entries equal to 1 called edge weights. The graph vertices denote nodes of a logic circuit, whereas edges represent connections between nodes through a logic gate (Fig. 1). Resulting graph gives a simplified model of the electrical network from the perspective of information flow. This is explained in more detail for the toy example circuit in the following subsection.

B. Methods & Results

Graph representation of a logic circuit is given in the form of a binary adjacency matrix, having an entry equal to 1 at position (i,j) if there is a connection between j^{th} and i^{th} node in that direction as seen by the information flow through a logic gate. Fig. 2(a) shows the adjacency matrix of a simple circuit encompassing one AND and one NOT gate (top row). Notice that there is also a connection at the entry (2,2), *i.e.* there is a self-loop edge at the node corresponding to the circuit output. This feature is automatically added by the used toolbox converting Verilog netlists, in order to reflect the notion of accumulation of information at the output, since it cannot flow anywhere from there.

The output of an information flow through the circuit can be modelled as $\mathbf{A}\boldsymbol{x}$, where $\boldsymbol{x}$ is the column vector of the input. Note that this is not the same as the function of the circuit in terms of voltages and currents as in modified nodal analysis. Here, all gates are represented as adders. Moreover, the final output of the simulated flow in the case of the circuit in Fig. 2 is given by $\mathbf{A}^2\boldsymbol{x}$. In general, this corresponds to a vector $\mathbf{A}^l\boldsymbol{x}$, where l is the number of gate levels in the hierarchy of the circuit's implementation.

Given a graph described by its adjacency matrix $\mathbf{A}$, the eigendecomposition of the graph/matrix refers to the set of n eigenvectors $\boldsymbol{u_i}$ concatenated into matrix $\mathbf{U}$, and the set of n scalar eigenvalues λ_i forming the diagonal matrix $\mathbf{\Lambda}$. These are found as the solution of the equation(s):

$$\mathbf{A}\cdot\mathbf{U}=\mathbf{U}\cdot\mathbf{\Lambda} \qquad (1)$$

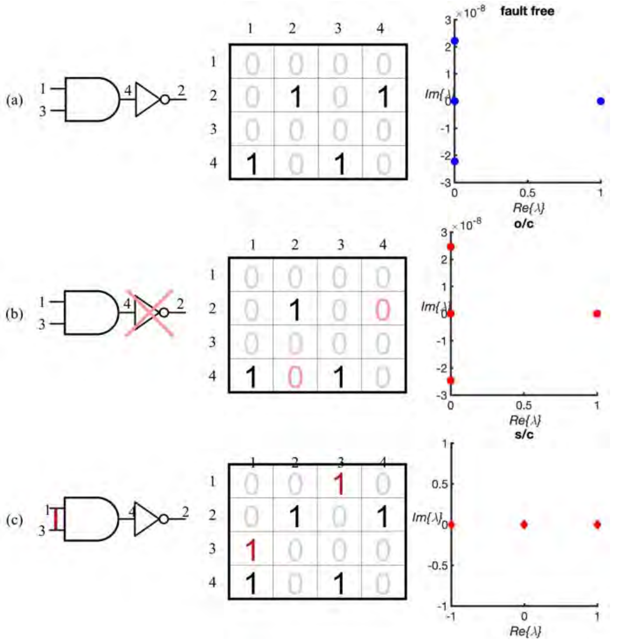


Fig. 2. Schematic of a logic circuit, adjacency matrix of its graph representation, and eigenvalues of the matrix in the case of fault free circuit (a), and in the presence of one open-circuit (b), or short-circuit (c) fault.

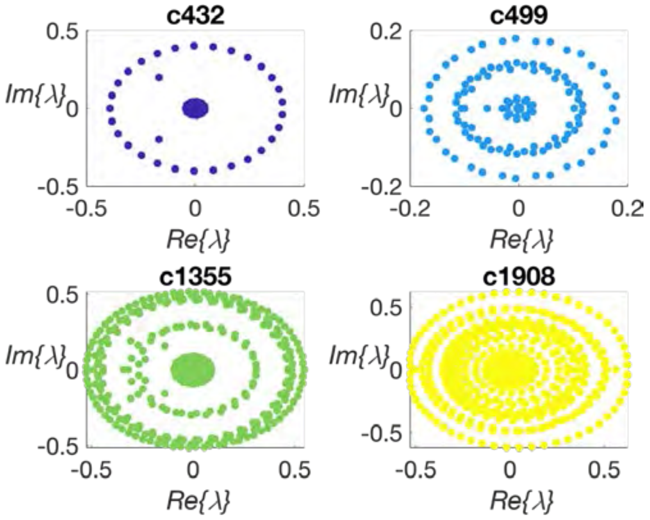


Fig. 3. Eigenvalues of benchmark circuits

Note that adjacency matrices of the available benchmark graphs are not symmetric, *i.e.* graphs are directed. Thus, $\mathbf{A}$'s eigenvalues and eigenvectors are complex-valued, and only *right* eigenvectors are considered. The set of graph eigenvalues λ_i is of crucial importance to many concepts of system analysis such as stability [8], and -as it will be demonstrated in the remaining of this paper- robustness and fault identification. Fig. 2 shows eigenvalues in the complex plane of the toy circuit. Furthermore, note in the middle and bottom rows of Fig. 2 how an open-circuit and a short-circuit fault can be modelled, *i.e.* how they affect the adjacency matrix and the consequent set of eigenvalues. Both types of faults change the resulting eigenvalue distribution, and in this particular circuit, the effect of short-circuit faults is more prominent.

Moving to the analysis of ISCAS'85 benchmarks, eigenvalue empirical distributions for the fault free, and faulty circuits are reported. Fig. 3 shows the eigenvalues (in the complex plane) of the four used benchmark circuits. The eigenvalues are situated in several concentric circles the number of which, together with the density of λ_i, increases with the graph size n.

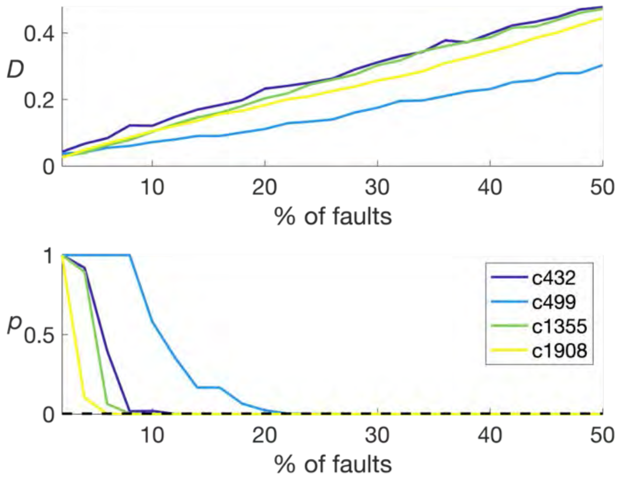


Fig. 4. Kolmogorov-Smirnov statistics (D) and p-values of test on eigenvalue distributions between the original and networks with open-circuit faults.

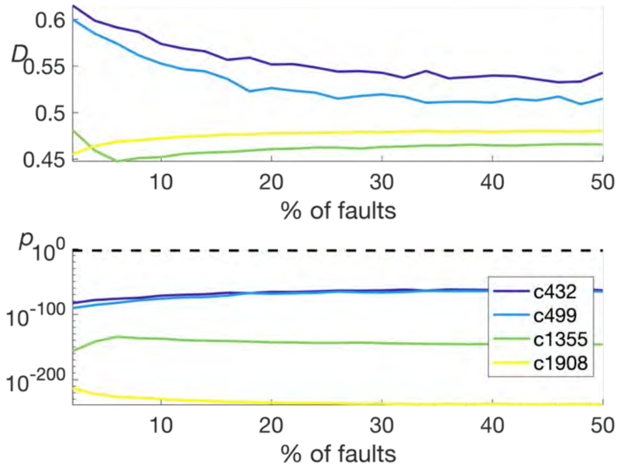


Fig. 5. Kolmogorov-Smirnov statistics (D) and p-values of tests on eigenvalue distributions between the original and networks with short-circuit faults

Two types of faults are considered: short circuit (s/c) and open circuit (o/c). An o/c is "put" in the original circuit by deleting an edge, *i.e.* setting a randomly chosen nonzero entry of **A** to 0 (Fig. 2., middle row). Different levels of gravity of the fault are modelled by choosing the number of put open circuits as a percentage of available connections in the network. In a similar fashion, an s/c fault is reflected in adding a certain value on top of an entry of **A** regardless of whether there was a nonzero edge or not. In this case, the percentage of faults is calculated *w.r.t.* all possible connections between the network nodes, that is $n(n-1)/2$. The added value is chosen to be equal to 1, so as to model the transfer of one unit of information. Note that in both types of faults, the modification of **A** is symmetric since, as opposed to gated connections which are by nature one-directional, o/c and s/c connections are bidirectional.

For each of the values of percentage of faults (from 2 to 50 with step size 2), there were 20 faulty circuits generated. The empirical distribution functions (EDF) of eigenvalues of these circuits are then compared to EDFs of the original circuits in Figs. 4 and 5, and illustrated in Figs. 6 and 7 (up to the fault percentage of 10%). The comparison involved Kolmogorov-Smirnov statistical test designed to determine if there is significant difference between the eigenvalue distributions of the fault free and a faulty circuit. The test rejects the null hypothesis of the same underlying distribution for a high enough value (depending on the chosen significance level α) of the test statistic $D = \sup_t |F_1(t) - F_2(t)|$, where $F_1(t)$ and $F_2(t)$ are empirical cumulative distributions of the given samples. A version of the test extended to 2-dimensional distributions was used here since the eigenvalues are complex in general [9,10]. A high (low) value of Kolmogorov-Smirnov statistic D suggests more (less) distinct probability distributions, further suggesting higher (lower) influence of the faults on the spectral features of the graph.

In the presented analysis, statistical test was performed multiple times, *i.e.* for each of the 4 circuits, each of 25 fault percentages, and for both types of faults, resulting in 200 tests. High number of tests increases the probability of false positive results. Hence, Bonferroni correction was applied before reporting the corresponding p-values. For the desired total significance level of $\alpha = 0.05$, this means that the reference significance level for each test should be $\alpha/200 = 0.0025$. Results of statistical tests are presented in Figs. 4 and 5, including corrected alpha level as dashed black line. For clarity, the p-values in the case of short-circuit faults are given in logarithmic scale (Fig. 5).

Finally, continuing on the observation that a dominant eigenvalue (with much higher magnitude that others) appears in s/c circuits (Fig. 7), the eigendecomposition was performed for single-fault circuits derived from c432 by modifying connections between specific pairs of nodes (red edges in Fig. 1). The magnitudes of eigenvectors corresponding to the emerged dominant eigenvalue are presented in Fig. 8. As it is explained in the following section, these could be exploited for fault identification.

C. Discussion

This subsection discusses how the previously reported results on eigenspectra relate to the concepts of robustness and fault detection in the analysis of electronic circuits. The postulates of linear algebra used as a basis of the following discussion are adapted from [11].

As the circuit flow is modelled with $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$, the problem of designing a robust circuit, that is, one with robust information flow, relates to the theory of perturbation and matrix sensitivity analysis. If the sensitivity, *i.e.* partial derivative of output signal with respect to circuit topology, is very low, one considers to have a robust circuit – $|\partial \boldsymbol{y}/\partial \boldsymbol{A}| \approx 0$. On the other hand, eigendecomposition of **A** provides an alternative model of circuit flow in terms of

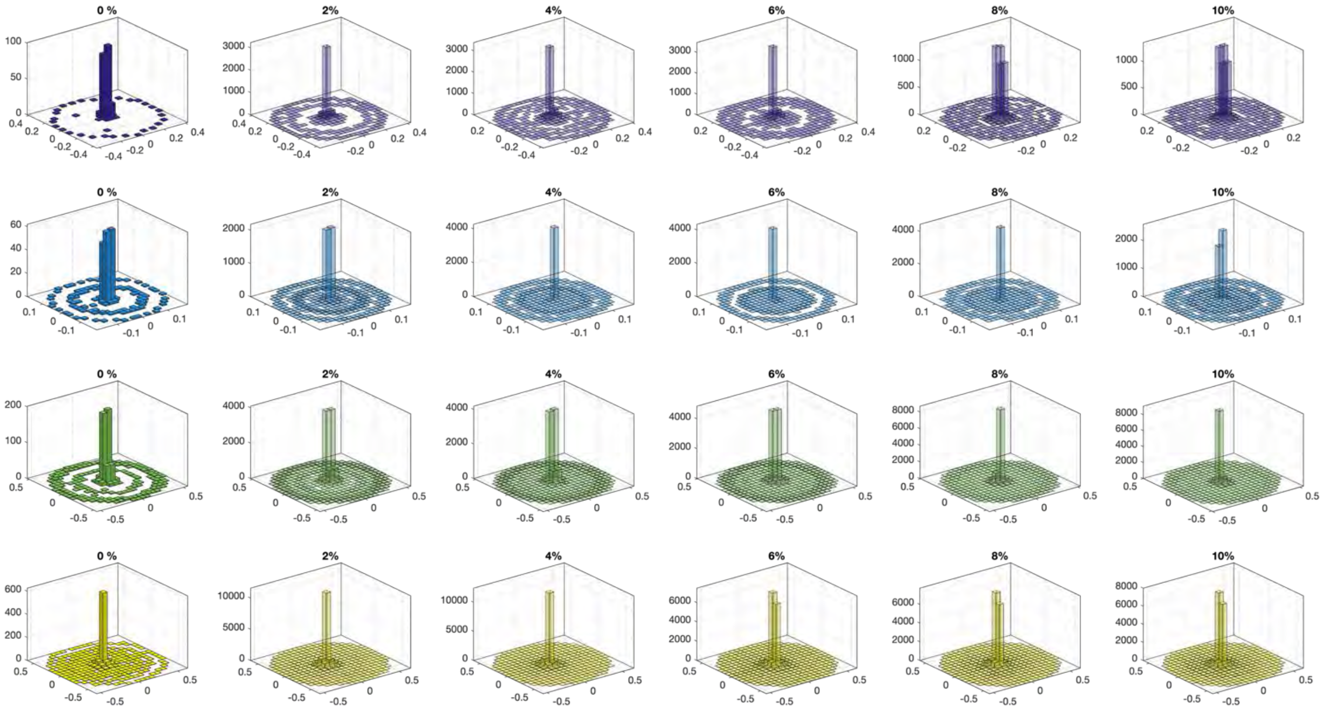


Fig. 6. Histograms of eigenvalues of benchmark circuits (without and with open-circuit faults)

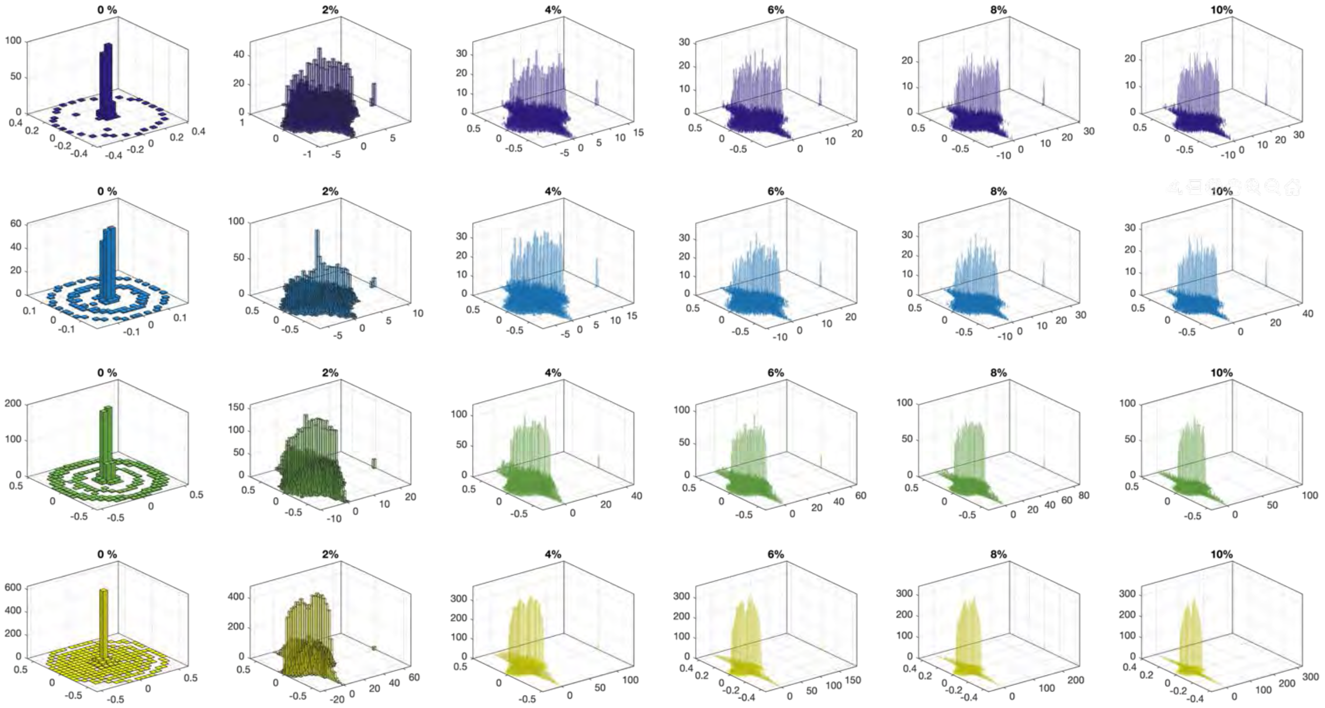


Fig. 7. Histograms of eigenvalues of benchmark circuits (without and with short-circuit faults)

eigenvectors $\boldsymbol{u_i}$ and eigenvalues λ_i as $\boldsymbol{y} = \sum_{i=1}^{n} \lambda_i \boldsymbol{u_i} \boldsymbol{v}_{\mathrm{i}}^{\boldsymbol{T}} \boldsymbol{x}$, where $\boldsymbol{v_i}$ is a corresponding left eigenvector (for which the matrix form is $\mathbf{V} = \mathbf{U}^{-1}$). Eigenvectors and eigenvalues are in fact functions of each other but for simplicity one can consider the sensitivity of circuit output to be defined by eigenvectors since they can be seen as specific topological principal components of the circuit. Hence, under a small perturbation matrix $\mathbf{E}$ that models s/c or o/c entries, a robust circuit $\mathbf{A}$ would be one for which the eigenvectors $\widetilde{\boldsymbol{u}}_{\boldsymbol{i}}$ of $\boldsymbol{A} + \boldsymbol{E}$ span similar vector subspaces as $\boldsymbol{u_i}$. Of general interest is to find features of circuits $\mathbf{A}$ which satisfy this property. As seen from the results of statistical testing (Figs. 4 and 5) a significant difference between eigenvalue distributions emerges at the highest percentage of (o/c) faults for c499. In other words, this circuit is the most robust, and its robustness can be related to the initial eigenvalue density. Indeed, from perturbation theory one finds that:

$$\sup \{\mathrm{dist}(\mathrm{span}\{\boldsymbol{u}_i\}, \mathrm{span}\{\widetilde{\boldsymbol{u}}_{\boldsymbol{i}}\})\} \propto \frac{1}{\min_{j \neq i} \{\lambda_i, \lambda_j\}} \tag{2}$$

If an eigenvalue is very distant from others, then the upper bound of the distance between the spans of its corresponding eigenvectors of original and perturbed matrices is lower. Hence, for a circuit with highly distinct eigenvalues, the effect faults have on its eigenvectors remain tightly upper-bounded. In other words, the circuit flow is preserved under small changes of its connections. Setting aside the particularly robust c499, a general trend is observed that if the circuit has higher number of nodes it is less robust (the significance is reached for lower percentage of o/c faults – Fig. 4, and D has much lower values for low percentages of s/c faults – Fig. 5). This happens since all presented circuits have eigenvalues in more or less similar range and yet higher number of nodes implies higher number of eigenvalues, finally rendering the mutual distances between them lower.

Though actual difference of information flow is reflected in the exact values of eigenvectors and eigenvalues of perturbed matrices, intuitively one can inspect the robustness just by looking at the compact support of the eigenvalue distribution. Hence, it can be observed that the benchmark circuits are much less robust to s/c faults than to o/c faults. Indeed, the support of EDFs with o/c remains circular regardless of the percentage of faults (Fig. 6), whereas the support of EDFs with s/c distorts into a star-shaped or a circular one with prominent tails (Fig. 7).

The high sensitivity to s/c faults (Fig. 5) is partially compensated by the potential ability to detect particular faults, *i.e.* to identify the location (pair of nodes/vertices) where the fault originates. Indeed, Fig. 7 shows that in presence of s/c faults the circuits exhibit a dominant eigenvalue with magnitude much higher than all others. Even though the output $\boldsymbol{y}$ is often completely destroyed with just a single-connection fault, the particular distribution of eigenvalues provides a way to (partially) identify the fault. This can be explained by the power iteration, a method for calculating the eigenvector corresponding to the dominant eigenvalue (assuming $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$). With k iterations, the vector $\boldsymbol{x^{(k)}} = \mathbf{A}^{\boldsymbol{k}} \boldsymbol{x}$ converges to the eigenvector $\boldsymbol{u_1}$ regardless of the input $\boldsymbol{x}$. There is a condition to be satisfied that $\boldsymbol{x}$ should contain a component in the direction of the dominant eigenvector ($\boldsymbol{x}$ not orthogonal to $\boldsymbol{u_1}$), but this condition is almost surely satisfied in practice. It is known that the distance between spans of the eigenvector $\boldsymbol{u_1}$ and its k-iterated approximation $\widehat{\boldsymbol{u}}_1^{(k)}$ is of order of magnitude that depends on the ratio between the two highest-magnitude eigenvalues:

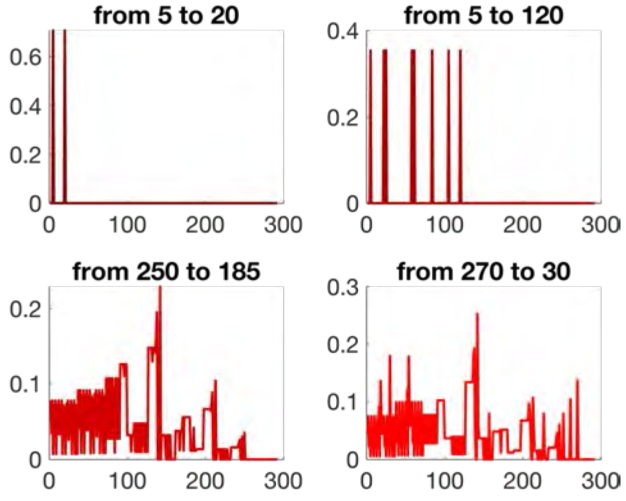


Fig. 8. Magnitude of dominant eigenvectors in cases of single short-circuit faults (red edges in Fig. 1)

$$\mathrm{dist}\left\{\mathrm{span}\{\boldsymbol{u_1}\}, \mathrm{span}\left\{\widehat{\boldsymbol{u}}_1^{(k)}\right\}\right\} = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \tag{3}$$

Indeed, if the dominant eigenvalue λ_1 has significantly higher magnitude than its successor λ_2 , the algorithm can reach convergence in just one iteration, implying that

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} \approx \boldsymbol{u_1} \neq f(\boldsymbol{x}) \tag{4}$$

This stable circuit output (at all considered nodes), that does not depend on the input vector $\boldsymbol{x}$, can be seen as a particular distribution of information driven by an internal connection fault (obstructed flow in case of o/c, enhanced flow in case of s/c). However, since a different location of an s/c fault may induce a different output in Eq. (4), one can identify the fault's location by recognizing the signature of the particular output, previously simulated as the dominant eigenvector. These eigenvectors are plotted in Fig. 8 for particular locations of faults (Fig. 1) in c432. As the eigenspectra are complex-valued in general, only the magnitudes are plotted here as an approximation of the real-world output that would be generated as a consequence of the fault. For shown faults, dominant eigenvectors can clearly be distinguished, allowing for the fault identification. However, in general, the dominant eigenvector as simulated here may be quite similar for two or more different single s/c faults. Still, this simulation provides at least a way to reduce the number of possible locations of a fault.

Since the considered eigenvectors are of the same nature as the graph model, *i.e.* the information flow, and not circuit function, a question is raised on how to measure them in practice during a physical test. In order to have physical measurements of the information flow in the circuit, one needs to simulate the propagation of a signal throughout the circuit unconstrained by the particular gates. Conceptually, this may be accomplished by exploiting the RF (radio-frequency) features of the circuit. The transistors exhibit parasitic impedances connecting their terminals, meaning that an AC signal of high frequency can be propagated regardless of the DC logic input, as long as the condition for the active operating mode is fulfilled. This may be significantly challenging in an existing circuit, since the AC output at the logic gate may not be the sum of inputs due to different delays imposed by transmission lines and capacitances, and consequently the phase difference. However, the circuit can be originally designed in IC technology in a way that allows the signal of certain frequency to be propagated throughout the circuit. Then, the fault identification would comprise of (a) making a dictionary of faults-to-eigenvectors based on the described methodology, (b), driving the circuit with the AC signal and recording the output (at different probing nodes; not necessarily just at the functional circuit output), and (c) finding the eigenvector closest in form to the recorded signal, ultimately informing about the corresponding fault location from the dictionary. Ideally, the amplitudes of AC signals at probed nodes should reflect the structure of the magnitude of an eigenvector. Analogously, the presence of multiple faults could also be reflected in the (larger) eigenvector dictionary. Implementation of the described fault identification procedure is certainly intricate due to complex behaviour of circuits at high frequencies, and its specific challenges fall beyond the scope of this paper. Nonetheless the method represents a conceptual framework one can explore further for practical use in testing and design for testability.

III. CONCLUSION

This paper provided insight into parallels between matrix algebra on graphs and logic circuits' (represented by graphs) robustness and capability of fault identification. Main advantages of the described simulation approaches are (a) availability of a general measure of robustness (percentage of faults at which divergence between eigenvalue distributions gets statistically significant) appropriate for comparing electrical networks of different size and function, and (b) possibility to -through a look-up table (fault location → dominant eigenvector)- ease the fault diagnostics (detection and identification).

On the other hand, certain limitations exist that need to be addressed before further extensions of the described approach: (a) computation of eigendecomposition of very sparse graphs (number of edges few orders of magnitudes lower than number of vertices) may be numerically unstable; (b) in the current form, the graphs are described with binary adjacency matrices thus hindering the possible modelling of partially obstructed flow (degrading/aging components); (c) the RF approach to fault identification needs to be extended from the conceptual to practical framework. Hopefully, work in this area will provide a fully functional framework for exploration of logic circuits' robustness, reliability, and similar concepts.

REFERENCES

[1] Hwang, S., Rajsuman, R., "VLSI Testing for High Reliability: Mixing IDDQ Testing with Logic Testing", *VLSI Design*, vol. 5, no. 3, pp. 299–311, Jan. 1997.

[2] Kundu, S., Reddy, S. M., Jha, N. K., "On the Design of Robust Multiple Fault Testable CMOS Combinational Logic Circuits", in *The Best of ICCAD*, Springer US, 2003, pp. 575–584.

[3] Raji, M., Pedram, H., Ghavami, B., "Soft error rate estimation of combinational circuits based on vulnerability analysis", *IET Computers & Digital Techniques*, vol. 9, no. 6, pp. 311–320, Nov. 2015.

[4] Dorfler, F., Simpson-Porco, J. W., Bullo, F., "Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 977–1005, May 2018.

[5] Dorfler F., Bullo, F., "Kron Reduction of Graphs with Applications to Electrical Networks", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, Jan. 2013.

[6] Hansen, M. C., Yalcin, H., Hayes, J. P., "Unveiling the ISCAS-85 benchmarks: a case study in reverse engineering," *IEEE Design & Test of Computers*, Vol. 16, No. 3, pp. 72–80, 1999.

[7] Schmitz, K., Stoppe, J., Drechsler, R., "Verilog2GEXF Dynamic Large Scale Circuit Visualization", *I4th Workshop on Design Automation for Understanding Hardware Designs DUHDe*, Lausanne, Switzerland, Mar. 2017.

[8] Zhou, B., "On asymptotic stability of linear time-varying systems," *Automatica*, vol. 68, pp. 266–276, Jun. 2016.

[9] Peacock, J. A., "Two-dimensional goodness-of-fit testing in astronomy", *Monthly Notices Royal Astronomy Society*, Vol. 202, No. 3, pp. 615-627, Mar. 1983.

[10] Muir, D., kstest_2s_2d(x1, x2, alpha), *MATLAB File Exchange, Online:* (https://www.mathworks.com/matlabcentral/fileexchange/38617-kstest_2s_2d-x1-x2-alpha), 2019.

[11] Golub, G. H., Van Loan, C. F., "Matrix computations", *The Johns Hopkins University Press*, Baltimore and London, 1996.

Approximately linear phase notch filters with magnitude characteristic symmetry

Goran Stančić, Ivana Kostić, Miloš Živković and Ivan Krstić

Abstract - A new method for design and software realization of approximately linear phase notch filters with magnitude characteristic symmetry is presented in this paper. The proposed configuration has parallel nature with pure delay in one path and allpass filter in the another. The resulting filter fulfills all predefined specifications in a wide range of maximal allowed attenuation i.e. for arbitrary given maximal passband attenuation and stopband width, passband boundary frequencies are symmetrical about given notch frequency. The efficiency of the presented method is illustrated on few examples.

Keywords - Notch filter, allpass filter, parallel connection, comb filter.

I. INTRODUCTION

The notch filters are invariably used in communication, control, instrumentation, and bio-medical engineering, besides a host of other fields, to eliminate noise and power line interferences [1-5]. The notch filter highly attenuates a particular frequency component in the input signal while leaving nearby frequency components in the ideal case unchanged. For example, the elimination of a sinusoidal interference corrupting a signal, such as the 50 Hz power-line interference in the design of a digital instrumentation system, is typically achieved with a notch filter tuned to the frequency of the interference. Usually very narrow notch characteristic is desired to filter out the single frequency or sinusoidal interference without distorting the signal of interest [6]. The notch filters find implementation in both, analog and digital domain. Analog implementations consume less bandwidth, but the signals are more likely to get deteriorated during transmission. In contrast, digital filters are generally noise-immune and flexible in implementation, although consume a lot more bandwidth to carry the same information [1][7].

Digital notch filters can be designed as infinite impulse response (IIR) as well as finite impulse response (FIR) structures. In situations where linearity of the phase is not important, IIR filters are preferred since these require much lower order than corresponding the FIR ones for the same set of magnitude response specifications. The standard analog notch filter can be starting point to obtain corresponding digital IIR notch filter by using bilinear transformation for example. One of the major problem in IIR filters design is nonlinear phase response and, therefore, phase distortion introduced in general [8]. However, it is possible to reduce phase distortion by cascading an all-pass phase equalizer [6].

FIR filters, on the other hand, are unconditionally stable and can be designed to provide exact linear phase characteristics. They find extensive use in applications where frequency dispersion due to nonlinear phase is undesirable, such as in speech processing, digital communication, image processing, etc.. This is the specific reason that a large number of commercial chips carry out signal processing with FIR filters. Standard FIR filter design methods, such as windowing, frequency sampling and computer-aided/optimization may be used for designing FIR notch filters. However, most of these methods result in ripples in the passbands [9].

The rest of the paper is structured as follows. In the Section II, the basic relations are derived between notch magnitude and allpass filter phase. A method for determination of filter coefficients and few examples are done in Section III. The results of the simulation of designed filter performance are given in Sections IV and V along with difference equations implemented in the Matlab® in order to obtain output of the notch filter.

II. PROBLEM DEFINITION

In this section the synthesis of linear phase filters with arbitrary number of notch frequencies with magnitude characteristic symmetry will be discussed. The notch filter will be realized as parallel connections of two allpass sub-filters. The coupled allpass structure offers convenient way to solve the problem of filter design thanks to straightforward dependence of notch filter's magnitude and phase characteristic of corresponding allpass sub-filter. A problem of design of notch filter magnitude is easy to reformulate as the allpass filter phase approximation problem.

The realized notch frequency will be positioned exactly at predefined location, while at the same time cut-off edge frequencies are symmetrical regardless the value of the notch frequency. Practically, the full control of the magnitude characteristic at three points is achieved in vicinity of every notch frequency [2]. The location of notch

Goran Stančić, Ivana Kostić and Miloš Živković are with the Dep. of Electronics, Uni. of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: goran.stancic@elfak.ni.ac.rs.

Ivan Krstić is with the Faculty of Technical Sciences at University of Priština, 38220 Kosovska Mitrovica, Serbia.

frequency ω_n, maximal allowed attenuation in passbands a (given in dB) and stopband width B_w are design input parameters. Stopband lower ω_l and upper ω_r edge frequencies are

$$\begin{aligned}\omega_{l_i} &= \omega_{n_i} - B_{w_i}/2 \\ \omega_{r_i} &= \omega_{n_i} + B_{w_i}/2\end{aligned} \tag{1}$$

based on the mentioned symmetry. Usually, for attenuation at cut-off frequencies the value of 3 dB is adopted. The proposed method allows arbitrary positive value for maximal attenuation in passbands to be chosen.

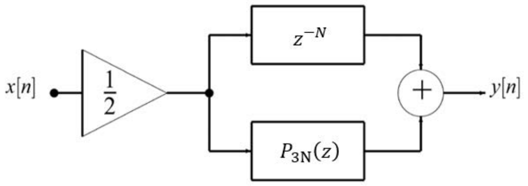


Fig. 1. Realization structure for linear phase filters with N notch frequencies using parallel connection of two alpass sub-filters

The transfer function of the proposed linear phase notch filter, realized using pure delay and allpass sub-filter, presented in the Fig. 1, can be written in the next form

$$H(z) = \frac{1}{2}[z^{-N} + P_{3N}(z)] \tag{2}$$

This solution also allows realization of complementary filter for signal extraction, using only one additional adder [10]. Transfer function of complementary filter is

$$G(z) = \frac{1}{2}[z^{-N} - P_{3N}(z)] \tag{3}$$

and it is represented in the Fig. 2.

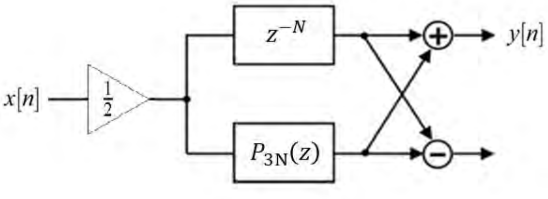


Fig. 2. Realization of complementary filter

The transfer function of subfilter $P_{3N}(z)$ is given with

$$P_{3N}(z) = \frac{p_{3N} + \cdots + p_1 z^{-(3N-1)} + z^{-3N}}{1 + p_1 z^{-1} + \cdots + p_{3N} z^{-3N}}$$

$$P_{3N}(z) = z^{-3N}\frac{D(z^{-1})}{D(z)} \tag{4}$$

$$D(z) = 1 + p_1 z^{-1} + \cdots + p_{3N} z^{-3N}$$

The phase of allpass filter $P_{3N}(z)$ is

$$\varphi(\omega) = -3N\omega + 2atan\frac{p_1\sin(\omega) + \cdots + p_{3N}\sin(3N\omega)}{1 + \cdots + p_{3N}\cos(3N\omega)} \tag{5}$$

By substituting Eq. (4), alongside with $z = \exp(j\omega)$, into Eq. (2), after simple mathematical manipulations the magnitude characteristic of the notch filter could be obtained

$$|H(e^{j\omega})| = \left|\cos\frac{-N\omega - \varphi(\omega)}{2}\right| \tag{6}$$

The phase of notch filter has value

$$argH(e^{j\omega}) = \frac{-N\omega + \varphi(\omega)}{2} \tag{7}$$

Proposed filter's phase approximates piecewise linear phase in all passbands. In order to provide a symmetric magnitude characteristic of notch filter it is necessary that the number of coefficients we determine to be three times greater than given number of notch frequencies. This approach allows that the maximum attenuation in all passbands can be independently controlled. Maximum phase approximation error e_i in the i-th passband is given by Eq. (8) where a_i is attenuation in i-th passband given in dB.

$$e_i = 2arccos\left(10^{-a_i/20}\right) \tag{8}$$

III. DETERMINATION OF FILTER COEFFICIENTS

Structure for realization of filter with one notch frequency is composed of first order delay and third order allpass phase corrector which is explained in detail in paper. Described approach can be generalized and applied to filter with N notch frequencies. In Fig. 3 points of interest are marked, notch frequency ω_n and passband egde frequencies ω_l and ω_r. Based on these frequencies, a system of equations is formed to determine the coefficients of filter [11].

The filter with one notch frequency, which structure is given in Fig. 1, can be made in two ways. The first approach is that bare wire i.e. zero-order delay line be in one path and second order allpass section in another path of parallel structure. Only two coefficients need to be determined i.e. to form two equations. It is valid to use two points of interest ω_n and ω_r for $\omega_n < 0.5\pi$ or ω_n and ω_l for $\omega_n > 0.5\pi$. In that case, magnitude characteristic will not be symmetric. Allpass filter with no multiple poles does not offer an equiripple phase approximation. In the other case, if an eqqiripple approximation of the linear phase should be achieved, allpass filter must be at least of 4th order. Double pair of complex conjugate poles will ensure the phase jump of 2π at notch frequency.

An explanation for determining the coefficients of filter with two notch frequencies is given below. The structure of the filter is consisted of 2nd order pure delay and 6th order allpass filter. The phase of proposed notch filter is a monotonically decreasing function of frequency. Phase approximation error is also monotonic function with maximum approximation error located at boundary frequencies, as given in Fig. 4.

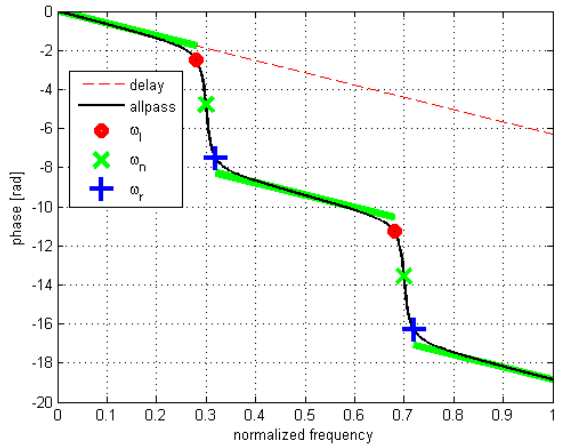


Fig. 3. Phase of delay line, allpass filter and points of interest ωn, ωl and ωr for filter with two notch frequencies

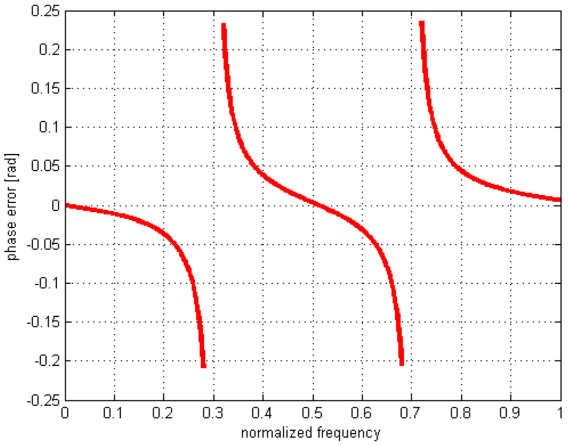


Fig. 4. Phase error of notch filter with two notch frequencies

According to Fig. 4, for IIR filter with two notch frequencies, next set of equations could be defined.

$$-2\omega_{n_1} - \pi = -6\omega_{n_1} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{n_1})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{n_1})} \tag{9}$$

$$-2\omega_{l_1} - e = -6\omega_{l_1} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{l_1})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{l_1})} \tag{10}$$

$$-2\omega_{r_1} - 2\pi + e = -6\omega_{r_1} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{r_1})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{r_1})} \tag{11}$$

$$-2\omega_{n_2} - 3\pi = -6\omega_{n_2} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{n_2})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{n_2})} \tag{12}$$

$$-2\omega_{l_2} - 2\pi - e = -6\omega_{l_2} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{l_2})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{l_2})} \tag{13}$$

$$-2\omega_{r_2} - 4\pi + e = -6\omega_{r_2} + 2arctg\frac{\sum_{i=1}^{6} p_i \sin(i\omega_{r_2})}{1 + \sum_{i=1}^{6} p_i \cos(i\omega_{r_2})} \tag{14}$$

Which could be given in matric form

$$\boldsymbol{Ap} = \boldsymbol{b} \tag{15}$$

from which coefficients

$$\boldsymbol{p} = [p_1 \quad p_2 \quad \cdots \quad p_{3N}]^T \tag{16}$$

will be obtainted.

The proposed approach could be generalized and applied to case of IIR notch filter with arbitrary number of notch frequencies N. After some manipulations, matrix ***A*** of sistem of Eq. (9) - Eq. (14) in case of N notch frequencies can be given with

$$\begin{gathered} A(i,j) = (-1)^{i+1}\cos\left((2N-j+1)\omega_{n_i}\right) \\ A(N+i,j) = (-1)^{i+1}\sin\left((2N-j+1)\omega_{l_i} + e/2\right) \\ A(2N+i,j) = (-1)^{i}\sin\left((2N-j+1)\omega_{r_i} - e/2\right) \\ i = 1,2,\dots,N \qquad j = 1,2,\dots,3N \end{gathered} \tag{17}$$

Elements of vector ***B*** have value

$$\begin{gathered} b(i) = (-1)^{i}\cos\left(N\omega_{n_i}\right) \\ b(N+i) = (-1)^{i+1}\sin\left(N\omega_{l_i} - e/2\right) \\ b(2N+i) = (-1)^{i}\sin\left(N\omega_{r_i} + e/2\right) \\ i = 1,2,\dots,N \end{gathered} \tag{18}$$

IV. EXAMPLES

For different values of stopband width B_w and allowed attenuation *a* in passband below are presented two figures which give position of the poles and zeros of notch filter with four notch frequencies ω_{n_i} = 0.1π, 0.2π, 0.4π and 0.8π.

According to the obtained results, displayed in Fig. 5 and 6, one can conclude that the reducing allowed attenuation in the passband make poles and zeros to approach closer to the unit circle. It can cause a problem in some implementations with restricted number of bits dedicated for filter coefficients representation because poles are too close to the edge of stability region and can leave the unit circle.

In case B_w = 0.09π and *a* = 3dB, phase characteristic of allpass sub-fitler is given in Fig. 7. It is noticeable that the highest phase slope is observed at the frequencies where the poles are located, i.e. around notch frequencies.

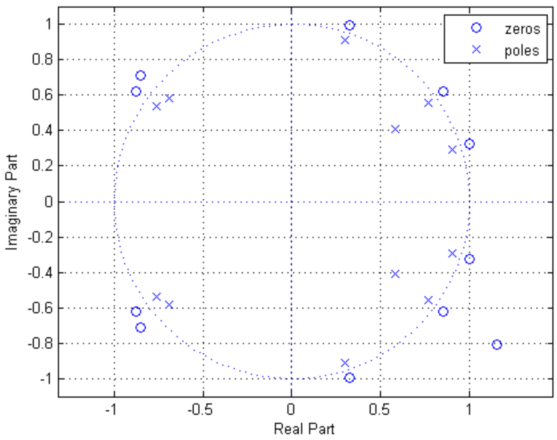


Fig. 5. Poles and zeros of allpass filter, $B_w = 0.09\pi$ and $a = 3$ dB

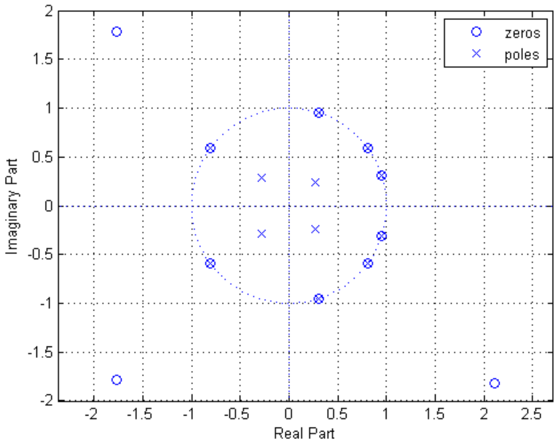


Fig. 6. Poles and zeros of allpass filter, $B_w = 0.09\pi$ and $a = 0.01$ dB

For any set of notch filter input specifications system of equations Eq. (15) can be solved. As mentioned earlier, one can not control passband magnitude. For some input parameters (when one notch frequency is far away from another) maximal obtained attenuation in passband could surpass the given value at passband boubdary frequency.

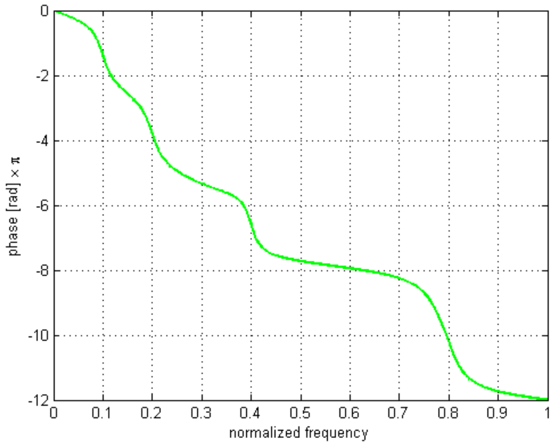


Fig. 7. Typical phase characteristic of allpass sub-filter

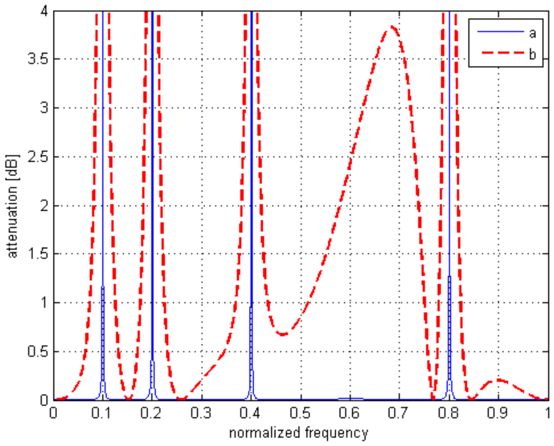


Fig. 8. Attenuation of filter with four notch frequencies, $B_w = 0.09\pi$ a) $a = 0.01$ dB and b) $a = 3$ dB

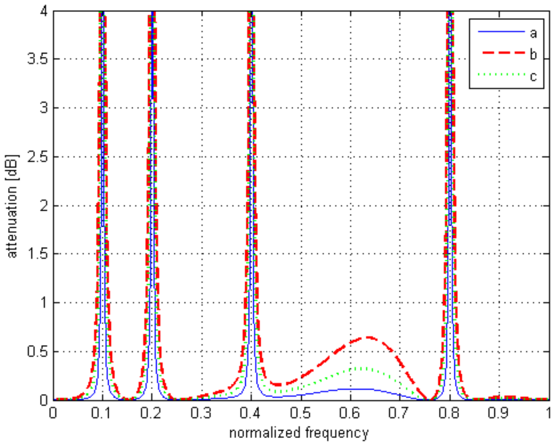


Fig. 9. Influence of bandwidth B_w on attenuation curve for $a = 1$ dB a) $B_w = 0.09\pi$, b) $B_w = 0.07\pi$ and c) $B_w = 0.03\pi$

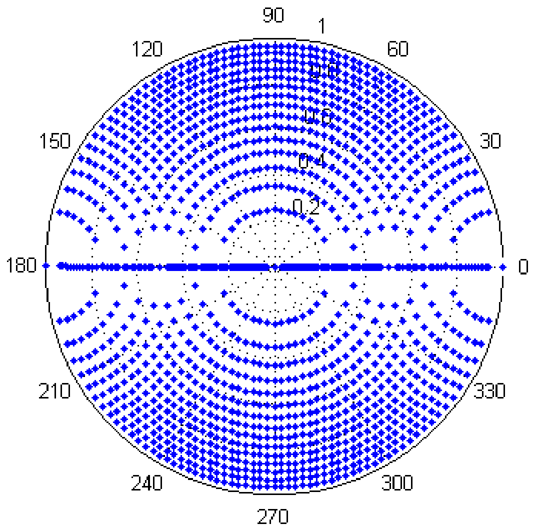


Fig. 10. Possible position of allpass filter poles for four bit fixed-point fraction representation

The Table I and Fig. 9 show that by reducing the value of width of stopband B_w and attenuation a at passband edges, the maximum attenuation in the passband (which is out of control by system of equations) is also reduced, but that the ratio a/a_{max} changes very little in the wide range of parameters B_w and a.

TABLE I

MAXIMAL ATTENUATION IN PASSBAND $a_{\max}$ AND $a/a_{\max}$ FOR DIFFERENT VALUES OF BANDWIDTH B_W AND ATTENUATION a

$a_{\max}$ / $a/a_{\max}$		a [dB]			
		3	1	0.1	0.01
B_w	0.09π	3.83 0.78	1.11 0.90	0.12 0.85	0.01 0.81
	0.07π	2.17 0.78	0.64 1.56	0.01 1.51	0.01 1.46
	0.05π	1.10 2.73	0.32 3.13	0.03 3.13	0.003 3.04
	0.03π	0.42 7.23	0.12 8.55	0.02 8.55	0.01 8.77

It can be seen from Fig. 10 that significant errors are possible in the implementation of the filter with the poles with low phase angles. This problem can be solved by larger number of bits dedicated for coefficients representation as well as spesific filter structures which go beyond the scope of this paper.

Described approach provides a solution for a wide range of input parameters (number of notch frequencies N, stopband width B_w and allowed attenuation a) which is shown in Table I. It can be seen from the table (for B_W = 0.09π) that if the bandwidth is too large, the maximal obtained attenuation in the passband (3.83 dB, which one can not control) may be greater than the prescribed attenuation (given at the boundaries of the passband (3 dB) and which need to be extremal value).

The proposed procedure can be also successfully used for realization of comb filters, as shown in the Fig. 11 for the case of a filter with nine notch frequencies $\omega_{n_i} = i\pi/10$.

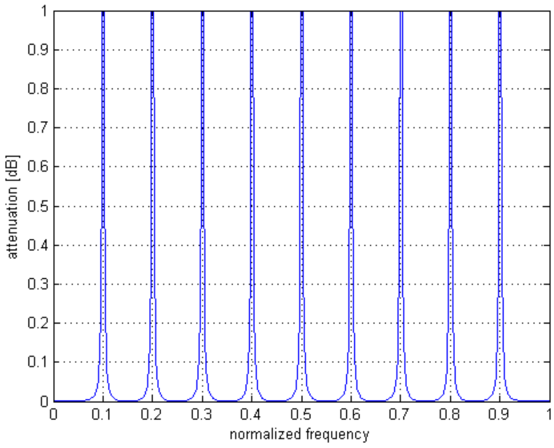


Fig. 11. Attenuation of the comb filter with nine notch frequencies $\omega_{n_i} = i\pi/10$, a =0.5 dB, B_w = 0.04π

V. SOFTWARE REALIZATION AND THE SIMULATION RESULTS

Power systems are designed to operate at frequencies of 60 (America and part of Asia) or 50Hz (in large parts of the world). However, certain types of loads (nonlinear loads) produce currents and voltages with frequencies that are integer multiples of the 50 or 60 Hz fundamental frequency. These higher frequencies are a form of electrical pollution known as power system harmonics. Notch filters have the application to remove these unwanted components.

The Matlab® software package is used for design and realization of the proposed linear phase notch filters with magnitude characteristic symmetry. Sinusoidal noise with amplitude 0.2, at power-line frequency F_n = 50 Hz, 100 Hz and 150 Hz are superimposed on the electrocardiogram (ECG) signal $s[n]$ downloaded from the database MIT-BIH [12], as it is given in Eq. (19). All available signals in MIT-BIH database are recorded after digitalization using sampling frequency F_s = 360 Hz.

Specifications of the digital filter in z domain are: $\omega_n = 2\pi F_n / F_s = 5\pi/18$, while for stopband width and maximal attenuation in passbands are adopted B_w = 0.02π and a = 1 dB, respectively.

The signal duration is 10 seconds so $n = 0, 1, \dots, 3599$.

$$x[n] = s[n] + \sum_{i=1}^{3} 0.2\sin(2\pi i 50 n/F_s) \tag{19}$$

$$y[n] = \frac{1}{2}\left[x[n-3] + \omega[n]\right] \tag{20}$$

$$\omega[n] = x[n-9] + \sum_{i=1}^{9} p_i\left(x[n-9+i] - \omega[n-1]\right) \tag{21}$$

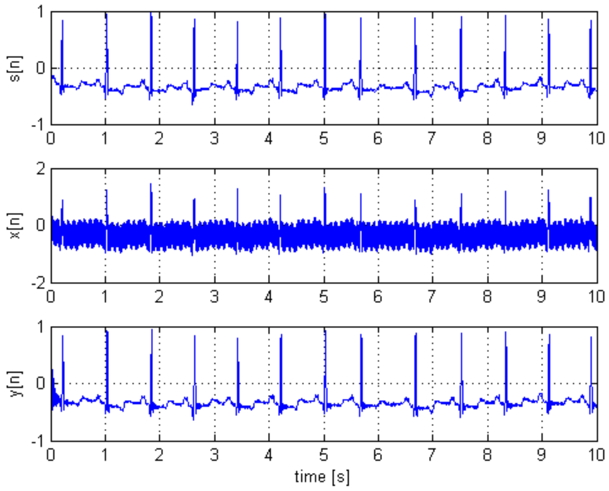


Fig. 12. ECG signal $s[n]$, ECG signal with added sinusoidal noise $x[n]$ and filtered signal at output of notch filter $y[n]$

In Fig. 12, ECG signal $s[n]$ which corresponds to 100th sample in MIT-BIH database is shown, together with corrupted version of this signal $x[n]$ after addition of sinusoidal noise. Filtered version of the signal at output of proposed notch filter $y[n]$ is also presented in Fig. 12.

It is evident that designed notch filter successfully eliminated sinusoidal noise at power-line frequency from ECG signal. In the steady state, output signal $y[n]$ is free of noise. A deviation at the beginning of output signal can be observed. It is normal behavior of digital filters since digital filter output data are valid after filter latency.

VI. CONCLUSION

Design and software realization of IIR digital filter with arbitrary number of notch frequencies (N) is presented in this paper. The proposed filter is realized as parallel structure with Nth order pure delay in one path and allpass sub-filter of order 3N in another path. Parallel connection enables realization of a complementary filter with only one additional adder and this solution exhibits low sensitivity on coefficients quantization.

Unlike existing methods for design of notch filter, the proposed configuration delivers solution with cut-off frequencies symmetry about notch frequency i.e. for arbitrary given maximal passband attenuation and stopband width, passband boundary frequencies are symmetrical about given notch frequency. The resulting filter has approximately linear phase in passbands and minimal order to provide symetry.

The functionality of the presented model is illustrated by filtering ECG input signal corrupted with sinusoidal noise at the power-line frequency and first two harmonics. The proposed method could be also applied on design of comb filter or filter for extraction of signal components at frequencies of interest.

ACKNOWLEDGEMENT

The research presented in this paper is financed by the Ministry of Education, Science and technological Development of the Republic of Serbia under the project TR33035.

REFERENCES

[1] C.T. Ma, P.-I. Mak.-I. Vai, P.-U. Mak, S.-H. Pun, W. Feng and R.P. Martins, “Frequency-bandwidth-tunable powerline notch filter for biopotential acquisition systems”, *Electronics Letters*, vol. 45, no. 4, 2009.

[2] X. Qian, Y. P. Xu and X. Li, “A CMOS continuous-time low-pass notch filter for EEG systems”, *Analog Integrated Circuits and Signal Processing*, vol. 44, pp. 231-238, 2005.

[3] Dr K. Kaur and M. Singh Saini, “Design IIR notch filter and comb filter to eliminate unwanted frequencies for environment”, *IJARCET*, vol. 3 issue 6, 2014.

[4] K. Du and M. Onabajo, “Digitally tunable lowpass-notch filter design for analog front-ends in brain signal measurement applications”, *Analog Integrated Circuits and Signal Processing*, vol. 44, pp. 231-238, 2005.

[5] K. Aboutabikh, I. Haidar and N. Aboukerdah, “Design and implementation of a digital FIR notch filter for the ECG signals using FPGA”, *IJARCCE*, vol. 5, no. 1, pp. 452-456, 2016.

[6] T. Yu, Sanjit K. Mitra and H. Babić, “Design of linear phase FIR notch filters”, *Sadhana*, vol. 15, part 3, pp. 133-135, 1990.

[7] M. Susca and P. Dobra, “Analog and digital notch filter implementation” *Conference: International Conference on Automation, Quality and Testing*, Robotics, Student Forum, At Cluj-Napoca, 2016.

[8] S. D. Roy, B. Kumar and B. Jain, “FIR notch filter design: A review”, *Facta Universitatis Series*, vol. 14, no. 3, pp. 295-327, 2001.

[9] Vaidyanathan, P.P., Regalia P.A. and Mitra S.K., “Design od double complementary IIR digital filters using a single complex allpass filter, with multirate applications”, *IEEE Trans. Cuircuits Syst.*, 34, pp. 378-389., 1987.

[10] I. Krstić, M. Živković, I. Kostić and G. Stančić, “Design of narrow band-stop recursive filter based on third order phase corrector application”, *Proceedings of ICEST 2019*, pp. 95-98, 2019.

[11] G. Stančić and S. Nikolić, “Design of narrow stopband recursive digital filter”, *Facta Universitatis*, vol. 24, no.1, pp. 121-132, 2011.

[12] Y. V. Joshi and S.C. Dutta Roy, “Design of IIR digital notch filters”, *Circuits, Systems and Signal Processing*, vol. 16, no. 4, pp. 415-427, 1997.

[13] G. Stančić and S. Nikolić, “Design of digital recursive notch filter with linear phase characteristic”, *11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services*, pp. 69-72, 2013.

[14] C. M. Wang and W. C. Xiao, “Second-order IIR notch filter design and implementation of digital signal processing system”, *ISCCCA-13*, 2013.

[15] K. Hirano, S. Nishimura and S. Mitra, “Design of Digital Notch Filters”, *IEEE Transactions on Communications*, vol. 22, no. 7, pp. 964-970, 1974.

[16] S. Nikolić , I. Krstić and G. Stančić, “Noniterative design of IIR multiple-notch filters with improved passband magnitude response”, *Int. Journal of Circuit Theory and Applications*, vol. 46, no. 12, pp. 2561-2567, 2018.

[17] G. B. Moody and R. G. Mark “The impact of the MIT-BIH Arrhythmia Database”, *IEEE Eng in Med and Biol,* vol. 20, no. 3, pp. 45-50, 2001.

[18] G. Stančić and S. Nikolić, “Digital linear phase notch filter design based on IIR all-pass filter application”, *Digital Signal Processing*, vol. 23, no. 3, pp. 1065-1069, 2013.

[19] S. Nikolić, G. Stančić: "Design of IIR notch filter with approximately linear phase", *Circuits, Systems & Signal Processing*, vol. 31, Issue 6, pp. 2119-2131, 2012.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

SPICE simulation of memristor logic functions

Luka Spahić, Miljana Milić

Abstract - In this paper, the short review of memristor models applicable for logic operations is given. Further, this paper describes elementary logical functions such as AND, OR and NOT and the application of the developped memristor model for their realization. Logical functions are verified in PSPICE.

Keywords – Memristor, simulation, logic functions, modelling.

I. INTRODUCTION

Currently, we know and we are using three basic circuit elements: a resistor, a capacitor and an inductor. But, in the 1971., the missing fourth element has been proposed by L. Chua [1]. Memristor is a two-terminal circuit element that links electric charge and magnetic flux, and later shows a resistive switching characteristics. Its state is definded by a time integral of current and a time integral of voltage.

At the time when Chua was theoretically proposing the idea of a memristor, it hasn't yet practically been manufactured. Finally, in 2008. R. Stanley William and his team, at HP Labs, made a discovery of the first physical memristor. This memristor was based on TiO_2 material doped with oxygen vacancies and had resistive switching characteristics. They have used a TiO_2 material placed between two platinum (Pt) electrodes, with one of the regions doped with oxygen vacancy (TiO_{2-x}) as shown in Fig. 1. The doped region, marked as *w*, has a lower resistivity than the undoped region. When applying a certain voltage across the element, the doped region will expand or shrink which will result in a resistance change.

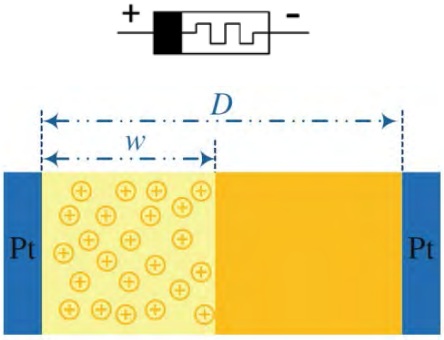


Fig. 1. HP memristor TiO_2 structure [3]

Luka Spahić and Miljana Milić are with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia.
E-mail: luka.spahic@elfak.rs, miljana.milic@elfak.ni.ac.rs

When the voltage is revoked, the states of the oxygen vacancy carriers remain unchanged, and that is why the device can "memorize" its previuse resistance value [2].

Since 2008, many researchers were trying to make a suitable model of memristor that could be used in future applications. The greatest advantage of the memristors is that they are extremely small as compared to the size of MOS transistors. In fact, a memristor can be manufactured with area as small as 9 nm^2 [4]. Due to their ability to memorize previuse states, memristors could be used for non-volatile resistive memory systems (ReRAM) and also because of their extremely small size, they could be very conveniently built in a crossbar array.

Several papers have been published with the description of memristors possible application in the implementation of logic functions. This will also be the subject of this paper. Down below, common design styles are briefly described.

Most common memristor logic models are IMPLY logic and MAGIC design. In IMPLY logic [5], memristors can be used to realize implication operation such as $p \rightarrow q = p' + q$. The initial values of *p* and *q*, and also the result are "stored" as resistance values in memristors. Here we are using two voltages for the initialization, but MAGIC uses only one.

In MAGIC memristor design style [6], memristors are used to implement logic functions/gates, where the inputs are applied as values of resistance. This also applies for the output values.

In this paper, two memristor SPICE models are reviewed. The first one, is a memristor model presented by HP (but SPICE description is given by Biolek [7]), which is considered to be the most efficient. But on the other hand, this model can properly operate for very low frequencies (1 Hz – 5 Hz). The second one, is an improved model [8] with ability to operate on frequencies between 1 kHz – 10 kHz. Next, the SPICE simulations that shows usage of new improved memristor model as type of logic function (OR, AND, NOT) are given in section III. Through out section III simulation results are discussed. Finally, we have come to conclusion of further reaserch of memristors based on discussion in section III.

II. REVIEW OF MEMRISTORS LOGIC MODELS

As previously mentioned, we have encountered two memristors models and one of them has been used for further simulations.

A. SPICE Model of Memristor with Nonlinear Dopant Drift[7]

The electric circuit of the first SPICE model is given in figure 2. To model the relation between voltage and current the following expression could be applied:

$$R_{MEM}(x) = R_{OFF} - x\Delta R, \Delta R = R_{OFF} - R_{ON} \quad (1)$$

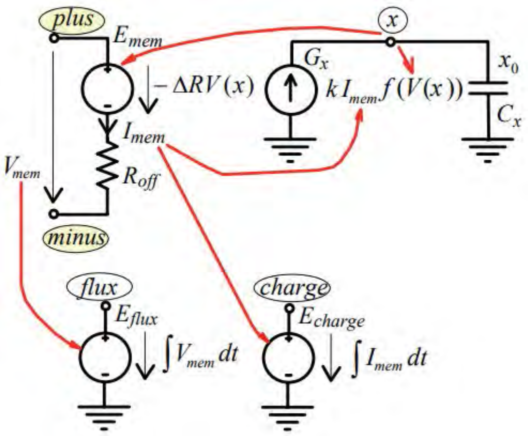


Fig. 2. Electric circuit of SPICE model [7]

As shown in fig. 2, eq. (1) corresponds to R_{OFF} resistor which is in series with E_{MEM} whose terminals are controlled according to eq. "- ΔRx". The width x of the doped layer is modeled by the voltage $V(x)$ of the capacitor C_X. The initial state of the normalized width x of the doped layer x_0, modeld by initial capacitor voltage which depends of the initial resistence R_{INIT} according to formula (1) [7]:

$$X_0 = \frac{R_{OFF} - R_{INIT}}{\Delta R} \quad (2)$$

The memristor library function for this model is as follows:

```
* HP Memristor SPICE Model * For Transient Analysis only *
+created by Zdenek and Dalibor Biolek
*************************
* Ron, Roff - Resistance in ON / OFF States
* Rinit - Resistance at T=0
* D - Width of the thin film
* uv - Migration coefficient
* p - Parameter of the WINDOW-function
* for modeling nonlinear boundary conditions
* x - W/D Ratio, W is the actual width
* of the doped area (from 0 to D) *
.SUBCKT memristor Plus Minus PARAMS:
+ Ron=1K Roff=100K Rinit=80K D=10N uv=10F p=1
************************************************ *
DIFFERENTIAL EQUATION MODELING *
************************************************
Gx 0 x value={ I(Emem)*uv*Ron/D^2*f(V(x),p)}
Cx x 0 1 IC={(Roff-Rinit)/(Roff-Ron)}
Raux x 0 1T
* RESISTIVE PORT OF THE MEMRISTOR *
*******************************
Emem plus aux value={-I(Emem)*V(x)*(Roff-Ron)}
Roff aux minus {Roff}
************************************************
*Flux computation*
************************************************
Eflux flux 0 value={SDT(V(plus,minus))}
************************************************
*Charge computation*
************************************************
Echarge charge 0 value={SDT(I(Emem))}
************************************************
* WINDOW FUNCTIONS *
* FOR NONLINEAR DRIFT MODELING *
************************************************
*window function, according to Joglekar
.func f(x,p)={1-(2*x-1)^(2*p)}
*proposed window function
.func f(x,i,p)={1-(x-stp(-i))^(2*p)}
.ENDS memristor
```

As can be shown in mentioned paper [7], described model above can not be used for frequencies higher than 5 Hz. This also, has been proved in paper [8], and that's why we have choose a new model which is based on Biolek model, but improved for higher frequencies.

B. A new improved model

This model is based on a Biolek SPICE model mentiond before and it represents an improvement in frequency domain. An improved electric circuit of the model [8] is shown in figure 3. After that, a SPICE model that describes circuit in Fig. 3 is given.

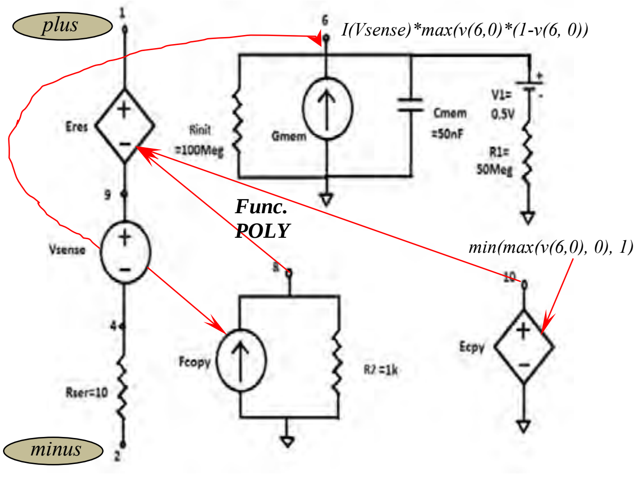


Fig. 3. New memristor model [8]

Here, POLY function that can directly calculate the value of integration is used instead of a regular integration function. A small value of R_{SER} is resistance used for current sensing between terminals 1 and 2. A clamping

circuit is used to initiate the operation of the memristor. Both, charge and flux depend on the state of the memristor, sensed at node six. The memristor terminals 1 and 2 are marked with red circles.
The used sub circuit PSPICE description is given next:

```
.SUBCKT newmem 1 2
****************************
**Squared quantity of voltage**
Eres 1 9 POLY(2) (8, 0) (10, 0) 0 0 0 0 1
****************************
Vsense 9 4 DC 0V
****************************
**Current sensing**
Fcopy 0 8 Vsense 1
R2 8 0 1k
Rser 2 4 10
**Differential equation**
Gmem 6 0 VALUE={I(Vsense)*max(v(6,0)*(1-v(6, 0)), +0)}
Cmem 6 0 50nF
****************************
**Limiting the window of 0 and 1**
Ecpy 10 0 VALUE={min(max(v(6,0), 0), 1)}
****************************
Rinit 6 0 100Meg
V1 6 7 DC 0.5V
R1 7 0 50Meg
.ENDS newmem
```

It can be noticed, that the new model doesn't have parametric structure. So, the values of D i.e., width of the thin film, and u_V – migration coefficient, etc., remain unknown. Figure 4 represents a well known I-V characteristic of memristor by HP [2]. The I-V characteristic of new model used in this paper is shown in Fig. 5.

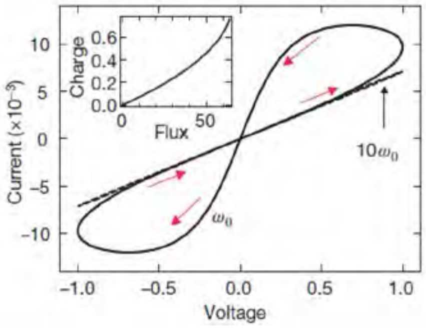


Fig. 4. Memrisor I-V characteristic [2]

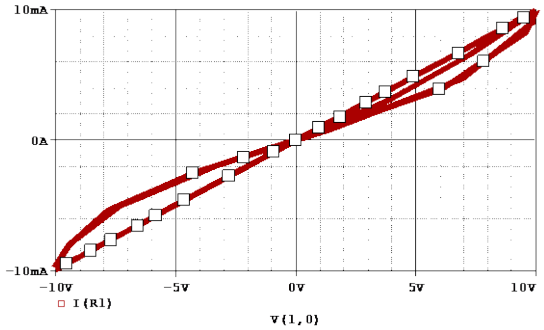


Fig. 5. A hysteresis curve on application of a sinusoidal signal 10V, 5kHz. The I-V characteristics of the memristor obtained by SPICE simulation

Further in the paper, authors will describe simulation results and memristors based realizationa of AND, OR and NOT logical functions.

III. SIMULATION RESULTS

Our simulations did not use MAGIC or IMPLY design. Instead, we have applied the input voltage [0 V, 5 V] while observing currents and voltages accros particular memristors. This was done only for AND and OR logical circuit. For NOT logic function, it was necesary to apply excitation voltage V_O, in order to obtain voltage accros input memristor and output memristor.

In oder to simulate AND, OR and NOT logical functions, memristors have to be connected in a certain way as given in figure 6.

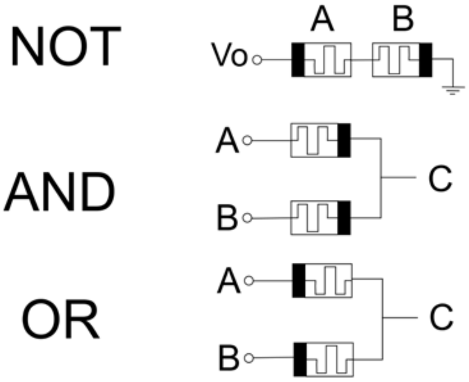


Fig. 6. Realised basic logic functions using memristors

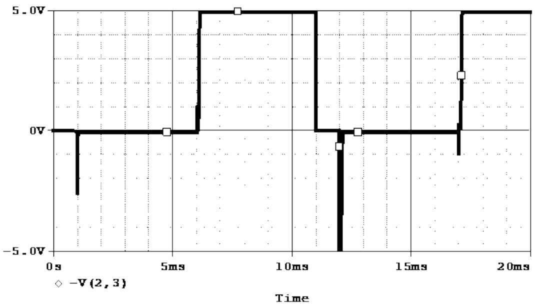


Fig. 7. Voltage accros memristor A of an inverter

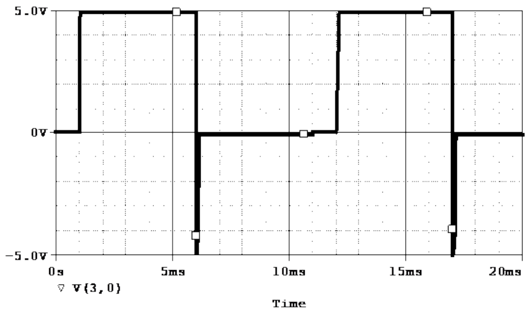


Fig. 8. Voltage accros memristor B for an inverter

For the simulation of an invertor, V_O voltage was applied. By applying V_O, resistances of memristor A and B was changing. According to that change, we can see that voltage accros memristor A represents inverted voltage of the memristor B. This confirms the assumptions that this circuit has properties of an invertor.

Signals of the memristors A and B are shown in Figures 7 and 8, respectively.

The AND and OR gate, there is no need for the V_O voltage. Simulations show that when logical "0" and "1" are applied at memristor A and B, memristor C gives voltage or current response that describes AND or OR logical function. Signals for AND gate are shown in figure 9, where red color relates to signal A and blue color relates to signal B. Green color represents the voltage of the output signal.

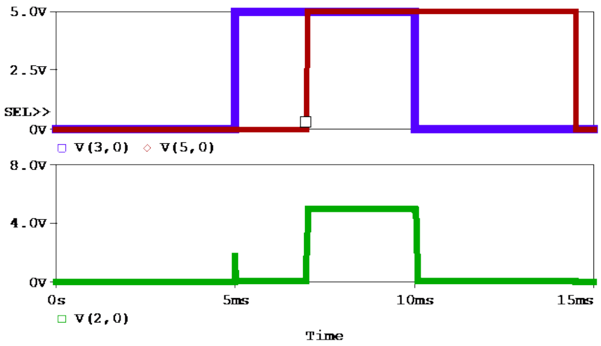


Fig. 9. Signals for AND gate

Signals for OR logic function are shown in figure 10, where red color coresponds to the voltage accros memristor A, while blue color corresponds to voltage accros memristor B. Green color represents the output signal voltage. As one can see from these waveforms, the output voltage, represents the OR logical function for A and B input combinations. The hazardous behavior of the voltage drop happens with transition of inputs high to low voltage level, when output drops to 0 V and after goes to 5V. This glitch represents an unwanted behavior and it can cause various readout errors.

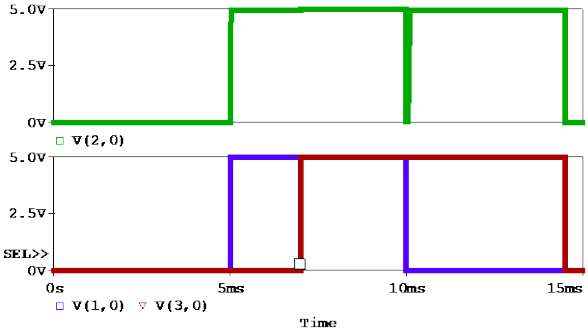


Fig. 10.Signals for OR gate

IV. CONCLUSION

The simulation of memristor logic model implemented in the software OrCad Pspice shows results relevant to the original logical AND, NOT and OR functons. The chosen way of simulation was the easiest to understand and to simulate.

Also, different SPICE memristor models are described. The first model is the most efficient, but it only works for very low frequencies which in digital electronics has no practical use. The second model, shows that with some changes in model design, it can work on higher frequiencies which is an andvatige, but with current technology it is still a simulation model.

Using memristors as logic gates would decrease power consumption and because of their nano scale structure thay could decrease or replace usage of traditional components such as CMOS transistors in logic gates. Further reaserch would be oriented into the elemenation of unwanted glitches. The increase of its speed should bring major advantage to memristor technology for its use in Crossbar arrays. The applications of memristors in RAM could make an important change in the memory storage systems realization.

ACKNOWLEDGEMENT

This research was patially funded by The Ministry of Education and Science of Republic of Serbia under contract no. TR32004.

REFERENCES

[1] L. Chua, “Memristor – The missing circuit element,” *IEEE Trans. Circuit Theory*, vol. 18, no. 5, pp. 507–519, Sep 1971.

[2] D. B. Strukov, G. S. Snider, D. R. Steward, and R. S. Williams, “The missing memristor found,” *Nature*, vol. 453, no. 7191, pp. 80–83, May 2008.

[3] Radwan, Ahmed G., Fouda, Mohammed E. – On the Mathematical Modeling of Memristor, Memcapacitor, and Meminductor DOI: 10.1007/978-3-319-17491-4

[4] A. Sinha, M. S. Kulkarni, and C. Teuscher, “Evolving nanoscale associative memories with memristors,” in Proc. of *11th IEEE Conference on Nanotechnology*, Aug 2011.

[5] S. Kvatinsky, G. Satat, N. Wald, E. G. Friedman, A.Kolodny, and U. C. Weiser, “Memristor-Based Material Implication (IMPLY) Logic: Design Principles and Methodologies,” *IEEE Trans. VLSI Syst.*, vol. 22, no. 10,pp. 2054–2066, Oct 2014.

[6] S. Kvatinsky, D. Belousov, S. Liman, G. Satat, N. Wald, E. Friedman,A. Kolodny, and U. Weiser, “MAGIC – Memristor-Aided Logic,” *IEEE Trans. Circuits Syst. II*, vol. 61, no. 11, pp. 895–899, Nov 2014.

[7] Z. Biolek, D. Biolek, and V. Biolkov´a. “SPICE model of memristor with nonlinear dopant drift.” *Radio engineering J.*, 18(2):211, 2009b.

[8] Ketaki Bhave, Dr. Nisha Sarwade “Memristor Modeling Using PSPICE” *Int. Journal of Engineering Research and Applications Vol. 4, July 2014, pp.139-143*

Unified Glare Rating as a measure of visual comfort

Dragan Vuckovic, Nenad Cvetkovic, Dejan Jovanovic and Miodrag Stojanovic

Abstract – Visual comfort is one of the essential parameters which is necessary to be considered during the designing process of indoor lighting. The International Committee for Illumination (CIE) adopted the Unified Glare Rating (UGR) formula in 1995, while the procedure for calculating and tabulating the UGR value of indoor lighting luminaires is adopted in 2010. Since then, most manufacturers have started to include data for UGR in their catalogues. Modern software packages for lighting design allow calculation of UGR and therefore provide the possibility to include UGR information in power engineering designs. From the other side, the measuring procedure of UGR, for realised installations, is very demanding, and consequently, it does not apply very often in practice. Because of that, the UGR information for existing installation is usually unknown. The UGR value depends of the light intensity and luminance distributions of luminaires, the luminaire orientation, as well as the level of eye adaptation and the luminaires arrangement (i.e. the angles between viewing direction and direction from the viewer and luminaires optical centres). Analyse and measurements of UGR in the case of the realised indoor lightning installation, show that it is possible to achieve satisfactory visual comfort evaluated by the UGR index using luminaires with worse UGR. It is also possible that using the luminaires with the satisfactory UGR does not always result with required visual comfort conditions. Based on that, one can conclude that it is not enough to choose a luminaire having a satisfied UGR. Still, it is necessary to consider other parameters that affect visual comfort and make the simulation of complete room scenario. That results with the need for establishing close cooperation of all professions during the design process.

Keywords - Unified Glare Rating (UGR), visual comfort, indoor lighting.

I. INTRODUCTION

There are many studies realized in the mid-twentieth century, in order to propose expression that includes glare influence from the luminaires located in the observer field of view. In 1963 the Illuminating Engineering Society defined the formula for the Visual Comfort Probability (VCP) [1]. Numerous experiments were realized in real workplaces, with observers who gave their subjective and personal opinion about visual comfort. The reflectance coefficients of the ceiling, walls and floor during experiment setup were 0.7, 0.5, 0.2 and 0.3 respectively. Various scenes and lighting scenarios were realized. Based on the obtained results, two diagrams corresponding to the two luminaire types are defined. In that experiment, two groups of luminaires have been used. Group A includes luminaires without luminous lateral side, while group B includes luminaires with luminous lateral sides higher than 30 mm. Eight luminance limiting curves have been defined for visual comfort estimation. Designers have used luminance limiting curves method for a long time, but the development of information technologies provided the possibility for proposing an expression that can quantify the level of visual comfort. The International Commission on Illumination (CIE) defined the formula for calculating UGR in 1995 [2]. According to the realised researches, obtained results and experience, the expression is proposed as following:

$$\mathrm{UGR} = 8\log\left[\frac{0.25}{L_b}\sum\frac{L^2\omega}{p^2}\right] \tag{1}$$

where is:

L_b - the background luminance

L - the luminance of the luminous part of each luminaire in the observer field of view

ω - the solid angle of the luminous part of each luminaire in the field of view of the observer

p - position index of the luminaire.

II. UGR AS CATALOGUE DATA

Procedure for calculating and tabulating the UGR value of indoor lighting luminaires is adopted by CIE in 2010 [3]. UGR data for luminaires are given for different room dimension ratios and different reflection coefficients of interior surfaces. The standard EN 12464 [4] defines the range of UGR values between 10 and 30. The typical UGR values are defined: 13, 16, 19, 22, 25 and 28. Also, the Standard defines minimal UGR values for certain visual tasks. E.g. for offices and classrooms, the maximal UGR value should be 19. Luminaires are usually declared for certain UGR value that corresponds to standard room dimensions. Standard room dimensions for UGR luminaires labelling are X=4H and Y=8H (H is room high). Besides that, reflection coefficients of ceilings, walls and the floors are 0.7, 0.5 and 0.2. Fig. 1 shows the correlation between the luminance limiting curves and UGR values. One can see that value UGR=19 corresponds to the luminance limiting curve for illuminance level of 500lx - Quality class 1.

Dragan Vuckovic, Nenad Cvetkovic, Dejan Jovanovic and Miodrag Stojanovic are with the University of Nis, Faculty of Electronic Engineering.

Address, E-mail: dragan.vuckovic@elfak.ni.ac.rs, neand.cvetkovic@elfak.ni.ac.rs, dejan.jovanovic@elfak.ni.ac.rs and miodrag.stojanovic@elfak.ni.ac.rs.

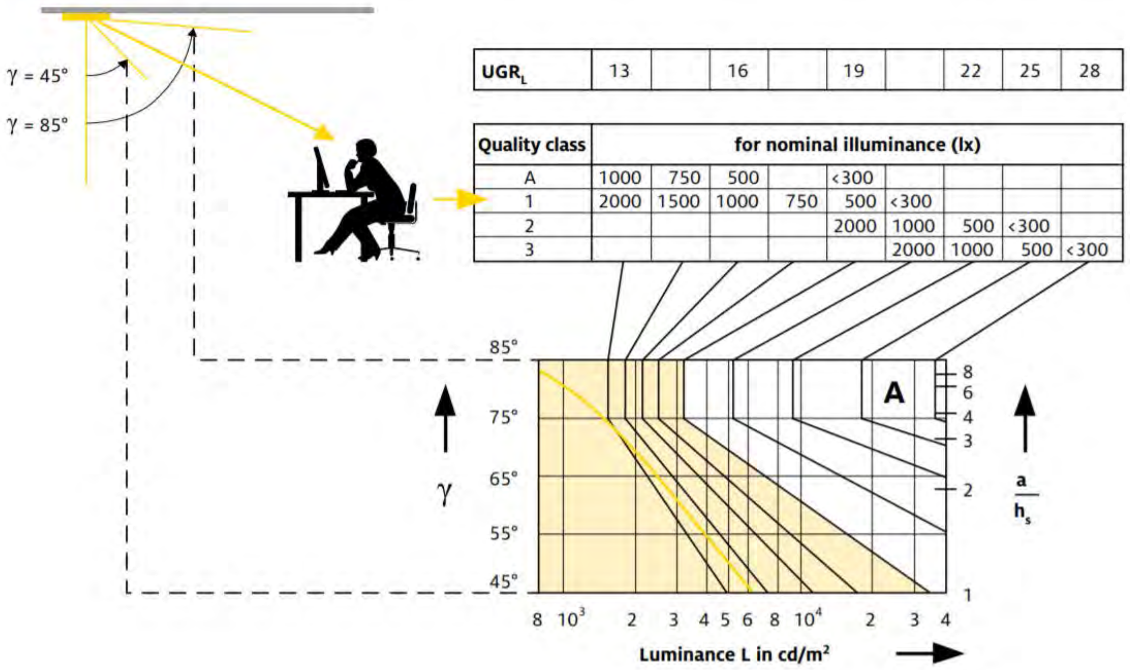


Fig. 1. Correlation between the luminance limiting curves and UGR values

III. UGR –APPLICATION IN PRACTICE

There are cases where calculations with luminaires having a declared value of UGR = 19, give UGR values slightly higher than 19. These results are obtained because of the large curtain-less glass surfaces. The glass influences on a slightly lower value of average brightness of the environment, so that the brightness of the luminaires significantly affects the psychological glare. Opposite to these cases, it happens that the reflection coefficients of the room surfaces, as well as the dimensions of the room, result with the highest value of UGR below 19, although installed luminaires for a standard room, has declared UGR = 22. So, it is obvious that during the designing process of the indoor lighting installations, it is necessary to define the reflection characteristics of the material as accurate as possible and then make a decision of luminaire type.

The LEDs luminaires usually have opal diffusers which have Lambertian radiation characteristics, so they consequently increase psychological glare. There are cases where the situation in terms of visual comfort is significantly deteriorated with the installation of opal diffuser LEDs luminaires. Fortunately, technology is in progress, and there are price-competitive materials for diffusers that provide better light distribution with UGR less than 19. Some manufacturers produce luminaires in which LED modules are retracted related to the surface of the luminaire. In this way, such luminaires are, for the standard room types, declared with UGR less than 16 (even less than 13). Unfortunately, such luminaires are still quite expensive compared to LED "panels" and because of that their massive application did not start.

Measuring of UGR on realized installations is a rather demanding procedure and requires appropriate equipment as well as time to perform all necessary measurements to obtain the results. UGR measurement is usually realized for the most critical observer position. It is interesting that even, the dark school board will affect the deterioration of the UGR value. Concerning necessity to satisfy the vertical illumination value of 500 lx for the board surface, the lamp for the illumination of school boards help to reduce the influence of its dark surface. Using whiteboards, the problem of the impact of the classroom board on visual comfort is reduced, but it is important to avoid physiological glare since their surfaces reflect light.

IV. NUMERICAL SIMULATION AND MEASURING RESULT

In this chapter, numerical results for UGR are given for classroom illuminated by luminaires for general lighting and luminaire for school board. Typical classroom luminaires and furniture setup is given in Fig. 2 and Fig. 3 respectively. The points for UGR calculations are given in Fig. 4, while the calculated results of the UGR are shown in Fig. 5. Dialux evo software package [5] has been used for the numerical analysis. LED luminaires that are used for classroom lighting have declared UGR of 22, for a standard room. Considering complete setup maximal obtained value of UGR is 19, and it corresponds to the calculating point placed on the back corner.

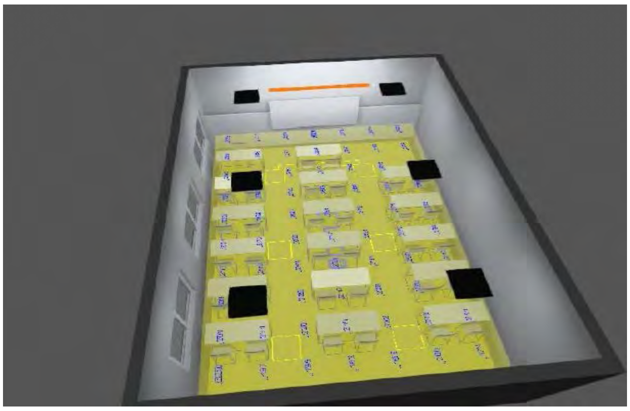


Fig. 2. Overview of classroom luminaires setup

Fig. 3. Overview of classroom furniture setup

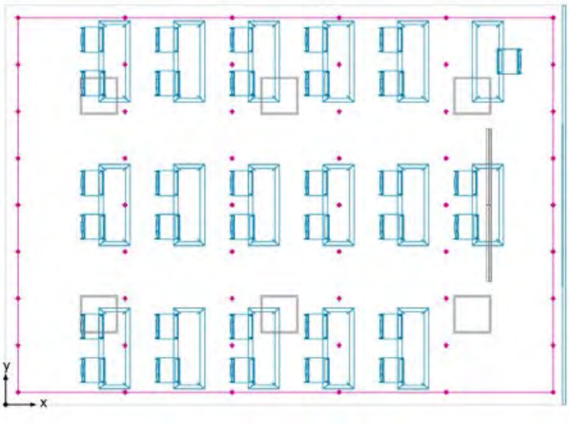


Fig. 4. UGR Calculation points

An overview of surfaces that have impact on UGR value, and measurement equipment are given if Fig. 6 and Fig. 7 respectively. According to the measured results of realized lighting installations, it can be calculated that UGR is in good agreement with one obtained by numerical simulation.

As previously mentioned, software for lighting calculations represent useful tools and helps a lot in the process of estimation of visual comfort.

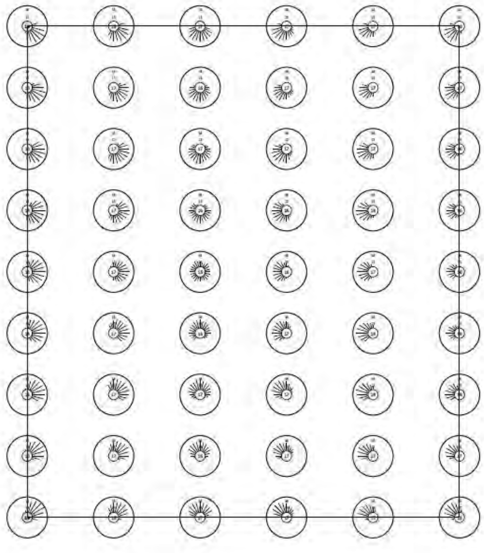


Fig. 5. UGR calculated values

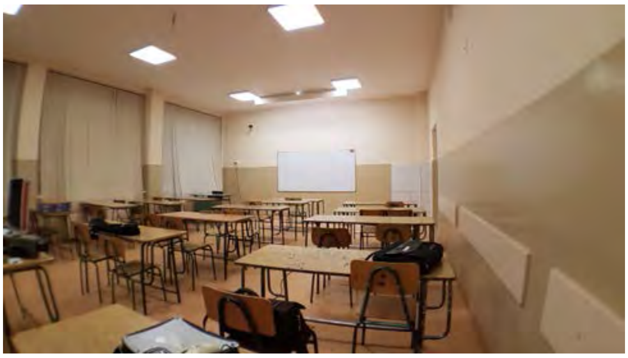


Fig. 6. View from the back corner - UGR measured point

Fig. 7. Measurement equipment for UGR evaluation

IV. CONCLUSION

This paper presents UGR as a measure of visual comfort. It is an essential datum that should be considered during the designing process of indoor lighting. There are a lot of studies dealing with discomfort glare due to the fact than the introduction of LED light sources with opal diffusor have significantly decreased visual comfort in the interior lighting. According to the obtained results, it can be concluded that it is not enough to choose a luminaire labelled with satisfied UGR, but it is necessary to make a proper simulation for complete room scenario. So, it is evident that during the designing process of the indoor lighting installations, it is necessary to define the reflection characteristics of the material as accurate as possible and then make a decision of luminaire type.

ACKNOWLEDGEMENT

This work is the result of research on project III44004 funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] IESNA, Ed.: Rea M.S.: Lighting Handbook - Reference & Application, IESNA 8th Edition, New York, 1993, Reprinted 1995.
[2] CIE Publication: “Discomfort Glare in Interior Lighting”, CIE Technical Report, CIE Publ. No. 117, 1995.
[3] CIE Publication: “Calculation and Presentation of Unified Glare Rating Tables for Indoor Lighting Luminaires”, CIE Technical Report, CIE Publ. No. 190, 2010.
[4] European Standard EN 12464-1, 2011, "Light and lighting - Lighting of work places - Part 1: Indoor work places"
[5] DIALux evo 8.2, available at: https://www.dial.de/en/home/

Adaptive Controller Based on LMS Algorithm for Grid-Connected and Islanding Inverters

Tatjana Nikolić, Goran Nikolić, and Branislav Petrović

Abstract - In the field of renewable energy, DC/AC inverters are crucial building blocks used in the various industrial and consumer devices. Grid connected power inverters should operate in synchronism with the grid voltage in order to get maximum active power. In this paper, the structure of a power system based on adaptive control is described. The main purpose of the adaptive controller is to continuously adapt the output signal of the inverter to the corresponding load and/or grid waveform. It is very important to enable the generation of appropriate excitation signal for the inverter. By involving adaptive controller the response time decreases and quality of power delivery to the load or grid increases. Matlab/Simulink model of the power system with adaptive control is used for simulation.

Keywords - DC/AC inverter, Pulse Width Modulation (PWM), Adaptive filter, System identification.

I. INTRODUCTION

Recently, there has been a pronounced increase of interest in the field of renewable energy. To meet the increased demand for electrical energy delivery alternative renewable energy sources with energy storage device (i.e. battery) are used. In an effort to utilize easily available solar energy effectively, the need for static inverters has increased significantly. With high energy efficiency and fast time response, inverters have possibility to convert DC energy stored in batteries to conventional controlled AC form. The applications require DC-AC power inverters with a high performance voltage regulation and high efficiency. To achieve these goals, power inverters need an optimal controller. The converted AC can be at any required voltage and frequency by using switching technique implemented by corresponding control circuits [1, 2]. With the rapid development of large-scale digital integrated circuits, power inverters are required to have a smaller area and a faster dynamic response [3].

Power inverters can be used as grid tie or stand-alone power sources. When the inverter is connected to the grid then the grid controls the amplitude and frequency of the inverter output and the inverter operates in current control mode. If the network is not available, the inverter will autonomously supply the load with adequate AC voltage in respect to amplitude and frequency. In this case, the inverter will control the voltage [4]. With aim to provide a correct operation of the power inverter connected to grid or/and load, with high stability and power quality, pulse width modulation (PWM) generator which drives the inverter is needed (see Fig. 1). The inverter will continuously generate the desired output signal that is consistent with the network waveform if it is excited by the appropriate input signal. Recently, research efforts have been devoted to system identification and adaptive and self-tuning control techniques for power converter applications [3, 5, 6]. The implementation of considered methods require significant signal processing, and higher computational complexity increases execution time and disables real-time operation. For this reason, there is still a need for a low complexity, online adaptive technique which can run continuously in closed loop, adapting to changes in the power system. In order to achieve this goal, an efficient design of an adaptive controller block is proposed in this paper (see Fig. 2). By inserting adaptive filtering between the inputs and outputs of the inverters, the transfer function of the inverter is first detected. Then, based on the known transfer function of the inverter and the desired output, using inverse filtering, an excitation PWM signal is generated. By introducing a feedback control loop, PWM waveform effectively tracks and compensates load variation and/or grid power disturbance. In this way, the output signal of the inverter is continuously adapted to the desired load and/or grid waveform.

The rest of the paper is organized as follows. Section II presents global structures of single-phase grid-connected conventional power system and power system with adaptive control. Detailed description of the proposed adaptive controller with two different processes, adaptive and inverse filtering, is given in Section III. Simulation results which relate to the generation of an input excitation signal based on the desired output signal are given in Section IV. Section V deals with conclusions.

II. POWER SYSTEM

General structure of the power system based on DC/AC inverter is presented in Fig. 1. It consists of a DC source, an inverter, PWM generator, filter block and a grid utility (or load). The inverter block acts as an interface circuit which connects the DC source to utility grid and/or local load. Primary function of the inverter is to supply active power to the load and grid in accordance with grid connection standards. Grid standards involve numerous requirements, but some of the more representative are the

Tatjana Nikolić, Goran Nikolić and Branislav Petrović are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: [tatjana.nikolic; goran.nikolic; branislav.petrovic] @elfak.ni.ac.rs

following: reactive power compensation, minimization of harmonic distortion produced by nonlinear load, and correct phase synchronism between the current and voltage with aim to achieve unit power factor [1]. The output voltage and current quality at the inverter output can be improved by using filter block. The most common approach is the use of LC or LCL filter.

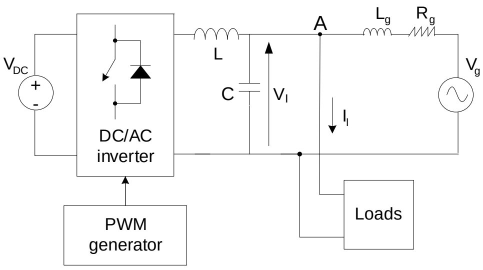


Fig. 1. Power system

PWM generator is implemented in the inverter with order to generate AC waveform of variable voltage and variable frequency. It generates pulses variable duration or duty cycle ratio and, in this way, controls the voltage/current/power at the output of the inverter. The main design challenge now is to obtain an optimal PWM signal for driving the inverter, which should ensure efficient operation and accomplishment of energy quality requirements [4]. The precision of DC/AC conversion depends of applied modulation method and inverter topology [7]. To fulfill the aforementioned requirements we propose adaptive control of the PWM generator block. Architecture of modified conventional power system with adaptive control is presented in Fig. 2. The main goal of this system is that the inverter continuously generates a signal that is consistent with the grid waveform [8].

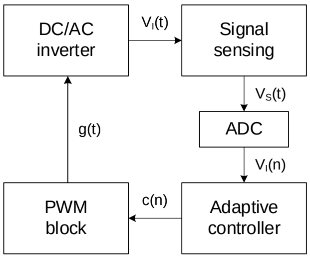


Fig. 2. Power system with adaptive control

As can be seen from Fig. 2, power system with adaptive control has two parts: an analog and a digital part. The digital control, in comparison with the analog control, improves the performance of the DC/AC inverters due to the following reasons: has more flexibility in the designs, the size design is smaller, the algorithms can be easily reprogrammed and improved, it is suitable for the adaptive system [8]. Fig. 2 shows a power system model, which consists of an A/D converter, an adaptive controller, a digital PWM block and a DC/AC inverter block. The A/D converter performs conversion of sensed analog voltage values to digital forms. The PWM block generates the duty ratio signal to control the inverter switches. Proposed adaptive controller is relied upon system identification to identify the parameters of the power inverter, operating under the assumption of certainty equivalence, whereby the estimated inverter parameters are treated as the true values for the purposes of controller design.

III. ADAPTIVE CONTROLLER

Adaptive controller from Fig. 2 consists of two blocks, adaptive filter and inverse filter. These blocks perform two different processes, adaptive and inverse filtering, following each other. First, adaptive filter models the relationship between the excitation and the output signal of the inverter in real time in an iterative manner. Knowledge of the transfer function of the power inverter is required by the controller so that a suitable control signal can be calculated and applied. Based on the transfer function of the inverter and the desired output signal, an appropriate excitation signal can be generated using inverse filtering.

A. Adaptive filtering process

Digital FIR filters are widely employed in practical real-time digital signal processing applications [9, 10]. Whereas any fixed filter is designed in advance with knowledge of the statistics of signals, the adaptive filter continuously adjusts to a changing environment through the use of adaptive algorithms that are needed in order to continuously update the filter coefficients. The adjustable parameters are dependent upon the applications. This paper will focus on applying adaptive filtering in power system identification by using least mean square (LMS) algorithm. An adaptive filter is used to provide a linear model that represents the best fit to an unknown system. In our case, the inverter block presents the unknown system. The system identification procedure has a three-step logic flow: (i) collect input–output data, (ii) choose a model set and (iii) pick the best parameters fit to this model. If the system is dynamic in nature, the model will be time varying.

Fig. 3 shows a MATLAB/Simulink model of adaptive filtering process. A resistor was introduced to simulate the dynamic characteristics of the inverter by changing the load. In order to identify the transfer function of the inverter block, the same input is applied to the inverter and to the adaptive filter. A sample from a digital input signal $x[n]$ is fed into the adaptive filter, that computes a corresponding output signal sample $y[n]$ at time n. This output signal $y[n]$ is computed based on adjustable parameters of the adaptive filter. The output signal is compared to the inverter response signal, $d[n]$, by subtracting the two samples at time n. This difference signal is the error signal, given by $e[n] = d[n] - y[n]$.

The LMS algorithm is used to emulate a desired filter by finding the filter coefficients that relate to producing the least mean squares of the error signal. This algorithm for every input sample first computes the output $y[n]$ using the current set of coefficients $\boldsymbol{h}_n[k]$, then computes the error between the desired response and the filter output, and then updates all the coefficients using the following equation:

$$\boldsymbol{h}_{n+1}[k] = \boldsymbol{h}_n[k] - \mu e[n]x[n-k] \qquad (1)$$

The factor μ determines the convergence of the algorithm. Large values of μ result in fast convergence [9].

If perfectly identified, the output $y[n]$ of the adaptive filter should be the same as the output $d[n]$ of the unknown system. However, under realistic environmental conditions (component tolerance, unpredictable load changes, ambient temperature change, ageing effect), the model parameters change so that the obtained coefficient values are not exact but approximate. Our goal is to determine the mathematical form of the transfer function that is closest to the real function of the inverter block. The adaptive filter coefficients, i.e. the transfer function of the inverter block, represent the outcome of the adaptive filtering process.

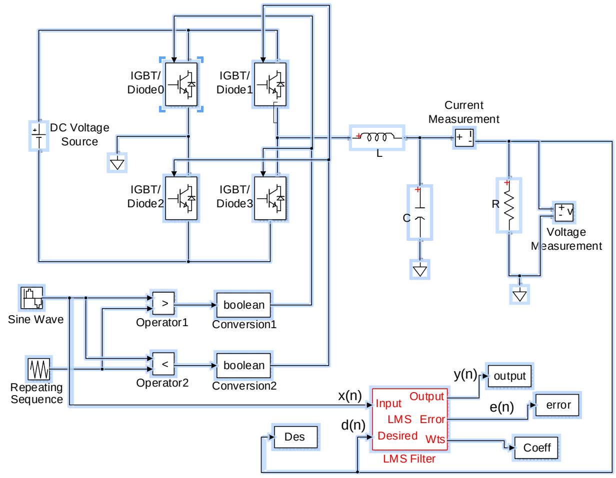


Fig. 3. MATLAB/Simulink model of a power inverter identification by using adaptive LMS filter

B. Inverse filtering process

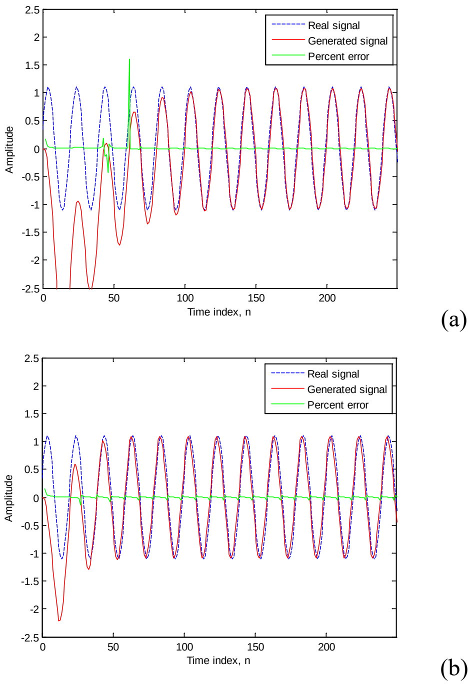
The transfer function of the inverter block is the connection of the output and the input voltage signal and it can be defined as $H(z) = Y(z)/X(z)$. By constantly monitoring the grid voltage, the output signal of the inverter, which should be in accordance with the grid voltage, will be known at all times. Therefore, if the transfer function and the inverter output voltage signal are known, then the input excitation signal can be determined by applying the inverse filtering process.

The pole–zero diagram provides important insight into the properties of systems. To find the poles and zeros of the rational function $H(z)$, we can use MATLAB tool. In the same way, we can find poles and zeros of the inverse function, $H(z)^{-1}$. The system is stable if all poles are inside the unit circle. Therefore, if we have an unstable filter, it is necessary that we first design a stable filter. If any pole outside the unit circle that causes instability is reflected back inside the unit circuit, then the magnitude of the frequency response will remain unchanged. For example, a pole p with $|p| > 1$ is replaced by the new pole $\tilde{p} = 1/p^*$, where $*$ denotes complex conjugation. In addition, a scaling factor is required. The correct scaling is obtained if each factor $(1 - p_i z^{-1})$ with $|p| > 1$ of the inverse function is replaced by $|p|(1 - z^{-1}/p_i^*)$. But the disadvantage of this method is that it will change the phase response. The filtered output signal is delayed with respect to the input signal. The delay caused by the filter equals half the filter order. When the shift is constant, we correct for the delay by shifting the output signal in time. Sometimes the filter delays some frequency components more than others so that a phase distortion occurs. To compensate for this effect, we can perform zero-phase filtering using appropriate function.

IV. SIMULATION RESULTS

With the aim to demonstrate the validity of the proposed adaptive control, simulation of the power systems is performed using MATLAB/Simulink environment. The performance measures discussed in these simulations are: filter length, convergence rate, error, and stability. The adaptive power system identification is primarily responsible for determining a discrete estimation of the transfer function for an unknown analog system. After a number of iterations of this process are performed, and if the inverter is designed correctly, the adaptive filter's transfer function will converge to, or near to, the unknown inverter block's transfer function. The coeficients of the inverter block are estimated with minimum error. For this configuration, the error signal does not have to go to zero, although convergence to zero is the ideal situation, to closely approximate the given system. There will, however, be a difference between adaptive filter transfer function and the unknown system transfer function if the error is nonzero and the magnitude of that difference will be directly related to the magnitude of the error signal.

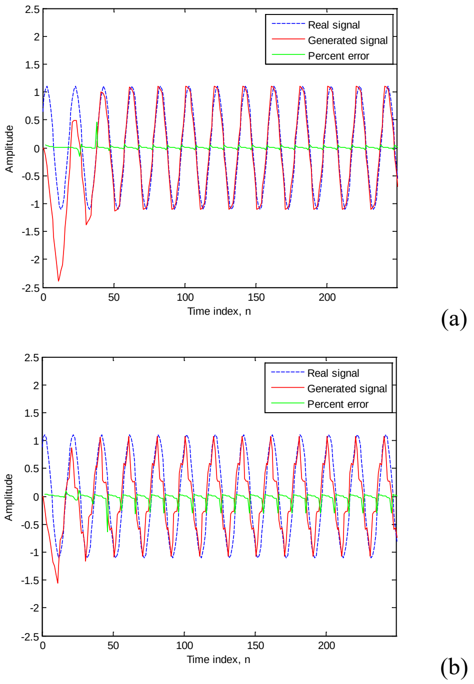
The results of simulation presented in Fig. 4 and 5 show agreement between the real (exact) excitation signal and the new signal obtained by the proposed adaptive method. In addition, in Fig. 4 and 5, the percent error of these two signals is presented. The percent error is chosen instead of an absolute error due to the normalized signal values. Fig. 4 shows the real and the generated signal for different parameters, filter order (N) and convergence factor (μ) where the parameters are: (a) $N = 2$, μ =0.1; and (b) $N = 2$, μ =0.01. As can be seen, faster dynamic response is achieved for a lower value μ. Fig. 5 shows the real and the generated signal for different parameters, where: (a) $N = 4$, μ =0.01; and (b) $N = 6$, μ =0.01. Note that the best design choice of producing inverter excitation signal is achieved by using a lower order filter.



(a)

(b)

Fig. 4. Real and new generated signal for (a) N = 2, μ=0.1, and (b) N = 2, μ=0.01



(a)

(b)

Fig. 5. Real and new generated signal for (a) N = 4, μ=0.01, and (b) N = 6, μ=0.01.

V. CONCLUSION

The adaptive controller of the power system based on LMS algorithm is proposed in this paper. Power inverter identification by using adaptive LMS filter and inverse filtering process are described. The excitation signal of the inverter is generated by the adaptive controller. It is shown that the output signal of the inverter is continuously adapted to the desired load and/or grid signal.

The proposed method can be implemented in other applications where it is necessary to generate an excitation signal based on a known output signal using a system identification method.

ACKNOWLEDGEMENT

This work was supported by the Serbian Ministry of Science and Technological Development, Project No. TR-32009 – "Low-Power Reconfigurable Fault-Tolerant Platforms".

REFERENCES

[1] Daniel W. Hart, "Power Electronics", McGraw Hill, Inc. New York, 2011.

[2] Arturo Y. Jaen-Cuellar, Rene de J. Romero-Troncoso, Luis Morales-Velazquez and Roque A. Osornio-Rios, "PID-Controller Tuning Optimization with Genetic Algorithms in Servo Systems", *International Journal of Advanced Robotic Systems*, 2013, Vol. 10(9):324

[3] Ling Lv, Changyuan Chang, Zhiqi Zhou, "An FPGA-Based Modified Adaptive PID Controller for DC/DC Buck Converters", *Journal of Power Electronics*, March 2015, Vol. 15, No. 2, pp. 346-355

[4] Haitham Abu Rub, Atif Iqbal, Jaroslaw Guzinski, "High performance control of AC drives with MATLAB/Simulink models", John Wiley & Sons Ltd, UK, 2012.

[5] M. Algreer, M. Armstrong, and D. Giaouris, "Adaptive PD+I control of a switch mode DC-DC power converter using a recursive FIR predictor" *IEEE Trans. Ind. Appl.*, Sep. 2011, Vol. 47, No. 5, pp. 2135-2144

[6] Amin Hajizadeh, Amir H. Shahirinia, Navid Namjoo, David C. Yu, "Self-tuning indirect adaptive control of non-inverting buck-boost converter", *IET Power Electronics*, 2015, Vol. 8, No. 11, pp. 2299-2306

[7] M. Valan Rajkumar, P.S. Manoharan, A. Ravi, "Simulation and an experimental investigation of SVPWM technique on a multilevel voltage source inverter for photovoltaic systems", *Electrical Power and Energy Systems*, 2013, Vol. 52, pp. 116-131

[8] H. Peng, A. Prodic, E. Alarcon, and D. Maksimovic, "Modeling of Quantization Effects in Digitally Controlled DC-DC Converters" *IEEE Tran. Power Electronics*, 2007, Vol. 22, pp. 208-215

[9] Shoab Ahmed Khan, "Digital Design of Signal Processing Systems, A Practical Approach", John Wiley & Sons, Ltd., 2011.

[10] Mohanty B. K., Meher P. K., "A High-Performance FIR Filter Architecture for Fixed and Reconfigurable Applications", *IEEE Trans. VLSI Systems*, 2016, Vol. 24, No. 2, pp. 444-452

Novel Method in Robust Radio Communication Emission Classification

Jovan Bajčetić and Davorin Mikluc

Abstract - This paper presents a new concept of deep learning technique usage for frequency spectrum image classification. The presented research gives the overview of the image database creation, acquired data processing and emission type recognition model developing. Image database is created using real time generated, differently modulated robust radio emissions using 25 kHz wide radio channels within 15 to 55 MHz frequency band. Those radio emissions are then represented using 300-time sample power distribution over given frequency band. Every each of those representations is defined with 500 data points. Using around 100 of spectrogram and polar images of the acquired data, we developed a deep learning model capable to perform multiclass classification for detected robust radio emission.

Keywords – Image classification, Deep learning, Radio emission, Robust radio communication.

I. INTRODUCTION

Deep learning (DL) procedures has enhanced image processing technology to the extent to be the inevitable technology of the 4th industrial revolution [1], [2]. It is the main driver of the autonomous driving technologies and is becoming more and more affordable for commercial use [3]. In data science, neural networks are in the core of DL techniques and can be considered as computing systems based on biological neural networks. Basic idea of artificial neural network is to simulate a large amount of densely packed, interconnected nerve cells within the computer, so that it can provide learning of terms, pattern recognizing or decisions making. Using this feature, today is achievable what was not possible before, thanks to the immensely improved computing capabilities of GPU based systems.

Tactical radio communications provide users with reliable narrow band communication links. In a harsh electronic warfare environment, it is essential to be undetectable, efficient in terms of information exchange while mobile, operating in the mash organized radio networks. Robust radio communication systems working in VHF (Very High Frequency – 30-300 MHz) and HF (High Frequency – 3-30 MHz) frequency bands yield for the possibility of real-time radio emission classification. Since modern tactical radio communications use various kinds of media access technologies that enable Low Probability of Interception (LPI) and power consumption [4], we initiated the research in order to develop the software that could automatically identify the received radio signal modulation technique and media access technology. That could furtherly allow intelligent spectrum utilisation in order to enable coexistence of multiple networks within the same area.

One of the conceivable DL use cases is to provide efficient, real-time radio emission classification out of the crumbled radio spectrum. That motivation drove us to conduct the research in order to generate the tool for such challengeable task.

The research was conducted in the following order:

1. Reference image database creation;
2. Image classification model defining;
3. Four class multiclass model training;
4. Model testing.

II. RADIO SPECTRUM EMISSION CLASSIFICATION

A. Reference image database creation

DL techniques require a huge amount of data in order to be properly trained. The research preparation was directed towards a suitable image database creation of different modulation emission recordings. The first step in software development was image acquiring into the database which was used as a source of data for DL processing. Images were the representations of radio emission spectrum within defined frequency band and the actual image creation was done using the Matlab software environment that provides connection to a wide variety of data acquisition hardware solutions.

The receiving signal spectrum acquiring was performed using Tektronix RSA306B USB Real Time Spectrum Analyzer which provides real time spectrum analysis, streaming capture, signal analysis and +20 dBm to -160 dBm measurement range for signals within 9 kHz to 6.2 GHz frequency band. The transmitter of this communication system was the genuine radio transceiver that generate defined modulated radio signal. Emission was produced using Thales HF and VHF transceivers as suitable transmitters since they offer possibility of different modulation modes which are between 15 MHz and 30 MHz and 30 MHz and 88 MHz, respectively. Those frequency bands are intensively used in the HF and VHF tactical military communication systems.

Modulations that were used in this research were Amplitude Modulation (AM), Frequency Modulation (FM), Frequency Shift Keying (FSK) and Fast Frequency Hopping (FFH). The communication system used was formed as presented in Fig. 1. Signal distortion due to interference and

Jovan Bajčetić and Davorin Mikluc, University of Defence in Belgrade, Military Academy, Pavla Jurišića Šturma 33, Belgrade, E-mail: bajce05@gmail.com, miklucd@yahoo.com.

noise was avoided using a proper attenuator in order to overcome the receiver overloading.

Fig. 1. Experiment setup – data acquisition

The spectrogram made from the sampled received signal was generated by VHF device between 30 MHz and 55 MHz for FFH, FSK and FM modulation techniques. For that frequency band, measurements were done for 100 different modulated carriers with 250 kHz step between them (100 recordings). For each frequency, the 15 MHz wide frequency band was sampled in 300 points. Each recording was stored using Matlab, in the structure that contains 3 matrixes. Those matrixes store time results in 500 samples of each recording for 15 MHz wide frequency band during the time of approximately of 2,5 seconds. The first matrix stores time of the recording, the second matrix is consisted of acquired IQ data for m=300, and the third matrix is consisted of produced signal spectrum generated with FFT (Fast Fourier Transformation).

AM modulation measurements were done with HF device as the trasmitter on 100 modulation frequencies between 15 and 30 MHz with 150 kHz step. Results were structured and stored in the same way as with VHF device. The analyzed spectrum bandwidth was 40 MHz broad with sampling points 80 kHz distant from each other. The acquired time moment spectrums ($SIG_{0 \dots m}$) can be represented with I and Q values for each frequency/time sample:

$$\begin{aligned} SIG_0 &= I_{0,1} + Q_{0,1}i + I_{0,2} + Q_{0,2}i + \ldots + I_{0,500} + Q_{0,500}i \\ SIG_1 &= I_{1,1} + Q_{1,1}i + I_{1,2} + Q_{1,2}i + \ldots + I_{1,500} + Q_{1,500}i \\ &\ldots \\ SIG_m &= I_{m,1} + Q_{m,1}i + I_{m,2} + Q_{m,2}i + \ldots + I_{m,500} + Q_{m,500}i \end{aligned} \tag{1}$$

The sum of spectrum data *SIG* that can be used for spectrogram presentation is defined in 500 frequency points (*n*) and 300 time recordings (*m*) with the acquired I and Q values:

$$SIG = \sum_{m=1}^{300} \sum_{n=1}^{500} I_{\mathrm{m},n} + Q_{m,n} i \tag{2}$$

Finally, $RFSIG_m$ is the representation of the emitted radio signal for each recorded moment of the radio emission:

$$RFSIG_m = \sum_{n=1}^{500} I_{m,n} \cos(2\pi f_n t) + Q_{m,n} \sin(2\pi f_n t) \tag{3}$$

The samples of produced spectrograms for each modulation technique are presented in Fig. 2.

B. Image classification model defining

Considering the database of the produced recordings coming from the experiment, the next stage was to produce adequate images which would be optimal for the multiclass image classification using DL implemented in a chosen neural networks API (Application Programming Interface). DL model used in the experiment is realized using nowadays one of the most popular APIs – Keras written in Python and capable of running on top of TensorFlow [5].

Keras is a very popular DL API with several advantages. First of all, it allows easy and fast prototyping because of user friendly code, modularity and extensibility (we used 3 classes multiclass classification model which we extended to 4 classes), Second, it supports both convolutional and recurrent networks, as well as combination of those two (we used convolutional). Finally, it can be run on both CPU and GPU (we used CPU). Four-class classification was implemented using model with the parameters shown in Table 1.

TABLE I
THE MOST IMPORTANT MODEL PARAMETERS

Type of DL model	Convolution
Number of layers	2
Number of classes	4
Number of training epochs	20
Number of the training set images	100
Number of validation images	6

Images were uniformly dispersed in four classes for the training process – 100 images representing spectrograms of the acquired data. In order to find more optimal image representation of the spectrogram for more efficient modulation recognition, we made two different types of the spectrogram depiction. The first one is classical 2D spectrogram representation of power distribution over time and frequency, and the second one is polar plotted representation of the same data.

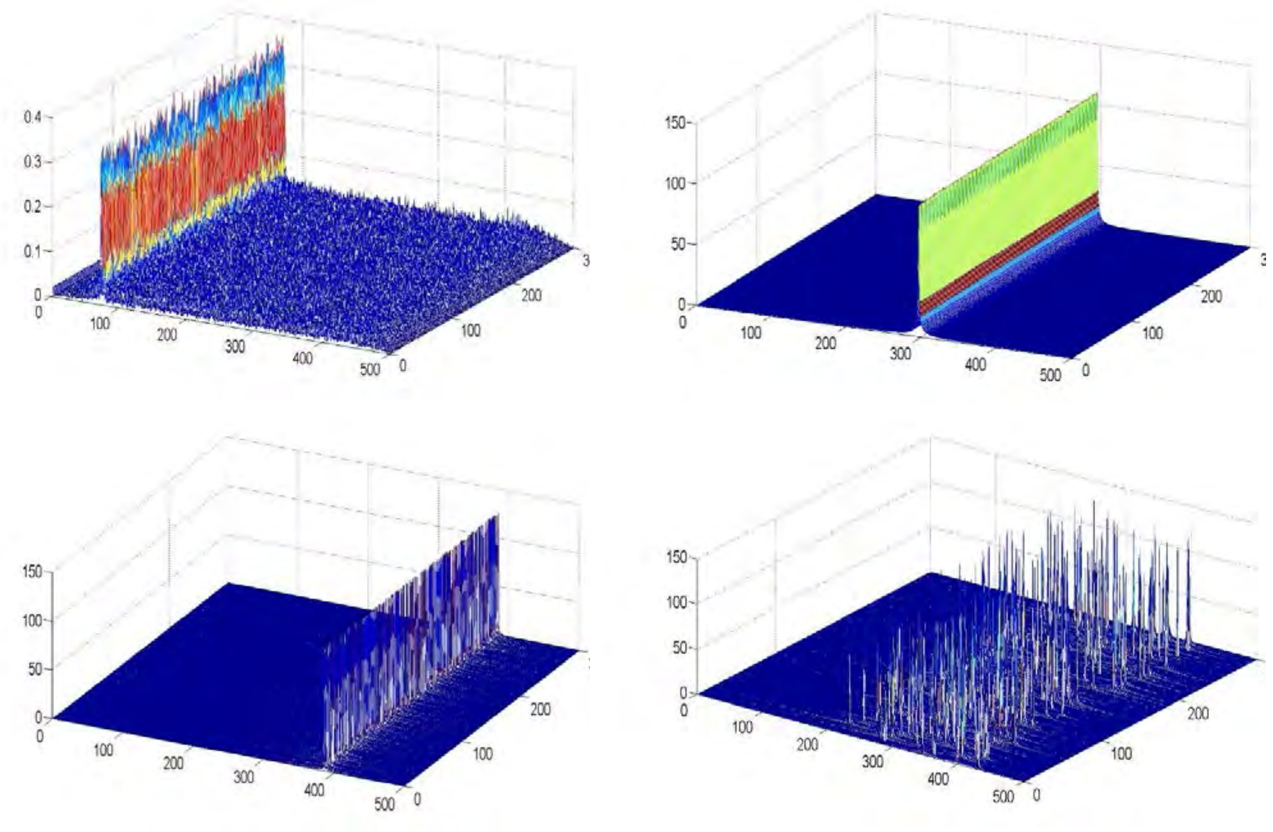


Fig. 2. Spectrogram images – AM (up left), FM (up right), FSK (down left), FH (down right)

C. Four class multiclass model training

One of the main problems in DL mechanism for image classification is model overfitting which occurs when there is not enough number of learn pattern examples (training samples) for feeding the model. It leads to the case that the model makes false predictions due to irrelevant features. To overcome this, it is advisable to make large training database, to use adequate number of layers and their size, to use “dropout” function and to ″augment″ samples, i.e. to transform images randomly in order to make our model never see twice the same image.

Those improvements were implemented in our model using *keras.preprocessing.image.ImageDataGenerators* class along with random image transformations and normalization operations. Each image was 1200 x 900 pixels in dimension, with the resolution of 96 dpi and 24-bit depth, compressed with JPG standard.

DL model was a simple stack of two convolutional layers with a ReLU activation and followed by max-polling layers. In top of it, we stick two fully connected layers and finish the model with a single unit and a softmax activation and categorical_crossentropy loss, since it is convenient for multiclass image classification. Dropout function was implemented with 0.5 factor in the last layer.

The randomizations that were used are rescaling, shear_range and zoom_range. The first one makes image scaling with 1/255 factor to enable model to perform faster with scaled image values between 0 and 1. The second one apply random shear transformations (displace each point in fixed direction by the amount proportional to its signed distance from the line that is parallel to that direction), while the last one makes random zooming inside images.

Training generator was constructed with 240x240 pixels image size, 32 batch size, 20 epochs and 1000 samples per epoch. In Fig. 3 are presented examples of images for each of four classes used in two different scenarios for DL image-based modulation recognition.

D. Model testing

Model performance is defined using training and validation accuracy which are the measures how well-developed model is. When training a machine learning model, one of the main things that should be avoided, as already being said is the overfitting. Practically, that term is related to the situation when a model fits the training data well, but it isn't able to generalize and make accurate predictions for data it hasn't seen before. In order to find out if a model is overfitting, a technique called cross-validation is used.

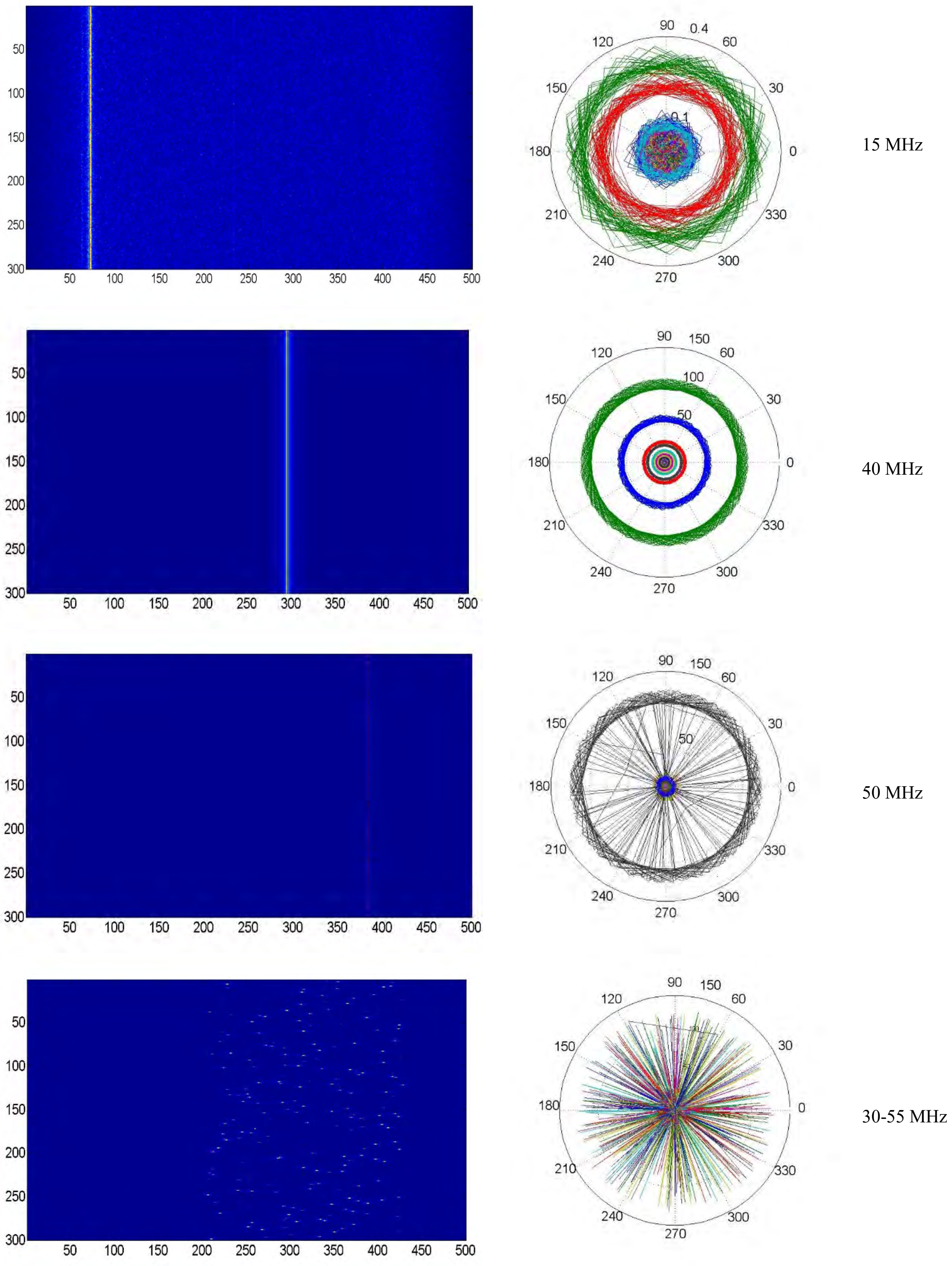


Fig. 3. Examples of images used in DL model (spectrogram-left, polar-right)

It requires the data to be split into two parts - the training set and the validation set. The training set which is in our case comprised of about 100 images in each class is used to train the model, while the validation set consisted of 6 images in our case is only used to evaluate the model's performance.

Metrics on the training set expressed as training loss and accuracy enable one to see how the model is progressing in terms of its training, and the metrics shown as validation loss and accuracy on the validation set that gives a measure of the quality of the model – how well it's able to make new predictions based on data it hasn't seen before.

III. CONCLUSION

Considering results given in Fig. 4, there are some conclusions that could be made on the developed model performing:

1. Training accuracy which is a measure of the training effectiveness is better on the same number of images in the training set for the polar representation of the spectrogram. It comes to the training accuracy of 90% in just 8 epochs, while for the same value of classical spectrogram image data set it is required to make at least 14 epochs;
2. Validation accuracy (the measurement of how the model performs when introduced to the data that hasn't been introduced to before) is quite better for the polar trained model, as well, since it comes to the high value of more than 85% very quickly (after only 2 epochs);
3. The same situation is considering validation loss as the summation of the errors made for each example in validation sets. Referring to the Fig. 4, it can be concluded that again polar based image classification model is much better since it classifies new images with fewer errors after just two epochs which is quite improvement related to classical spectrogram model.

Our future work is going to be directed towards the real-world application designing for radio emission pattern recognition. This indicates development of optimized DL models for near-real-time multiclass emission classification which is going to be a quite demanding challenge. In order to perform such a task, it is an obligate to use GPU (Graphics Processing Unit) as a platform for deep learning model training and validation because of the much faster processing capabilities.

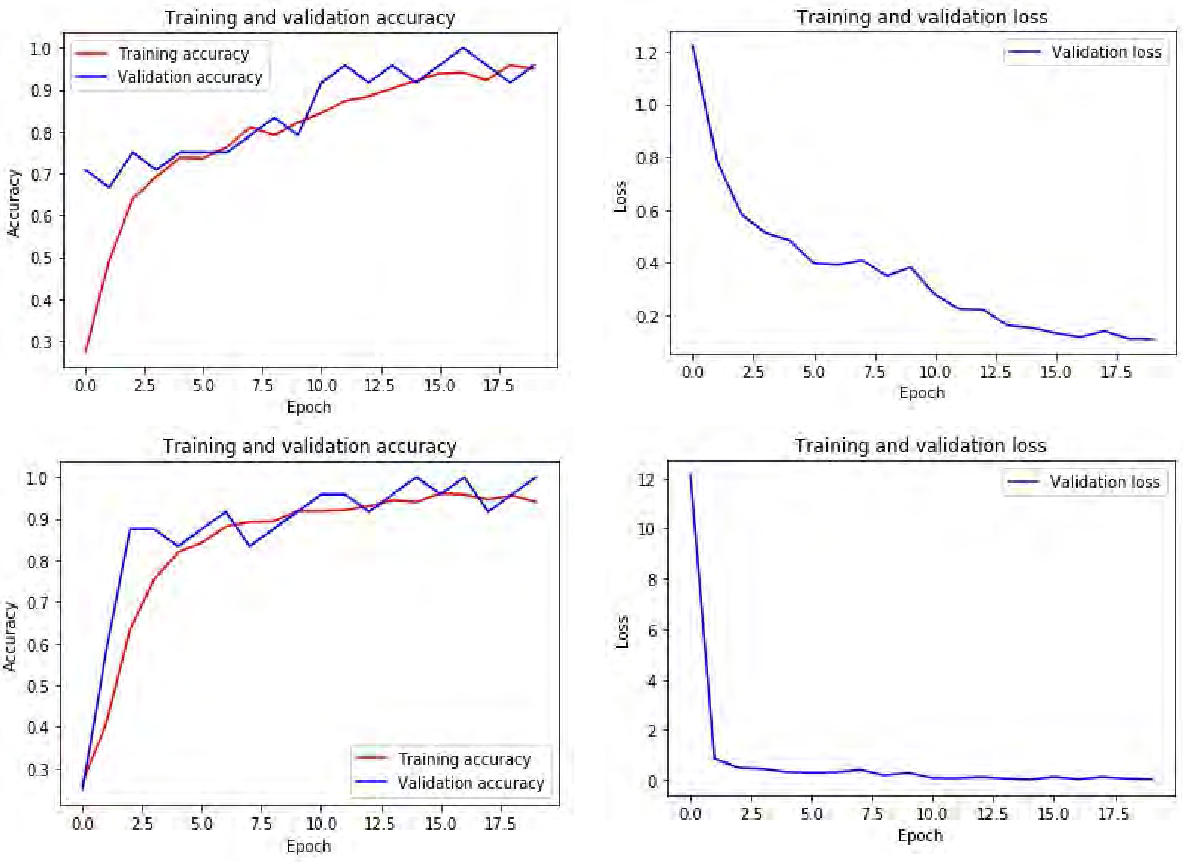


Fig. 4. Model performance – spectrogram (up), polar (down)

ACKNOWLEDGEMENT

This research is conducted under the project VA-TT/3/20-22 funded by University of Defence in Belgrade.

REFERENCES

[1] LeCun, Y., Bengio, Y., Hinton, G., “Deep Learning”, *Nature,* Vol. 521, pp. 436-444, May. 2015.

[2] Russell, S. J., Norvig, P., “Artificial Intelligence: A Modern Approach (3rd ed.)”, *Prentice Hall*, Upper Saddle River, New Jersey, 2009.

[3] Uçar, A., Demir, Y., Güzeliş, C., “Object recognition and detection with deep learning for autonomous driving applications”, *Simulation,* Vol. 93, pp. 759-769, 2017.

[4] Pischella, M., Le Ruyet, D., Pischella, M., Le Ruyet, D. (2015). „Multi-Carrier Modulations. In Digital Communications 2”, *John Wiley & Sons, Inc.*, 2015.

[5] Chollet, F., „Deep Learning with Python”, *Manning Publications Co*, Shelter Island, New York, 2018.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

New Classes of the Orthogonal Filters - An Overview

Saša S. Nikolić, Dragan Antić, Nikola Danković, Marko Milojković, and Staniša Perić

Abstract - This paper gives some results and contributions achieved in recently time on the topic of theory of orthogonal polynomials, i.e., orthogonal filters. This theory is based on new definitions and specific generalizations of orthogonal functions and polynomials, derived directly in complex domain, and their possible applications in the field of identification, modeling, signal processing and control of dynamical systems. Necessary short mathematical background is given, also. In this paper we give some main results for almost, improved almost, quasi-, generalized, and trigonometric orthogonal polynomials.

Keywords - Orthogonal polynomials, Almost orthogonal filetrs, Quasi-orthogonal filetrs, Generalized trigonometric functions.

I. INTRODUCTION

In last 10 yeats several classes of rational functions orthogonal in the complex domain, were developed. They were used for designing new classes of orthogonal filters. Orthogonal rational functions with fixed poles on the unit circle were published in the middle of the last century [1], and they represented generalizations of classical orthogonal polynomials of Szegö [2]. History of orthogonal polynomials can be find in [2]-[4]. There are many papers considering orthogonal systems and their applications in electronics, circuit theory, signal processing, communications, and control system theory [5-7]. On the other hand, concept of quasi-orthogonality is introduced for the first time in 1923 [8] as a tool for solving the problem of moments in mechanics. Quasi-orthogonal functions and especially quasi-orthogonal polynomials as well as numerous applications have been discussed in many papers [9-12]. In practice there is often need for transfer functions of general type i.e., with difference in orders of polynomials higher than one. This can be accomplished by using quasi-orthogonal functions and filters [13].

On the other side, almost orthogonal functions [13-16] can be successfully used for modeling and analysis of the systems with imperfections. First mentioned quasi-orthogonal functions [9] are also suitable for modeling and design of imperfect systems, especially if used in combination with almost orthogonal functions.

Authors of this paper have paid significant attention in recent years to the investigating new types of orthogonal polynomials and their possible applications in control systems and others applications. The main contributions of these researches are new types of orthogonal functions and polynomials with use of several different types of transformation functions for their generating [11-22]. Based on these functions we obtained new classes of improved almost and quasi-orthogonal polynomials as well as practically implemented orthogonal filters in the form of a printed circuit boards [13, 15]. As some possible application of the newly derived almost orthogonal filters can find in [14], [15].

It should be mentioned that in [13], new types of filters based on orthogonal Legendre and Malmquist quasi-orthogonal polynomials have been proposed. In [18], a new class of orthogonal Legendre type filters with complex poles and zeroes were designed with exemplar application in modelling of real cascade-connected industrial system.

II. GENERALIZED ORTHOGONAL FILTERS

A. Almost orthogonal polynomials and filters

Now, we introduce the concept of generalized orthogonal functions using well known Legendre polynomials for the sake of simplicity (the same concept can be used on other classical orthogonal polynomials as well) [15]. Our design is based on shifted Legendre polynomials orthogonal over interval (0, 1). On the other side, technical systems operate in the real time, so we need the corresponding approximation over interval (0, ∞). The solution is to use the substitution $x = e^{-t}$. In this manner, polynomial sequence orthogonal over (0, 1), become exponential polynomial sequence orthogonal over interval (0, ∞).

Now, we consider the orthogonal Legendre polynomials in their explicit form [14, 23]:

$$P_n(x) = \sum_{i=0}^{n} A_{n.i} x^i \ , \tag{1}$$

where:

$$A_{n.i} = \frac{1}{n!}(-1)^{n-i}\binom{n}{i}\frac{(n+i)!}{i!}. \tag{2}$$

These polynomials are orthogonal over interval (0, 1), with weight function $w(x)=1$, and the following definition of orthogonality based on inner product:

$$\int_0^1 P_m(x) P_n(x)\,dx = \begin{cases} 0, & m \neq n \\ N_i, & m = n, \end{cases} \tag{3}$$

and they can be successively used for modelling, and control of dynamical systems as well as for identification of specific systems parameters.

Corresponding almost orthogonal polynomials $P_n^{(\varepsilon)}(x)$ can be defined as [10, 23]:

$$\int_0^1 P_m^{(\varepsilon)}(x) P_n^{(\varepsilon)}(x)\,dx = \begin{cases} \varepsilon, & m \neq n \\ N_n^{(\varepsilon)}, & m = n, \end{cases} \tag{4}$$

where ε represent a very small positive constant $(0 < \varepsilon << 1)$. The connection between classical orthogonal and almost orthogonal polynomials is proved in [14] with the following relation:

$$P_n^{(\varepsilon)}(x) = P_n(x) + \sum_{k=1}^{n-1} \frac{b_k}{\left\| P_k \right\|^2} P_k(x), \tag{5}$$

where $\left\| P_k \right\|^2$ represents the square of the norm and b_k are polynomials dependent on ε. Obviously, for ε=0 almost orthogonal polynomials become strictly orthogonal, i.e. $\lim_{\varepsilon \to 0} P_n^{(\varepsilon)}(x) = P_n(x)$. For design of almost orthogonal filters, it is very convenient to use the following three term recurrent relation, derived in [14], [16]:

$$W_{n+1}^{(\varepsilon)}(s) = W_n^{(\varepsilon)}(s) + W_{n+1}(s) + W_n(s) k_n, \tag{6}$$

where: $k_n = \dfrac{b_n(\varepsilon)}{\left\| P_n \right\|^2} - 1, k_0 = \varepsilon$.

The first few members of the almost orthogonal polynomials $\left\{ P_n^{(\varepsilon)}(x) \right\}$ over interval (0, 1) with weight function $w(x)$=1 sequence are:

$$\begin{aligned} &P_0^{(\varepsilon)}(x) = 1, \\ &P_1^{(\varepsilon)}(x) = 2x - (1 - 2\varepsilon), \\ &P_2^{(\varepsilon)}(x) = 6x^2 - 6\left(1 - 12\varepsilon + 12\varepsilon^2\right)x + \\ &\qquad + \left(1 - 30\varepsilon + 36\varepsilon^2\right). \end{aligned} \tag{7}$$

On the Fig. 1 we can see the practical realization of the almost orthogonal filters with six section.

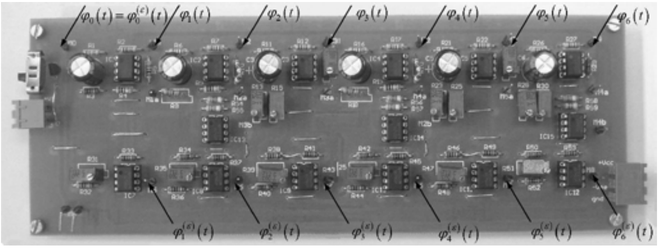


Fig. 1. The almost orthogonal filter of Legendre type with six sections - printed circuit board

Signals $\varphi_i^{(\varepsilon)}(t)$ from Fig. 1 represent Legendre type almost orthogonal functions generated by almost orthogonal filter [15]. This filter consists of the 15 µA 741CN amplifiers, 60 resistors, 12 potentiomets, and 6 capacitors.

B. Improved almost orthogonal polynomials

We defined almost orthogonality in different manner in order to accomplish further simplifications and improvements in filters design. First, we defined almost orthogonal Legendre polynomials $P_n^{(\delta)}(x)$ [15]:

$$P_n^{(\delta)}(x) = \sum_{i=0}^{n} A_{n,i}^{\delta} x^i, \tag{8}$$

where $A_{n,i}^{\delta}$ represent coefficients defined by:

$$A_{n,i}^{\delta} = (-1)^{n+i} \frac{\Gamma(n\delta + i + 1)}{\Gamma(n\delta + 1)\, i!\,(n-i)!}. \tag{9}$$

δ is a constant near to one: δ=1+ε≈1 and Γ is a symbol for the gamma function [2]. Parameter δ is uncertain quantity, which describes imperfection of the system. Variations of δ contain cumulative impacts of all imperfect elements, model uncertainties, and measurement noise on the system output. Range of variations can be determined by conducting several experiments. Hence, it is expected that responses obtained from different experiments are mutually different. The responses are in certain boundaries, which depend on parameter δ i.e., on the real system components quality, or the noise level present in signal.

After applying the substitution $x = e^{-t}$ to (8) and Laplace transform, we obtain transfer function suitable for designing almost orthogonal filters:

$$W_n^{(\delta)}(s) = \frac{\prod_{i=1}^{n}(s - i\delta)}{\prod_{i=0}^{n}(s + i)} = \frac{(s-\delta)(s-2\delta)\cdots(s-n\delta)}{s(s+1)(s+2)\cdots(s+n)}. \tag{10}$$

As a mapping function use transformation $\overline{s} + s = 0$, i.e. $\overline{s} = -s$. In this case, left semi plane of the complex plane s is being transformed into the right semi plane [7], [15]. Almost orthogonality can be analyzed from the following inner product:

$$N_{nm} = \oint_C W_n^{(\delta)}(s) \overline{W}_m^{(\delta)}(s) w(s)\,ds \tag{11}$$

with weight $w(s)$=1, m>n.

Now, applying the Cauchy theorem, we obtain:

$$N_{nm} = 2\pi j \sum_{k=1}^{m} \operatorname{Res}\left[W_n^{(\delta)}(s) \overline{W}_m^{(\delta)}(s) \right], \tag{12}$$

i.e.:

$$N_{nm} = 2\pi j \sum_{k=1}^{m} \frac{(-1)^{n+1} \prod_{i=1}^{n}(k + i\delta) \prod_{i=1}^{m}(k - i\delta)}{k^2 \prod_{i=1}^{n}(k - i) \prod_{i=1}^{m}(k + i)}. \tag{13}$$

The first few members in time domain of the improved almost orthogonal polynomials $\left\{ \varphi_i^{(\delta)}(t) \right\}$ over interval (0, ∞) with weight function $w(x)$=1 sequence are:

$$\begin{aligned}
&\varphi_0^{(\delta)}(t)=1,\\
&\varphi_1^{(\delta)}(t)=(\delta+1)e^{-t}-\delta,\\
&\varphi_2^{(\delta)}(t)=(\delta+1)(\delta+2)e^{-2t}-(\delta+1)(2\delta+1)e^{-t}+\delta^2.
\end{aligned}\tag{14}$$

On the Fig. 2 we can see the practical realization of the improved almost orthogonal filters with five section.

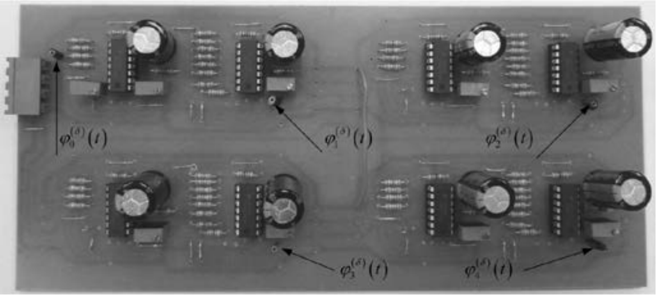


Fig. 2. The improved almost orthogonal filter of Legendre type with five sections – printed circuit board

This filter consists of the 8 LM 324N amplifiers, 48 resistors, 9 potentiomets, and 8 capacitors.

C. Quasi-orthogonal and generalized quasi-orthogonal filters

The final generalization of the concept of orthogonality can be introduced by the following definition of quasi-orthogonality for the polynomial set $P_n(x)$ [17]:

$$\begin{aligned}
&\left(P_n^k(x),P_m^k(x)\right)=\int_a^b w(x)P_n^k(x)P_m^k(x)dx=\\
&=\begin{cases}0, & 0\le m\le n-k-1,\\ N_{n,m}^k\ne 0, & n-k\le m\le n,\end{cases}
\end{aligned}\tag{15}$$

where k represents the order of quasi-orthogonality, a and b are the limits of quasi-orthogonality interval, and $w(x)$ is the weight function.

In our case of Legendre quasi-orthogonal polynomials of the first order [17]:

$$\int_0^1 P_n^k(x)P_m^k(x)xdx=\begin{cases}=0, & 0\le m\le n-k-1\\ \ne N_{n,m}^k, & n\ge k+1,\end{cases}\tag{16}$$

$$P_n^k(x)=x\sum_{i=0}^{n}A_{n,i}^k x^{i-1},\tag{17}$$

where:

$$A_{n,i}^k=\frac{(-1)^{n-k+i}(n+i-k-1)!}{i!(i-1)!(n-i)!}.\tag{18}$$

The first few members of quasi-orthogonal polynomials of order k=1, $P_n^1(x)$ over interval (0, 1) with weight function $w(x)$=x sequence are:

$$\begin{aligned}
&P_1^1(x)=-x+1,\\
&P_2^1(x)=-2x^2+3x-1,\\
&P_3^1(x)=-5x^3+10x^2-6x+1,\\
&P_4^1(x)=-14x^4+35x^3-30x^2+10x-1.
\end{aligned}\tag{19}$$

If we apply (15) on the almost orthogonal Legendre polynomials defined by (8) and (9), we can obtain generalized quasi-orthogonal Legendre polynomials [19] over interval (0, 1) with weight function $w(x)$=1:

$$P_n^{(k,\delta)}(x)=\sum_{i=0}^{n}A_{n,i}^{k,\delta}x^i,\tag{20}$$

where:

$$A_{n,i}^{(k,\delta)}=(-1)^{n+i+k}\frac{\prod_{j=1}^{n-k}(i+j\delta)}{i!(n-i)!}.\tag{21}$$

For example, the first member of first order (k=1) generalized quasi-orthogonal Legendre polynomials of this sequence are:

$$\begin{aligned}
&P_1^{(1,\delta)}(x)=-x+1,\\
&P_2^{(1,\delta)}(x)=-\frac{(\delta+2)}{2}x^2+(\delta+1)x-\frac{\delta}{2},\\
&P_3^{(1,\delta)}(x)=-\frac{(\delta+3)(2\delta+3)}{6}x^3+(\delta+1)(\delta+2)x^2-\\
&\qquad-\frac{(\delta+1)(2\delta+1)}{2}x+\frac{\delta^2}{3},
\end{aligned}\tag{22}$$

On the Fig. 3 we can see the practical realization of the generalized quasi-orthogonal filters with five section.

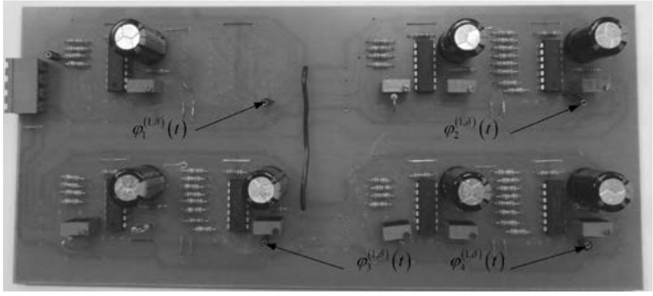


Fig. 3. The generalized quasi-orthogonal filter of Legendre type of first order (k=1) with five sections – printed circuit board

This filter consists of the 5 µA 741 amplifiers, 16 resistors, 10 potentiomets, and 7 capacitors.

III. COMPLEX RATIONAL FUNCTIONS AND TRIGONOMETRIC FILTERS OF LEGENDRE TYPE

Complex rational functions were intensively investigated in the 1970s and the 1980s. Rational functions with complex poles (and appropriate filters) can be also orthogonal [18]. These functions provide generating a sequence of rational functions with arbitrary complex-conjugate poles using three term recurrent relation [24]. Later, the complex rational functions were used for designing appropriate filters whose transfer function has complex-conjugate poles. In this paper, we present complex rational functions which are used for derivation of complex Legendre polynomials of trigonometric form. Now, we will briefly define these new complex rational functions and appropriate polynomials and practical implementation will be shown in designing appropriate orthogonal filters [25].

Denote poles of function:

$W_n(s)=\dfrac{\prod_{k=1}^{n-1}\left(s^2+\bar{a}_k s+\bar{b}_k\right)}{\prod_{k=1}^{n}\left(s^2+a_k s+b_k\right)}$, $n\geq 2$ with s_k. Then the following relation is valid: $s^2+a_k s+b_k=\left(s-s_k\right)\left(s-s_k^*\right)$, where s_k^* is complex-conjugate related to s_k i.e., $s_k=\sigma_k+j\omega_k$, $s_k^*=\sigma_k-j\omega_k$. If we apply Heaviside's expansion on function $W_n(s)$ we obtain:

$$W_n(s)=\sum_{i=1}^{n}\frac{A_i}{s-s_i}+\sum_{i=1}^{n}\frac{B_i}{s-s_i^*}\,. \tag{23}$$

Taking into account the residues of function (6), we obtain:

$$\begin{aligned} A_i&=\frac{\prod_{k=1}^{n-1}\left(s_k^2+\bar{a}_k s+\bar{b}_k\right)}{\left(s_i-s_i^*\right)\prod_{k=1}^{i-1}\left(s_i^2+a_k s_i+b_k\right)\prod_{k=i+1}^{n}\left(s_i^2+a_k s_i+b_k\right)},\\ B_i&=\frac{\prod_{k=1}^{n-1}\left(\left(s_i^*\right)^2+\bar{a}_k s_i^*+\bar{b}_k\right)}{\left(s_i-s_i^*\right)\prod_{k=1}^{i-1}\left(\left(s_i^*\right)^2+a_k s_i^*+b_k\right)\prod_{k=i+1}^{n}\left(\left(s_i^*\right)^2+a_k s_i^*+b_k\right)}. \end{aligned} \tag{24}$$

Applying inverse Laplace transformation to function (23) we obtain:

$$\begin{aligned} W_n(t)&=\sum_{i=1}^{n}A_i e^{s_i t}+\sum_{i=1}^{n}B_i e^{s_i^* t},\\ W_n(t)&=\sum_{i=1}^{n}e^{\sigma_i t}\left(\alpha_i\cos\omega_i t+\beta_i\sin\omega_i t\right)\,. \end{aligned} \tag{25}$$

These functions are orthogonal on (0, ∞) with weight function $w(t)=1$, i.e., the following expression is valid [19]:

$$\int_0^\infty W_m(t)W_n(t)\,dt=\begin{cases}0, & m\neq n\\ N_n^2, & m=n\,.\end{cases} \tag{26}$$

Complete orthogonality mathematical proof is given in [18, 23]. The first few first member of Legendre polynomials in time domain are:

$$\begin{aligned} P_1(t)&=e^{-t}\sin t,\\ P_2(t)&=\frac{3}{10}e^{-2t}\left(4\cos 2t+3\sin 2t\right)+\frac{2}{5}e^{-t}\left(-3\cos t+\sin t\right),\\ P_3(t)&=\frac{2}{39}e^{-3t}\left(63\cos 3t-16\sin 3t\right)+\\ &+\frac{12}{65}e^{-2t}\left(-11\cos 2t+23\sin 2t\right)-\frac{3}{5}e^{-t}\left(2\cos t-\sin t\right). \end{aligned} \tag{27}$$

On the Fig. 4 we can see the practical realization of the generalized orthogonal filters with three section. This filter consists of the 17 LM 324N amplifiers, 68 resistors, and 11 capacitors.

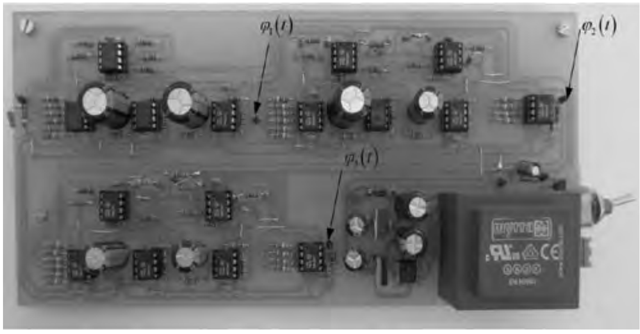


Fig. 4. The generalized orthogonal filter of Legendre type with three sections – printed circuit board

IV. CONCLUSION

In this paper we gave survey study on new classes of orthogonal filters (almost, improved almost, quasi-, generalized quasi-, and trigonometric) which have been derived in recent 10 years. These polynomials could have wide spectrum of applications in various science and technical fields. We have already designed the appropriate filters based on new classes of filters and gave experimentally printed circuit board in this paper.

We are developing some new class of orthogonal rational functions which can be used for designing new orthogonal filters.

APPENDIX

We present some block diagrams and electronic schemes of the proposed filetrs. On the Fig. 5 we present the improved almost orthogonal filter of Legendre type. The experimental signals (Fig. 6) sensed on implemented filter for the value of the parameter δ=1.02 and the amplitude of the input signal 2.5V rebound, is shown in Fig. 5. On the Fig. 7 we shown the experimental signals sensed on implemented generalized quasi-orthogonal filter of Legendre type of first order (k=1)for the value of the parameter δ=1.02 and the amplitude of the input signal 2.5V rebound. On the Fig. 8 we present the generalized orthogonal filter of Legendre type with complex zeroes and poles.

ACKNOWLEDGEMENT

This paper was realized as a part of the projects III 43007 and TR 35005, financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia within the framework of integrated and interdisciplinary research for the period 2011-2019.

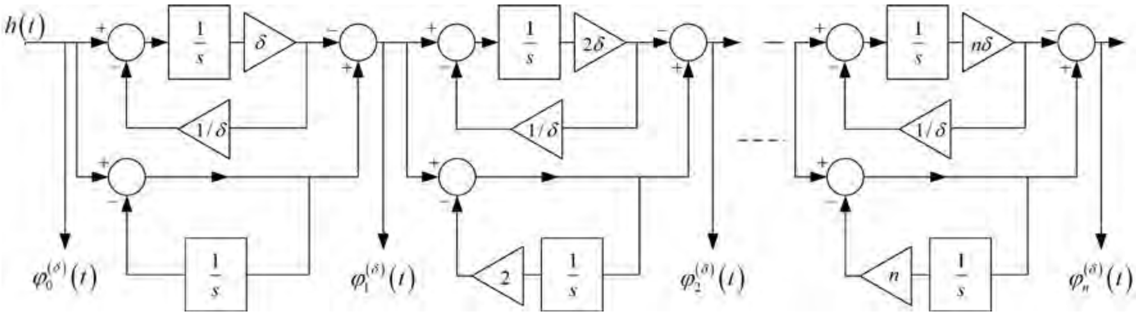


Fig. 5. The improved almost orthogonal filter of Legendre type – analogue scheme

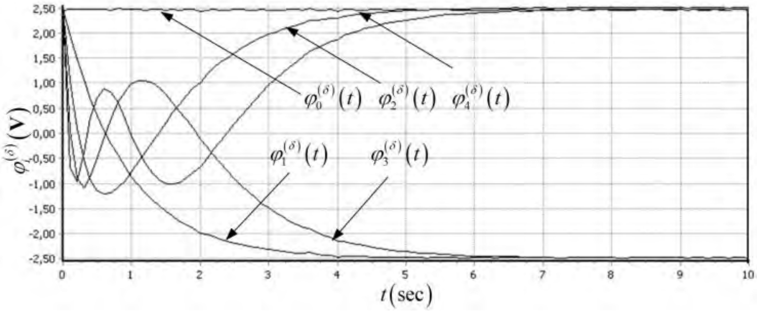


Fig. 6. The improved almost orthogonal filter of Legendre type – signals sensed on printed circuit board

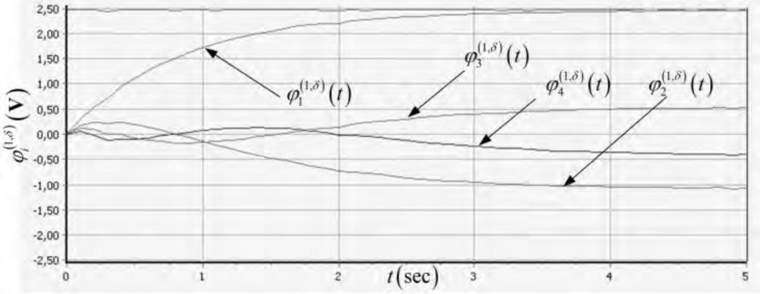


Fig. 7. The generalized quasi-orthogonal filter of Legendre type of first order (k=1) – signals sensed on printed circuit board

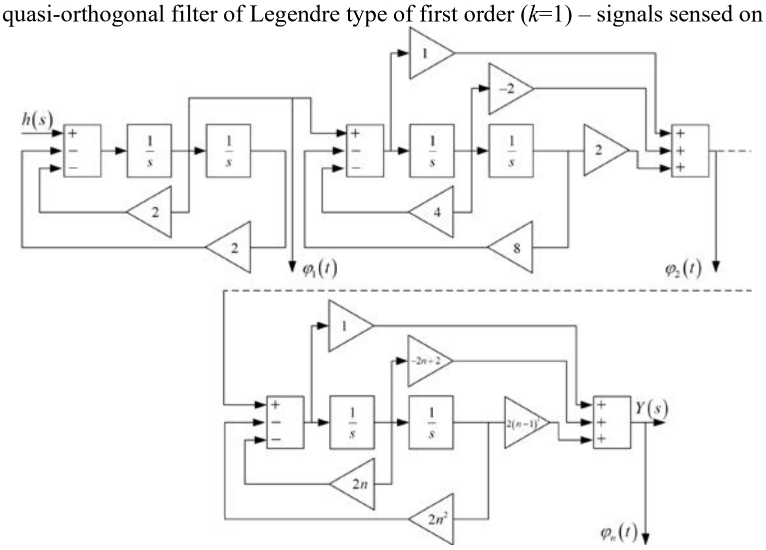


Fig. 8. The generalized orthogonal filter of Legendre type with complex zeroes and poles – analogue scheme

REFERENCES

[1] M. M. Djrbashian, "Orthogonal systems of rational functions on the circle with given set of poles", *Dokl. Akad. Nauk SSSR*, vol. 147, pp. 1278–1281, 1962.

[2] G. Szegö, *Orthogonal Polynomials*, Providence, American Mathematical Society, Colloquium Publications, 1975.

[3] M. M. Djrbashian, "A survey on the theory of orthogonal systems and some open problems", In P. Nevai (ed.) Orthogonal polynomials. Springer, Netherlands, pp. 135–146, 1990.

[4] G. V. Badalyan, "Generalisation of Legendre polynomials and some of their applications" *Akad. Nauk. Armyan. SSR Izv. Ser. Fiz.-Mat. Estest. Tekhn. Nauk*, vol. 8, no. 5, pp. 1–28, 1955.

[5] B. Danković, G. V. Milovanović, S. Rančić, "Malmquist and Müntz Orthogonal Systems and Applications", In Rassias TM (ed.), Inner product spaces and applications, Harlow: Addison-Wesley Longman, pp. 22–41, 1997.

[6] P. V. Hof Heuberger, B. Wahlberg, *Modelling and identification with rational orthogonal basis functions*. London: Springer-Verlag, 2005.

[7] S. Nikolić, D. Antić, B. Danković, M. Milojković, Z. Jovanović, S. Perić, "Orthogonal functions applied in antenna positioning", *Advances in Electrical and Computer Engineering*, vol. 10, no. 4, pp. 35–42, 2010.

[8] M. Riesz, "Sur le probleme des moments", *Troisième Note, Arkiv för Matematik, Astronomi och Fysik,* vol. 17, no. 16, pp. 1–52, 1923.

[9] T. S. Chihara, "On quasi-orthogonal polynomials", *Proc. American Mathematical Society*, vol. 8, pp. 765–767, 1957.

[10] J. S. Dehesa, F. Marcellan, A. Ronveaux, "On orthogonal polynomials with perturbed recurrence relations", *Journal of Computational and Applied Mathematics*, vol. 30, pp. 203–212, 1990.

[11] D. Antić, Z. Jovanović, V. Nikolić, M. Milojković, S. Nikolić, N. Danković, "Modeling of cascade-connected systems using quasi-orthogonal functions", *Electronics and Electrical Engineering*, vol. 18, no. 10, pp. 3–8, 2012.

[12] M. Alfaro, L. Moral, "Quasi-orthogonality on the unit circle and semi-classical forms", *Port. Mathematica*, vol. 51, pp. 47–62, 1991.

[13] M. T. Milojković, D. S. Antić, S. S. Nikolić, Z. D. Jovanović, S. Lj. Perić, "On a new class of quasi-orthogonal filters", *International Journal of Electronics*, vol. 100, no. 10, pp. 1361–1372, 2013.

[14] B. Danković, S. Nikolić, M. Milojković, Z. Jovanović, "A class of almost orthogonal filters", *Journal of Circuits, Systems, and Computers*, vol. 18, no. 5, pp. 923–931, 2009.

[15] D. Antić, B. Danković, S. Nikolić, M. Milojković, Z. Jovanović, "Approximation based on orthogonal and almost orthogonal functions", *Journal of the Franklin Institute*, vol. 349, no. 1, pp. 323–336, 2012.

[16] M. Milojković, S. Nikolić, B. Danković, D. Antić, Z. Jovanović, "Modelling of dynamical systems based on almost orthogonal polynomials", *Mathematical and Computer Modelling of Dynamical Systems*, vol. 16, no. 2, pp. 133–144, 2010.

[17] D. Antić, S. Nikolić, M. Milojković, N. Danković, Z. Jovanović, S. Perić, "Sensitivity analysis of imperfect systems using almost orthogonal filters", *Acta Polytechnica Hungarica*, vol. 8, no. 6, pp. 79–94, 2011.

[18] S. S. Nikolić, D. S. Antić, S. Lj. Perić, N. B. Danković, M. T. Milojković, "Design of generalised orthogonal filters: Application to the modelling of dynamical systems", *International Journal of Electronics*, vol. 103, no. 2, pp. 269–280, 2016.

[19] S. S. Nikolić, D. S. Antić, M. T. Milojković, M. B. Milovanović, S. Lj. Perić, D. B. Mitić, "Application of neural networks with orthogonal activation functions in control of dynamical systems", *International Journal of Electronics*, vol. 103, no. 4, pp. 667–685, 2016.

[20] S. Lj. Perić, D. S. Antić, M. B. Milovanović, D. B. Mitić, M. T. Milojković, S. S. Nikolić, "Quasi-sliding mode control with orthogonal endocrine neural network-based estimator applied in anti-lock braking system", *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 2, pp. 754–764, 2016.

[21] N. Danković, D. Antić, S. Nikolić, S. Perić, M. Spasić, "Generalized cascade orthogonal filters based on symmetric bilinear transformation with application to modeling of dynamic systems", *FILOMAT*, vol. 32, no. 12, pp. 4275–4284, 2018.

[22] N. Danković, D. Antić, S. Nikolić, M. Milojković, S. Perić, "New class of digital Malmquist-type orthogonal filters based on generalized inner product; application to modeling DPCM system", *FACTA UNIVERSITATIS Series: Mechanical Engineering*, vol. 17, no. 3, pp. 385–396, 2019.

[23] S. S. Nikolić, "Application of generalized classical filters in intelligent control systems", *PhD Dissertation*, University of Niš, Faculty of Electronic Engineering, Republic of Serbia, 2014.

[24] K. Deckers, A. Bultheel, "Orthogonal rational functions with complex poles: The Favard theorem", *Journal of Mathematical Analysis and Applications*, vol. 356, no. 2, pp. 764–768, 2009.

[25] N. B. Danković, D. S. Antić, S. S. Nikolić, M. T. Milojković, S. Lj. Perić, "New class of digital orthogonal filters based on bilinear transformation with one application", *Proceedings of the XIV International Conference on Systems, Automatic Control and Measurements, SAUM 2018,* Niš, Serbia, November 14.-16., 2018., pp. 110–113.

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

SARIMA and ANN approaches in day-ahead power consumption forecasting

Igor Jovanović, Andrija Petrušić, Miona Andrejević Stošović and Dragan Mančić

Abstract - In this paper, we will present two models, a seasonal autoregressive integrated moving average model (SARIMA) and an artificial neural network (ANN) model in order to forecast hourly electricity consumption of energy in industry for a day-ahead. We will start with a brief analysis of the global electricity market with special reference to the Serbian market. Next, the daily electricity consumption amounts between August 1st and December 19th 2019. will be analyzed using statistical tools. According to the obtained results, we will compare the two models.

Keywords – ANN, SARIMA, energy consumption, forecasting, seasonality.

I. INTRODUCTION

Liberalization of electric power sector in the last two decades led to very dynamic wholesale markets and deregulated retail markets. At wholesale market level, traders are focused on price dynamics trends, while on retail level for supplying electricity to final consumers the consumption forecast can make all the difference for fostering cost-efficiency [1].

Furthermore, forecasting energy consumption on short-term, mid-term and long-term level is essential when defining strategies for production planning and expansion of infrastructure capacities for electric power systems.

A good approach to consumption prediction of individual consumers can have a far-reaching effect on the aggregated level. Strategic decision-making tools are more reliable when based on proper models for determining patterns of individual consumptions for various types of industries.

It can be observed on the example of electric power suppliers as balance responsible parties at liberalized energy markets how important is having an adequate mathematical model for forecasting individual consumption. The balance responsibility refers to the penalty-based financial mechanism developed to force market participants to properly anticipate their energy needs on short-term level.

Electric power supplier is obliged to inform in a day-ahead manner transmission system operator (TSO) how much of electric power needs to be provided on hourly level for all his electricity consumers. This is important, because TSO needs to maintain the balance in real-time between production and consumption for entire energy system, based on the data provided from all balancing parties active on the market. There is a certain tolerance that TSO will allow to balancing parties, but at the end if anticipated hour consumptions are not matching actual consumed energy and the tolerance for imbalance goes above established limits, the balancing party will be penalized through settlement prices defined for each hour retroactively based on pre-established methodology issued by Regulatory body.

In liberalized power markets, balance responsibility and imbalance settlement are two closely related elements that constitute the essential part of a balancing market, which is an institutional arrangement establishing market-based balancing [2].

In the end, market participant proportionally to the imbalance they created in the system for each hour pays generated imbalance costs. If Electric power supplier lacks adequate forecasting tool and gives poor judgment on how much electricity is needed in each hour for all its consumers that are in his balancing group, unnecessary expense occurs for each estimation that at the end is above the allowed imbalance limits for measured consumption.

If consumption of the balancing group is observed only on an aggregated level, many studies have shown that there is a correlation with outside temperature, but balancing group dynamics is mostly out of focus. Therefore, the smaller balancing group is (fewer metering points and less volume) and if bigger, the fluctuation of client is on yearly level, there is more need for Electric power supplier to estimate balancing group needs through forecasting models of individual energy consumptions for each consumer or metering point.

The importance of electricity consumption forecasting on one hand, and its complexity on the other hand, have motivated many researchers in this area. In the literature, there are numerous studies on electricity consumption and demand estimation. In these studies, some of commonly used methods are stationary time series models [3], regression models [4], [5], [6] econometric models [7]. However, most of time series models are linear predictors, while electricity consumption is inherently a nonlinear function. So, the behavior of electricity consumption series may not be completely captured by the time series techniques. To solve this problem, some other research papers have proposed Artificial Neural Networks (ANN) and genetic algorithms for electricity consumption

Igor Jovanović, Andrija Petrušić, Miona Andrejević Stošović and Dragan Mančić are with the University of Niš, Faculty of Electronic Engineering, E-mail: igor.jovanovic@elfak.ni.ac.rs, andrija.petrusic@gmail.com, miona.andrejevic@elfak.ni.ac.rs, dragan.mancic@elfak.ni.ac.rs.

forecasting [8], [9].

The main purpose of this study using measured hourly-based consumption data is to develop an adequate forecasting model for predicting electricity consumption for individual consumers that can be used as a proper tool by Electric Power Supplier to avoid unnecessary imbalance costs to properly estimate energy needs on an hourly level.

For the purpose of this study, United Green Energy Ltd. [10] consulting company for Energy market has provided data of an individual consumer. Data was strategically chosen so irregular intermittent consumption profile can be used as the input data set. Furthermore, this consumption profile lacks correlation with global factors (e.g. outside temperature), as it is mostly dependable on the business potential company that creates for itself.

Collected hourly data is presented in the form of the time series for the period August 1st –December 19th 2019.

II. FORECASTING MODELS

A. Seasonal Autoregressive Integrated Moving Average Model

Autoregressive Integrated Moving Average Model (ARIMA) is a widely used time series analysis model in statistics. ARIMA(p, d, q) is a kind of short-term prediction model in time series analysis, where p, d and q are non-negative integers that correspond to the order of the autoregressive, integrated and moving average parts of the model, respectively. The ARIMA model is also applicable for non-stationary time series that have some clearly identifiable trends.

The periodicity, which has an impact not only in the electricity market but also in other sectors of the economy, is an important indicator for planning and policy-making [6]. Periodicity of the periodical time series is usually due to seasonal changes. For such time series, a Seasonal ARIMA (SARIMA) is used. Therefore, SARIMA(p, d, q)(P, D, Q)$_m$ is used for time series with seasonality, where P, D and Q are relevant seasonal autoregressive parameter, seasonal integrated parameter and seasonal moving average parameter, respectively, and m is period of time series. The seasonal part of the model is very similar to the non-seasonal part, but it is involved in backshifts of the seasonal period.

Fig. 1 displays the time series plot for electricity consumption data in kWh. The series displays considerable fluctuations over time, and a stationary model does not seem to be reasonable.

To stabilize the variance, data need to be rescaled using the formula:

$$a_i = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)}, \qquad (1)$$

where a_i is the rescaled value, y_i represents the original data, and $\min(y_i)$ and $\max(y_i)$ are the minimum and maximum values of the original data set, respectively. After stabilization of variance, the standard deviation is reduced from 22.1195 to 0.1991.

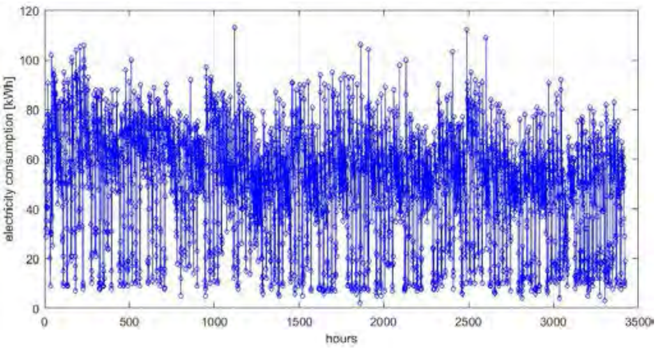


Fig. 1. The hourly electricity consumption from August 1^{st} –December 19^{th} 2019

The sample autocorrelation function (ACF) for the rescaled data is displayed in the Fig. 2. From that figure, one can see that there is a seasonal period of time (24 lags). In addition, software implementation of an augmented Dickey–Fuller (ADF) test for level stationarity applied to the rescaled data leads to a test statistic of −9.3863 for p<0.01. Although the test showed that the data was stationarity, a seasonal difference of the rescaled data series must be applied.

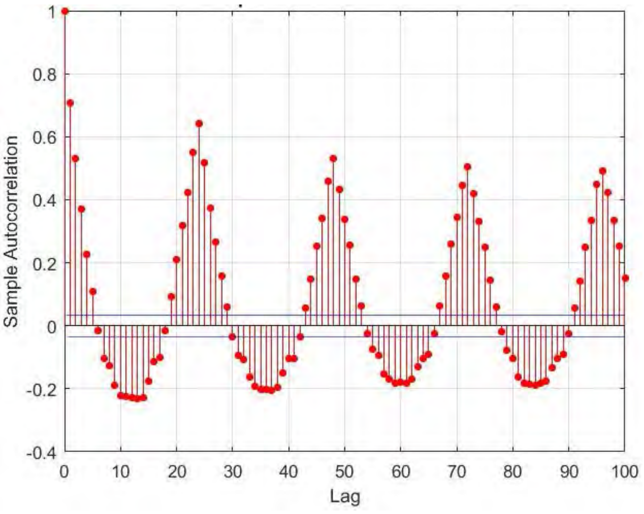


Fig. 2. The sample ACF for the rescaled electricity consumption data

The sample ACF and partial autocorrelation function (PACF) for the seasonal differences of the rescaled electricity consumption data are shown in Fig. 3.

On the basis of this plot, one can well consider a stationary model as appropriate. After first differencing, the standard deviation is reduced to 0.1711, and ADF test for level stationarity applied to data leads to a test statistic of −87.5573 for $p<0.01$. Taking the second seasonal differences also results in stationarity (i.e., the trend is removed), but leads to an over differences series with a standard deviation which moved to a level higher than the series after first differencing ($\sigma^2=0.4876$).

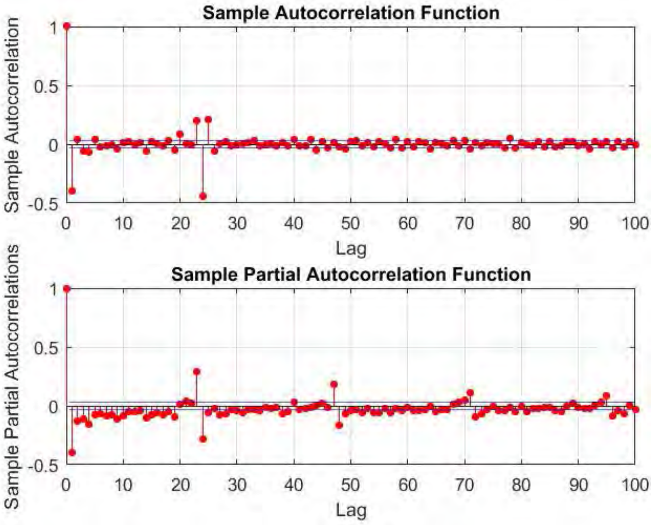


Fig. 3. Sample ACF and PCF for seasonal differences of the rescaled electricity consumption data

After the order of seasonal I(D) terms has been identified (here, 1), in SARIMA$(p,d,q)(P,D,Q)_m$, the next step is to determine whether the pattern of autocorrelation can be better explained by (S)AR terms, (S)MA terms, or a combination of both.

ACF indicating that there are two significant seasonal and non-seasonal spikes and PACF will die out over time. In this case, theoretically, we should use SMA(2) and MA(2) after differencing the data. The orders p and P are selected from a reasonable range of non-negative values to create several SARIMA models.

In order to maximize the probability of obtaining the data that we have observed the best SARIMA$(1,0,2)(2,1,2)_{24}$ model (Table 1) was selected based on the lowest Akaike's Information Criterion value (AIC=-3580). The lower value of Schwarz Bayesian Information Criterion (BIC=-3537.1) and lower root mean square error were desired. The low RMSE indicates that the dependent series is closest to the model predicted levels and thus, the predictive model is useful for forecasting purposes.

Through a combination of regression modelling and one seasonal differencing proposed model successfully extracted all the autocorrelation from the data in order to achieve more efficient forecasts.

The relative success of statistical models in reproducing the measured time-series can also be measured in terms of residuals of error. The frequency distributions of the residuals of the SARIMA models for all data are presented in Figure 4. The points of a quantile-quantile plot of the residuals seem to follow the straight line fairly closely. This graph would not lead us to reject normality of the error terms in this model. In addition, the Ljung-Box Q-test for residual autocorrelation applied to the residuals, produces a test statistic of 66.8844, which corresponds to a p-value <0.01, and we would not reject normality based on this test. To check on the independence of the error terms in the model, we consider the sample ACF of the residuals in Figure 5. In other words, the residuals follow a linear trend. Thus, the residuals are normally distributed. In general, the model shows good forecasting accuracy and can be used to predict future values.

TABLE I

PARAMETERS VALUE OF THE SARIMA$(1,0,2)(2,1,2)_{24}$ MODEL

Parameter	Value
Constant	0.401641
AR{1}	0.032015
SAR{1}	0.00657
SAR{2}	0.005856
MA{1}	0.000825
MA{2}	0.000824
SMA{1}	0.033497
SMA{2}	0.02343

Cross-validation prediction checks of proposed SARIMA model can predict and verify the value for the observed period, and the forecast future trends.

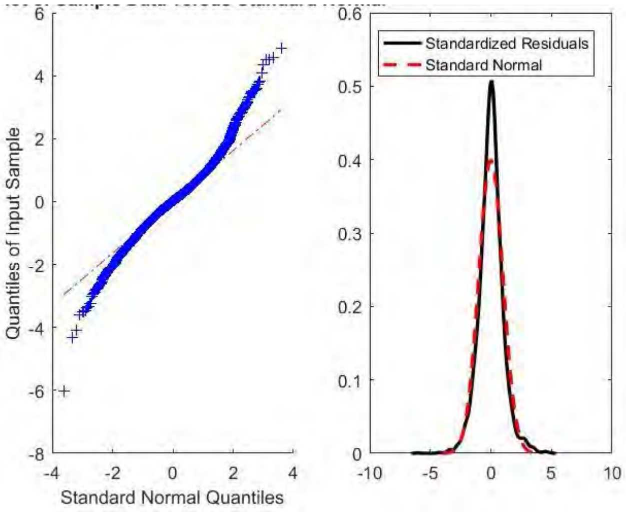


Fig. 4. Plots of the fitted SARIMA$(1,0,2)(2,1,2)_{24}$ model ACF and PCF for seasonal differences of the rescaled electricity consumption data

The model predicts the rescaled electricity consumption data (testing set, or observed data) for the next 139 days, based on the 37 hours of past data (presample data) with RMSE=19.2581, while minimum presample data for the selected model is 27 hours (Fig. 6). From Fig. 6 it can be concluded that overall model performance is satisfactory for all the electricity consumption data, although the observed peaks are reproduced in some cases with higher amplitude.

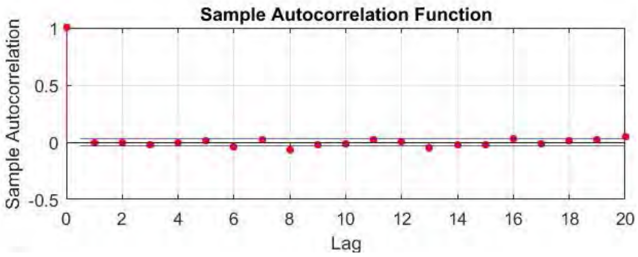


Fig. 5. Sample ACF for the SARIMA residuals

Verified value for the observed period can serve as confirmation of the adequacy of the proposed model, while forecasted future trend for electricity consumption data can be an indication of a future state. Fig. 7 shows a comparison between testing set and training set. For the sake of clarity, Fig. 7 shows values for the presample (black line) and observed (red dotted line) data only for the last four days (the data are already shown in Fig. 6).

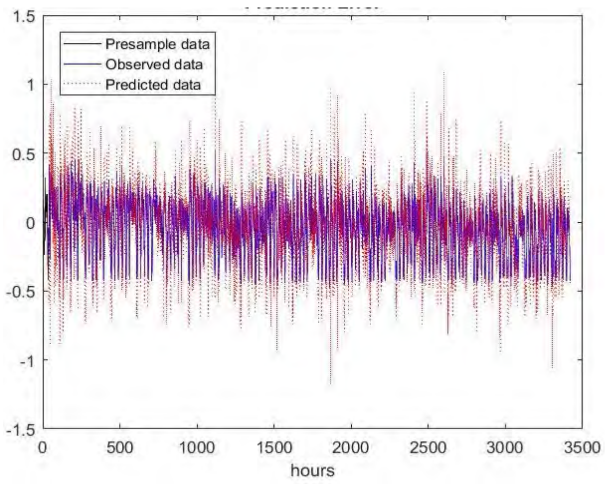


Fig. 6. Diagnostics proposed SARIMA model with rescaled data

Fig. 8 shows forecasted values (red line) for the day-ahead period. Evaluating forecast accuracy is accomplished by examining the residuals for any systematic pattern of misspecification. Forecasts should ideally be located within the 95% confidence limits, and formal statistics can be calculated from the model residuals in order to evaluate its adequacy. In other words, according to Fig. 8, the real values appear within the 95% confidence intervals (between red dotted lines).

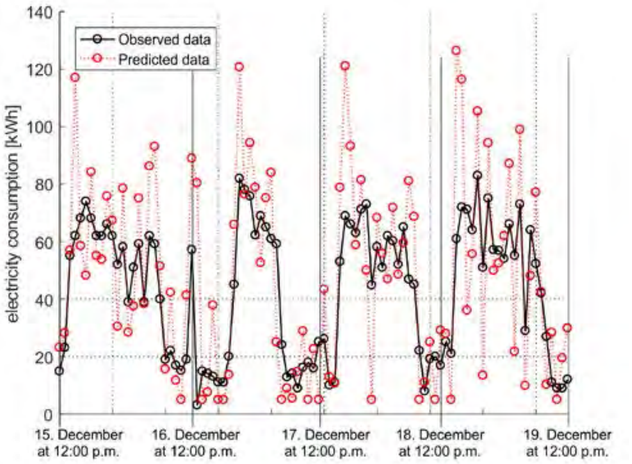


Fig. 7. Diagnostics proposed SARIMA model with original power consumptions data

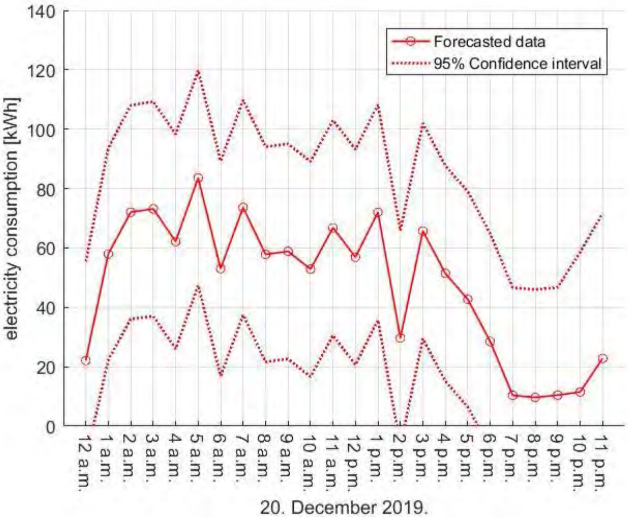


Fig. 8. Forecasted power consumption data using SARIMA model for the day-ahead period

B. ANN approach

Artificial neural networks have been used for short-time forecasting in electronics [11], [12]. Models were also developed for prediction of electricity consumption on a monthly and weekly level.

In this paper we will present a model based on a recurrent neural network. In fact, we used data from 24 previous measurements (measurement is done every hour), so we use consumption from a day before in order to predict consumption in the next hour. So, the artificial neural network has 24 inputs, describing what was happening for the previous 24 hours. For the network training we used the next consumption value as a network output. The obtained network has one hidden layer with 40 hidden neurons. After the training was completed, a training error was $1.23*10^{-17}$, what is negligible. That means that when data used in the training process are used as an excitation to the obtained ANN, no error is occurred.

In the Fig. 9 we present forecasted data for the next day, i.e. for next 24 hours. Forecasted data is obtained when we excited ANN with values for next 24 hours. We also should stress that part of the data used for prediction are also predicted in earlier iterations.

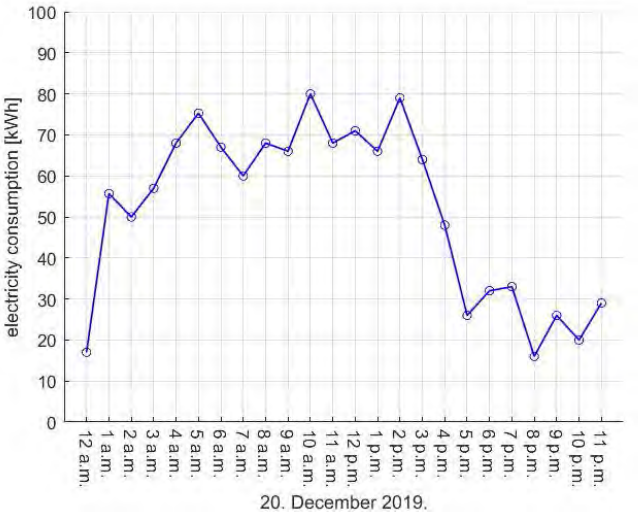


Fig. 9. Forecasted power consumption data using ANN model for the day-ahead period

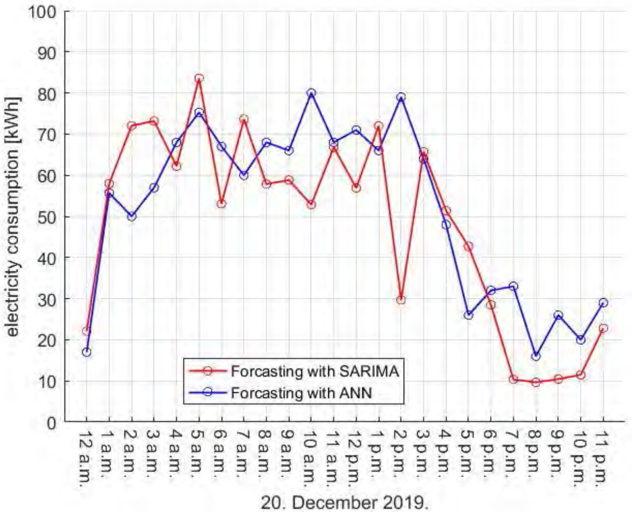


Fig. 10. Comparison of forecasted power consumption data using SARIMA and ANN model for the day-ahead period

III. CONCLUSION

In this paper, two approaches, statistical and neural network approach, were deployed for obtaining models for time series, i.e. day-ahead power consumption forecasting.

By analyzing and comparing forecasted data obtained by these two approached, we can see from the Fig. 10 that forecasted data have the same trend. We can also notice from the paper that ANN approach is much simpler.

It can be seen from the measured data that the energy consumption is correlated with outside temperature and with few more factors related to the daily operations of the observed company. In our future work, we plan to do further research and develop some hybrid forecasting methods which will take into account most of the above parameters.

ACKNOWLEDGEMENT

The research presented in this paper is financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia under the projects TR33035 and TR32004.

REFERENCES

[1] Pepermans, G., “European energy market liberalization: experiences and challenges”, *IJEPS*, Vol. 13, pp. 3–26, 2019.

[2] Van der Veen, R. A. C., Hakvoort, R. A., “Balance responsibility and imbalance settlement in Northern Europe — An evaluation”, *6th International Conference on the European Energy Market*, Leuven, pp. 1-6, May 2009.

[3] Fosso, O. B., Gjelsvik, A., Haugstad, A., Mo B., Wangensteen, I., “Generation scheduling in a deregulated system. The Norwegian case”, *IEEE Transactions on Power Systems*, Vol. 14, No. 1, pp. 75-81, February 1999.

[4] Bianco, V., Manca, O., Nardini, S., “Electricity consumption forecasting in Italy using linear regression models”, *Energy*, Vol. 34, No. 9, pp. 1413-1421, September 2009.

[5] Abdel-Aal, R.E., Al-Garni, A.Z., “Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis”, *Energy*, Vol. 22, No. 11, pp. 1059-1069, November 1997.

[6] Akdi, Y., Gölveren, E., Okkaoğlu, Y., “Daily electrical energy consumption: Periodicity, harmonic regression method and forecasting”, *Energy*, Vol. 191, January 2020.

[7] Gori, F., Takanen, C., Forecast of energy consumption of industry and household & services in Italy”, *International Journal of Heat and Technology*, Vol. 22, No. 2, pp. 115-121, January 2004.

[8] Azadeh, A., Ghaderi, S.F., Sohrabkhani, S., “Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors”, *Energy Conversion and Management*, Vol. 49, No. 8, pp. 2272-2278, August 2008.

[9] Amber, K.P., Aslam, M.W., Hussain, S.K., “Electricity consumption forecasting models for administration buildings of the UK higher education sector”, *Energy and Buildings*, Vol. 90, pp. 127-136,

March 2015.

[10] http://www.uge.rs/

[11] Milojković, J.B., Litovski, V.B., “Short term forecasting in Electronics”, *International Journal of Electronics*, Vol. 98, No. 2, pp. 161-172, February 2011.

[12] Milojković, J, Litovski, V.B., “Dynamic Short-Term Forecasting of Electricity Load Using Feed-Forward ANNs“, *Int. J. of Engineering Intelligent Systems for Electrical Engineering and Communication*, Vol. 17, No. 1, pp. 38-48, March 2009.

Analysis of existing methods for detection of source of harmonic pollution

Dejan Stevanović, Miona Andrejević Stošović, Predrag Petković

Abstract – In this paper we will give an overview of the methods used for the detection of sources of harmonic pollution at power grid. The aim is to compare the original method based on the use of distortion power proposed by the authors with other known methods. The comparison will be done by simulation using benchmark examples. The obtained simulation results will confirm that the value of the distortion power can be effectively used to detect the source of harmonic pollution at the point of common coupling (PCC).

Keywords – Harmonic source detection, power meters.

I. INTRODUCTION

The last few decades are characterized by a wide use of smart electronic appliances and a massive use of the electronic control systems in industrial production. Most of the electronic equipment needs DC supply for operation, thus making the AC to DC converters become the dominant loads at the power grid. The non-linear nature of these converters introduces a huge amount of harmonics. It is well known that any deviation of the line voltage produces serious problems [1, 2]. The continuous rise of the number of non-linear loads - gives rise to all problems related to the influence of harmonics. To face this regulations, the limit of the allowed amount of each harmonic is created. Two widely used standards in this area are the IEEE 519-1992 and IEC/EN61000-3-2 [1].

The standard IEC/EN61000-3-2 is in use in the European Union since 2001. It restricts the value of input current distortion up to the fortieth harmonic in electrical household appliances. It considers consumers up to 16A per phase, with the nominal voltage from 240V up to 415V.

So far, there are no regulations that specify what happens if a customer exceeds the allowed amount of harmonic pollution. However, in order for the utility to carry out any measure one needs a proper and efficient method to detect and quantify the level of distortion of the respective customer (PCC *Point of Common Coupling*). Thereafter, it needs a procedure to discipline the irresponsible customer.

Practically, the utility could apply one of two approaches. The first is to disconnect the harmonic producer. The softer alternative is to charge for extra-losses. From the point of view of the general public, the former is better because it encourages customers to avoid using appliances that pollute the grid. Actually, the main problem in the tax driven regulation is the lack of the measurement tools to detect the dominant source of harmonics producer. One of the suggestions is to implement new electronic power meters able to measure harmonic distortion. However, measuring *Total harmonic distortion* (*THD*) is not sufficient for analysis of the effects of the harmonic polluter [4]. Therefore, the computation of another quantity named *Power Quality Index* (*PQI*) was suggested.

A number of experts deal with the problem of power quality measurement and location of non-linear loads [3-9]. In our previous paper [10- 12] we offered a solution that could be implemented using existing solid-state electronic meters. This we consider to be the main advantage of the proposed idea.

The paper is organized in five parts. The next section reviews existing solutions for detection of non-linear loads on the grid together with known shortcomings. After that we give the basic definition of fundamental power quantities. In the fourth paragraph we compare all metods at real benchmark examples by simulation. The conclusion is given in the fifth section.

II. REVIEW OF EXISTING METHODS FOR HARMONIC SOURCE DETECTION

At present, there are several proposed solutions for harmonic source detection at power grid. All of them can be divided in two groups:

- multi-point methods and
- single-point methods.

The first group of methods is based on distributed and synchronous measurement system and data collection from different points at power grid. This method provides precise and complete information about propagation of harmonics at grid. However its practical realization is difficult because it requires complex and expensive measurement instrumentation. On the other hand, single-point methods are more convenient for implementation, but less precise. Table I summarizes the state of the art of single-point methods that are used for harmonic source detecting or sharing harmonic responsibility between utility and consumer.

DejanStevanović, Miona Andrejević Stošović and Predrag Petković are with University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, Niš, Serbia,
{dejan.stevanovic,miona.andrejevic,predrag.petkovic}@elfak.ni.ac.rs.

TABLE I
THE-STATE-OF-ART FOCUS ON THE HARMONIC SOURCE DETECTION AND SHARING HARMONIC RESPONSIBILITY METHODS

Method (Indices)	**Required Data**	**Aims**
Active Power Direction (APD)	Voltage and current obtained by single point measurement	Harmonic Source Detection
Reactive Power Direction (RPD)		
Nonactive Power Method (NP)		
Harmonic Polluted Ranking (HPR) method	Current obtained by single point measurement	
Critical Impedance Method(CI)	Thevenin's equivalents of utility and consumer sides	Sharing harmonic responsibility between utility and consumer

One of the most common used single-point strategies from Table I (APD) is based on monitoring the sign of the harmonic active power P_H [1, 3, 8]. According to this method a consumer represents source of harmonic pollution just in case when $P_H< 0$. Otherwise, the utility-side pollutes the power grid. This identification of the nonlinear consumer is widespread and has been used in industry for many years [1], [3]. Many manufactures present this possibility as a key feature of their equipment despite the proofs that it is not 100% accurate [3].

RPD proposes tracking the sign of the harmonic reactive power, Q_H. It is complement to APD. Which one will be applied depends on the balance of resistances and reactances of the loads [1], [3]. However, determining the character of impedance is, in most cases, a difficult problem.

NP is based on comparison of three non-active power components [6], [7] that complicates DSP software of power meter.

The authors of HPR method introduce new power quality index to monitor the affect of each nonlinear load on a *Point of Common Coupling* (PCC) of a power distribution system. It is defined as a product of *Load Composition Rate* and THD. Both parameters are derived from the load current waveform using a complex *Reduced Multivariate Polynomial* model [4].

An alternative group of methods able to share responsibility between harmonic sources relies on measurement of impedance at the grid and the customer side (CI). Theoretically, the method is reliable. Therefore it appears in numerous variations [9]. However this method cannot be implemented easily. The practical problem is that one can determine the impedances only if he intrudes the system.

The main drawbacks of aforesaid methods are:

- neither of them gives exact quantitative information of a customer which pollutes the grid;
- they cannot be easy implemented within ordinary electronic power meter, consequently cannot be used for each customer.

Considering all disadvantages of the methods proposed in the literature, the authors of this paper tried to fix and eliminate them in their own method. [10-12]

III. THE DEFINITIONS OF THE FUNDAMENTAL QUANTITIES

Traditional power system characterization quantities such as RMS values of current and voltage, power (active, reactive, apparent) are defined for ideal sinusoidal conditions. However, in the presence of non-linear loads, these definitions need correction. The instantaneous values of a quantity rich with harmonics (voltage or current) can be expressed as:

$$x(t) = \sum_{h=1}^{M} X_h \sin(\omega_h t + \alpha_h), \tag{1}$$

where h is the number of the harmonic, M denotes the highest harmonic, while X_h, ω_h *and* α_h, represent amplitude, frequency and phase angle of the h-th harmonic.

The RMS value of the signal expressed by (1) is defined as:

$$X_{\mathrm{RMS}} = \sqrt{\sum_{h=1}^{M} X_{\mathrm{RMSh}}^2}\,, \tag{2}$$

where X_{RMSh} is the RMS values of the h-th harmonic.

Product of the voltage and current having the same harmonic frequency gives the harmonic power. Total active power is defined as:

$$P = \sum_{h=1}^{M} V_{\mathrm{RMS}_h} I_{\mathrm{RMS}_h} \cos(\theta_h) = P_1 + P_{\mathrm{H}}, \tag{3}$$

where θ_h denotes phase angle between voltage and current, while P_1 and P_H represent *fundamental active power* and *harmonic reactive power*, respectively.

According to Budeanu [1, 8] reactive power is defined as:

$$Q_{\mathrm{B}} = \sum_{h=1}^{M} V_{\mathrm{RMS}_h} I_{\mathrm{RMS}_h} \sin(\theta_h) = Q_1 + Q_{\mathrm{H}}\,, \tag{4}$$

where, similarly to (3), Q_1 and Q_H denote *fundamental reactive power* and *harmonic reactive power*, respectively.

By analyzing the equations for calculating active (3) and reactive power (4), it can be concluded that only harmonics of current and voltage of the same order affect the total value of active and reactive power. The contribution of the harmonic component of active and reactive power to their total value is small, usually less than 3% [1]. The harmonics influence the most to the effective value of current and voltage and therefore on apparent power, which is calculated as the product of the effect of value of current and voltage:

$$S = I_{RMS} * V_{RMS} = \sqrt{\sum_{h=1}^{M} V^2{}_{RMSh}} * \sqrt{\sum_{h=1}^{M} I^2{}_{RMSh}}\,. \tag{5}$$

Using (3), (4) and (5) one gets the inequality:

$$S^2 > P^2 + Q_B{}^2 \tag{6}$$

However, for the sinusoidal case $S^2=P^2+Q_B{}^2$. Consequently, it is quite clear that a difference exists due to harmonic distortion. Following the logic about defining active and reactive power, Budeanu introduced the term *distortion power* D_B that quantifies the discrepancy within the inequality (6):

$$S^2 = P^2 + Q_B{}^2 + D_B{}^2 . \tag{7}$$

The essence of this revision is contained in the fact that in absence of harmonics, $D_B=0$ and $S^2=P^2+Q_B{}^2$. Therfore, in our previous paper we suggested method for detection source of harmonic pollution at power grid based on measuring power distortion [9-12] calculated by using equation (8).

$$D_B = \sqrt{S^2 - P^2 - Q_B{}^2} . \tag{8}$$

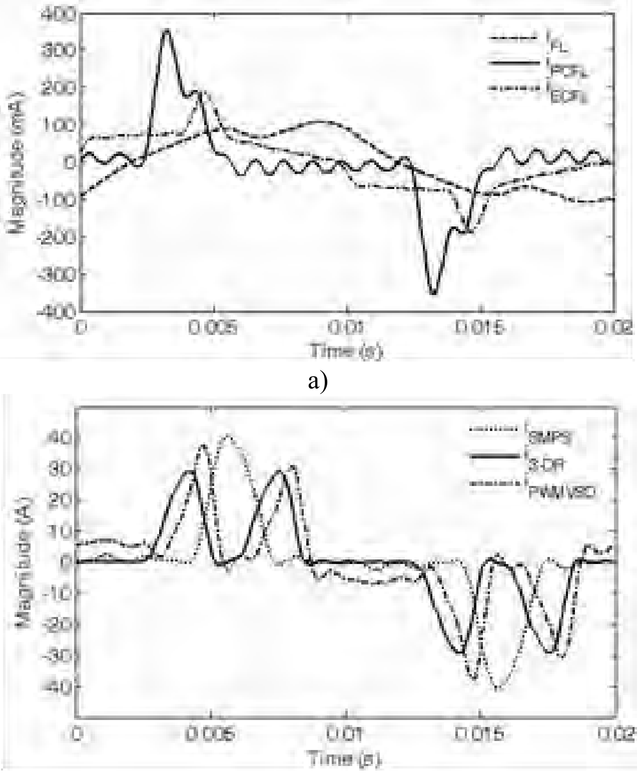
IV. COMPARISON OF EXISTING METHODS FOR DETECTION SOURCE OF HARMONICS POLLUTION BY SIMULATION

The comparison between the methods is done by simulation. Based on the equations (2), (3), (4), and (8), a Matlab script was generated. In the simulation we tried to make realistic environment, so we modelled voltage polluted with 3rd, 5th and 7th harmonic. Aiming to get sinewave with flattened top and bottom peaks, we added 3% of 3rd harmonic, 2.5% of the 5th and the 7th harmonic in respect to the value of the fundamental harmonic V_{RMS1} = 230V. The following ten cases of different load types were considered:

a) Incandescent light bulb (ILB)
b) Fluorescent lamp (FL)
c) Phillips Compact Fluorescent Lamp (PCFL)
d) EcoBulb Compact Fluorescent Lamp (ECFL)
e) 6-pulse 3-phase diode rectifier dc power supply (3-DR)
f) 6-pulse switched-mode power supply (SMPS)
g) 6-pulse PWM controlled variable speed drive (PWM VSD)

The aforementioned loads were not selected accidentally. They were chosen because they can serve here as benchmarks to the power electronic community, since measured data for these loads were already published in [1, 13]. Fig. 1.a illustrates the waveforms of the currents for FL, PCFL, ECFL, while Fig. 1.b presents waveforms of the currents through SMPS, 3-DR and PWM VSD. Among these simulated loads, only ILB has linear characteristic, because the current tracks the waveform of the voltage. Other loads are nonlinear introducing new harmonics into current.

In the Table II all the simulation results are collected and compared. The quantities in the Table are labelled according to the notation given in equations (2)-(8). The additional parameters are THD_V and THD_I, standing for total harmonic distortion of voltage and current waveforms, respectively. THD_I has been lately suggested as a parameter for distortion power detection and quantification [13]. Further, we obtain D_{IEEE} and D_1 when we replace Q_B with Q_{IEEE} and Q_1 in (8), respectively. Besides, we presented D_I as one more form for distortion power estimation. It defines distortion power as a product of THD_I and S_1 (fundamental apparent power), according to [14].



a)

b)

Fig. 1. Current waveforms for a) FL, PCFL, ECFL and SMPS b) SMPS, 3-DR and PWM VSD

We can also notice from the Table II, that for linear loads active power, P, equals to the apparent power, S. Using (12), we can easily calculate that the value of distortion power is equal to zero no matter what definition of Q is used. However, the value of the quantity D_I is quite inaccurate in the case of linear resistive load (ILB), since nonlinearities in nonlinear voltage supply are mapped into the nonlinearities of the current, making quantity D_I inaccurate. We can thus conclude that this power quantity is not reliable to be used for identification of the source of harmonic pollution.

When taking nonlinear loads into account, all four methods for determining D offer approximate, comparative results. From the given values we can conclude three important issues. First, all the values for distortion power (D_B, D_{IEEE} and D_1) are greater than 0 for all nonlinear loads, thus confirming the suggested method. Second, all three definitions of D (D_B, D_{IEEE} and D_1) give very similar results. Slight differences in definitions of Q map into slight differences in D, but this disagreement can be compensated with convenient billing policy. The last, but maybe the most important conclusion is that the amount of distortion power is not negligible, because in some cases it is of the same order of magnitude as the active power. Disregarding D could cause considerable losses at power grid [2].

TABLE II
SIMULATION RESULTS

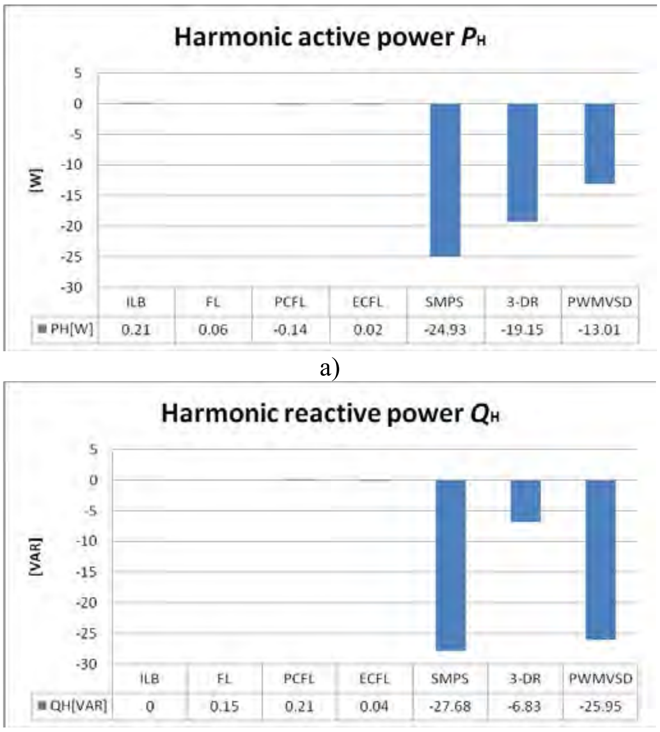
	ILB	**FL**	**PCFL**	**ECFL**	**SMPS**	**3-DR**	**PWMVSD**
I_{RMS}[A]	0.42	0.10	0.13	0.09	14.84	13.53	14.23
V_{RMS}[V]	230.25	230.25	230.25	230.25	230.25	230.25	230.25
P_1[W]	96.60	17.31	16.09	18.59	2249.74	2251.39	2300.00
P_H[W]	0.21	0.06	-0.14	0.02	-24.93	-19.15	-13.01
P[W]	96.81	17.36	15.95	18.61	2224.81	2232.25	2286.99
Q_1[VAR]	0.00	15.15	-10.06	-6.04	478.20	470.34	0.00
Q_H[VAR]	0.00	0.15	0.21	0.04	-27.68	-6.83	-25.95
Q_B[VAR]	0.00	15.30	-9.84	-6.00	450.52	463.51	-25.95
Q_{IEEE}[VAR]	0.00	15.15	10.07	6.04	481.42	472.68	41.92
S[VA]	96.81	23.62	29.71	20.91	3416.27	3114.89	3276.55
THD_V[%]	4.64	4.64	4.64	4.64	4.64	4.64	4.64
THD_I[%]	4.64	22.85	120.28	37.66	109.61	91.12	101.25
D_B[VAR]	0.00	4.72	23.06	7.41	2553.07	2122.45	2346.23
D_{IEEE}[VAR]	0.00	5.18	22.96	7.38	2547.42	2120.42	2346.00
D_1[VAR]	0.00	5.18	22.96	7.38	2548.03	2120.94	2346.37
D_I[VAR]	4.48	5.26	22.82	7.36	2521.08	2095.65	2328.69
Q*sq*[VAR]	0.00	15.88	21.99	7.44	2139.69	1949.59	1864.67
Q*x*[VAR]	0.00	16.01	24.78	9.53	2511.77	2109.96	2314.21
N[VAR]	0.00	16.01	25.07	9.54	2592.51	2172.47	2346.37

TABLE III
COMPARISON OF OUR METHOD (M) WITH METHOD PUBLISHED AT [6-7] [M1 AND M2]

		ILB	**FL**	**PCFL**	**ECFL**	**SMPS**	**3-DR**	**PWM VSD**
Q_1[VAR]		0.00	15.15	-10.06	-6.04	478.20	470.34	0.00
Q_{SH}[VAR]		0.00	15.88	21.99	7.44	2139.69	1949.59	1864.67
Q_X[VAR]		0.00	16.01	24.78	9.53	2511.77	2109.96	2314.21
Q_F=N[VAR]		0.00	16.01	25.07	9.54	2592.51	2172.47	2346.37
Detected as a source of harmonics pollution	M1	none	consumer	consumer	consumer	consumer	consumer	consumer
	M2	none	consumer	consumer	consumer	consumer	consumer	consumer
	M	none	consumer	consumer	consumer	consumer	consumer	consumer

Table III presents results of comparison between the method proposed in [10-12] (denoted in Table III as M) and two alternative methods published in [6] (denoted in the Table as M1) and in [7] (denoted in the Table as M2). In case of linear resistive load (ILB) all methods registered no harmonic pollution. That is because either Q or D do not exist on linear resistive load. In this case harmonics in current can appear only if they already exist in voltage. The utility is responsible to keep voltage distortion at low level defined by standards IEEE 519-1992 and IEC/EN61000-3-2. Besides, the utility has interest to keep non-distorted voltage in order to protect own equipment from damages. Method M1 specifies the customer as distortion producer if value of *Sharon reactive power* (Q_{Sh}) is closer to *Fryze reactive power* (Q_F) than to Q_1. If it is not the case the inducer is at utility side. In method M2 we replaced *Sharon reactive power* (Q_{SH}) with *fictitious reactive power* (Q_X) and used same principe.

The common drawback of methods M1 and M2 is the following. First, for loads with small nominal power (in our case FL), the differences between the three non-active powers are small. Therefore, for precise identification, the instrument with high accuracy is required. Second, the value of *fictitious and Sharon's* reactive power can only be positive and can create confusion when the load has capacitive nature. This is common case for CFL. Finally, the method requires measurement of non-active power defined in three different ways while our method is based only on one value. Results presented in Table III confirm that our method undoubtedly detects the inducer of harmonic pollution.



a)

b)

Fig 2. Graphical representation of obtained value for simulated load a) harmonic active power b) harmonic reactive power

In addition we compared our approach with the method based on observing the sign of active harmonic power P_H [1, 3, 8]. According to this method, when P_H>, 0 the source of harmonic pollution is located upstream from the metering point (the utility). Otherwise, the source of harmonic pollution is located downstream (consumer). Fig. 2.a shows the graphical presentation of obtained value for P_H for the same set of loads as given in the Table II. In all cases (except ILB) we have nonlinear loads and the inducer of harmonics is at the customer side. However, the sign of harmonic active power changes. It is positive for ILB, FL, ECFL, but negative for PCFL, SMPS, PWM VSD and 3-DR. Consequently this shows that this method is not reliable [3]. Alternatively, Fig 2.b presents harmonic reactive power for the same loads. Obviously none of these parameters can be used for precise location of harmonic pollution source, as well. Namely, in all the cases the inducer is the customer while the sign of harmonic reactive power is negative for SMPS, PWM VSD and 3-DR and positive for other loads.

V. CONCLUSION

The simulation results of different types of loads, presented in Table III, clearly show that the value of the distortion power can be successfully used to detect source of harmonic distortion. Moreover, the method based on measuring the distortion power successfully identifies and quantifies the degree of harmonic distortion introduced by each consumer at PCC. The method is compared on real loads with recently published methods based on APD, RPD and NP methods.

ACKNOWLEDGEMENT

This work has been partly funded by the Serbian Ministry of Science, Education and Technological development under the contract No. TR32004.

REFERENCES

[1] Wakileh. J. G., Power Systems Harmonics, *Springer*, 2001

[2]Stevanović D., Petković P., "Utility needs smarter power meters in order to reduce economic losses", Facta Universitatis Series: Electronics and Energetics Vol. 28, No 3, pp. 407-421, doi:10.2298/FUEE1503407S, Sep. 2015.

[3] Xu W., Liu X., Liu Y.: An investigation on the validity of power-direction method for harmonic source determination, *IEEE Trans. Power Delivery*, vol. 18, no.1, January 2003. pp. 214-218

[4] Lee S., Park J. W., New power quality index in a distribution power system by using RMP model, *IEEE Tran. On Industrial Applications*, vol. 46, n.3, May/June 2010, pp. 1204 – 1211

[5]Chun L., Xu W., and Tayjasanant T., A critical impedance based method for identifying harmonic sources, *IEEE Trans. Power Del.*, vol. 19, no. 2, Apr. 2004 pp. 671–678.

[6] P. V. Barbaro, A. Cataliotti, V. Cosentino, and S. Nuccio, "A novel approach based on nonactive powers for the identification of disturbing loads in power systems," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1782–1789, Jul. 2007.

[7] Cataliotti A., Cosentino V., A single-point approach based on IEEE 1459-2000 for the identification of detection of prevailing harmonic sources in distorted three phase power systems, *Proc. of the Metrology and Measurement Systems*, 2009.

[8]Emanuel A. E., Summary of IEEE standard 1459: definitions for the measurement of electric power

quantities under sinusoidal, nonsinusoidal, balanced, or unbalanced conditions, *IEEE Tran. On Industrial Applications*, Vol. 40, No. 3, May 2004 pp. 869–876.

[9]Chun L., Xu W., and Tayjasanant T. (2004). A critical impedance based method for identifying harmonic sources. *IEEE Trans. Power Del.,* vol. 19, No. 2, 671–678, 2004.

[10]Stevanović D., Petković P,"The Efficient Technique for Harmonic Sources Detection at Power Grid", Przegląd Elektrotechniczny, pp. 196-199, 2012

[11]Stevanović D., Petković P.: "A single-point method based on distortion power for the detection of harmonic sources in power system". *Metrol. Meas. Syst.*, Vol. 21, No 1, pp.3-14, March, 2013.

[12]Petković P., Stevanović D., "Detection of power grid harmonic pollution sources based on upgraded power meters," *Journal of Electrical Engineering*, vol. 65, no. 3, pp. 163-168, 2014.

[13]IEEE Task Force on Harmonics Modeling and Simulation, "Test Systems for Harmonics Modeling and Simulation IEEE Trans", Power Del., vol. 14, No. 2, 579–587, April 1999.

[14]IEEE Std 1459-2010, "IEEE Standard Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Nonsinusoidal, Balanced, Or Unbalanced Conditions", 2010.

Library of Combinational Logic Cells Resistant to Side Channel Attacks

Milena Stanojlović Mirković, Miljana Milić, Dejan Mirković and Vančo Litovski

Abstract – This paper presents an overview of research in the field of hardware cryptography. An effort was made to develop the library of CMOS cells that are resistant to Side Channel Attack. The No Short-circuit current Dynamic Differential Logic method is implemented. Characteristics of encrypted cells are compared with the standard, (not encrypted), cells under various operational conditions in order to prove the SCA resistance. Designed encrypted cells represent the basis of developing more complex crypto system and increase its overall security.

Keywords – CMOS, IC design, SCA, Cryptography, NSDDL method.

I. INTRODUCTION

The content of an encrypted data in digital systems is protected by utilizing specific algorithms which should harden obstruct the decrypting. The protection is usually based on application of complex keys which require hacker to apply a large number of combinations in order to break them. The longer time it takes for trying of each bits combination, the protection is better. However, time for key breaking may be significantly reduced if, besides logical states, other characteristics of the signal are observed. Usually, analysis of power consumption i.e. power supply current time profile are used for this purpose. Every unauthorized collecting of such information about crypto system behaviour is referred to as Side Channel Attack (SCA) [1-2].

The main source of information about the behaviour of a circuit is the circuit activity expressed through the change of the supply current. Observing changes in the supply current (I_{DD}) and correlating them with known input vector can be used as valuable information for breaking the coding key. Physical background for this approach lies in the fact that an abrupt change of the I_{DD} in a CMOS digital circuit occurs only during transition of a logic state. For example, during the 0-to-1 transition of the signal an additional charge is needed to load capacitances. Besides, some "short-circuit" current flows when PMOS and NMOS transistors are turned on simultaneously. During this transition, I_{DD} changes produce electromagnetic field variations which the attackers may detect using special probes.

The encrypted library of CMOS cells, that are resistant to SCA attacks, is developed in LEDA Laboratory at the Faculty of Electronic engineering, University of Nis. The key achievement in this study is the development of the part of this library which applies only to combinational logic.

The SCA resistance is measured by the degree of the information hiddenness and it is larger if the correlation between the I_{DD} and the circuit behaviour is suppressed. For the design of encrypted cells, the No Short circuit current Dynamic Differential Logic method is adopted [3].

This paper is organized as follows: the section II presents the basics of the NSDDL method; the section III presents the design methodology of combinational encrypted cells; while in the section IV considers the difference between power supply currents (I_{DD}) for standard and encrypted cells. The final section summarizes key contributions of this research.

II. NSDDL METHOD

The encrypted cells' functioning exploits the idea that each combination of input signals results in the same power consumption. This can be realized 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 hiding the true function of the cell is achieved.

The NSDDL method requires three different operation phases. During the first, precharge, phase both outputs (true and false) of all logic cells are driven to high logic level. In the second phase, known as the evaluation phase, the desired value is set at the true output and the complementary value is established at the false output. The third phase is named discharged because all outputs achieve low logic level. These phases are illustrated in Fig. 1.

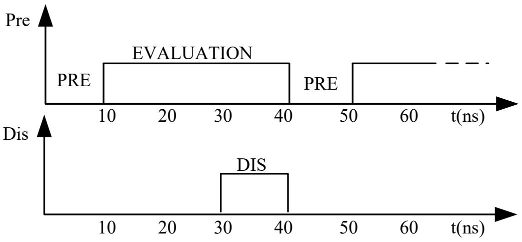


Fig. 1 Waveforms of control signals for the Dnor cell

Milena Stanojlović Mirković, Miljana Milić, Dejan Mirković and Vančo Litovski are with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail:
stanojlovim@gmail.com, milana.milic@elfak.ni.ac.rs,
dejan.mirkovic@elfak.ni.ac.rs vanco.litovski@elfak.ni.ac.rs.

The advantage of this method compared to other popular solutions, like WDDL [4-5], is its immunity to imbalance loads at true and false outputs. This is achieved by using a dynamic NOR circuit (DNOR) which minimizes the impact of short circuit currents in the CMOS circuit. It is an integral part of the control logic and NSDDL cells. Figure 2 illustrates this circuitry.

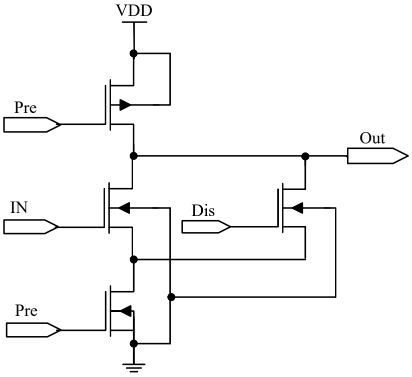


Fig. 2. Dnor cell

III. NSDDL CELLS - COMBINATIONAL LOGIC

All logic cells resistant to SCA are designed in CMOS technology TSMC 0.35um. The idea of the NSDDL methodology requires that the responses of each input signal combination consume the same power. This is possible with the doubled hardware that contains a cell with the complementary properties. Therefore, any input combination will imply the transitions on both true and false outputs.

A. Designing an inverter / buffer cell (INV / BUFF)

Due to the complementary design an inverter (INV) cell will behave as a buffer (BUFF) at the false output[6-7]. However, the simplicity of the logic function requires a different design approach comparing to the general NSDDL concept. Obviously, complementary signals A and notA are required at true and false inputs At the output of the cell two signals denoted with OT (Output True) and OF (Output False) are generated. The signal OT represents the true output signal (inverted A signal), while OF represents the false one, as depicted in Fig. 3.

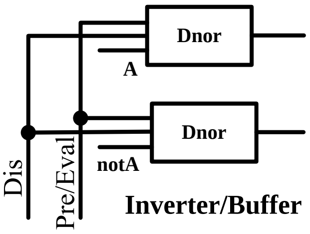
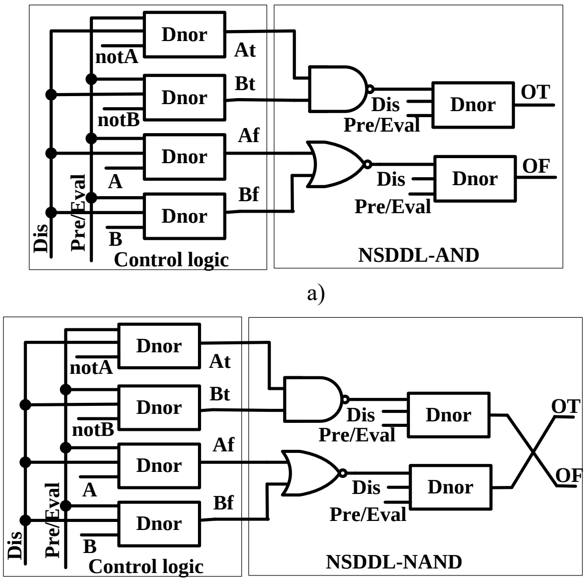


Fig. 3. Block diagram of the NSDDL INV/BUFF cell

In the case of a BUFF cell, OT and OF signals, meaning, are swapped. At the same time this cell represents a part of the logic that controls others NSDDL cells. This practically means that the inverting function in the NSDDL logic is obtained with crisscrossing of true and false outputs.

B. Design of two - input AND/NAND/OR/NOR cell

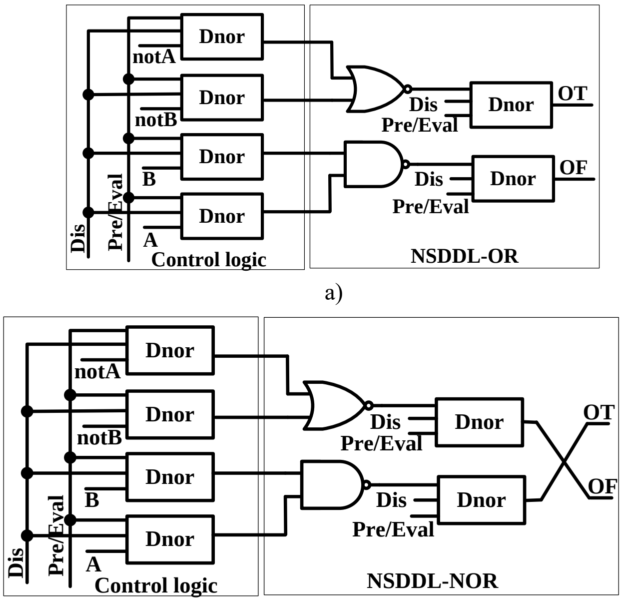
Consider mutually complementary NAND and NOR cells, as a part of a SCA resistant structure. Block schemes of AND/NAND and OR/NOR NSDDL, SCA resistant cells are presented in figures 4 and 5, respectively [8].



a)

b)

Fig. 4. Block diagrams of a) NSDDL AND cell; b) NSDDL NAND cell



a)

b)

Fig. 5. Block diagrams of a) NSDDL OR cell; b) NSDDL NOR cell

Besides, from figures 4 and 5 can be seen that cells are excited with mutually complement input signals A and notA i.e. B and notB. Using de Morgan rules it can be shown that with simple input signal (A, notA, B, notB).

It is obvious that the same hardware structure implement AND, NAND, OR and NOR functions. That is why this structure is referred to as the AND/NAND/ OR/NOR SCA resistant cell. It is important to notice that all functions are implemented using negative logic circuits with negative logic (NAND and NOR) which can be easily implemented using CMOS technology.

The DNOR circuit represents the basic element for all SCA resistant cells within the NSDDL technique. It provides inverting function when transforming from standard to the NSDDL logic.

C. Design of the XOR/XNOR cell

This cell consist of two NSDDL AND cells and one NSDDL OR cell. The block diagram of the XOR/XNOR SCA resistant cell is presented on Fig. 6.

The control logic generates the input signals for this cell. Since that the input (true and false) signals, are complementary, the same structure provides both functions: the XOR function at the true output (OT) and XNOR function at the false output (OF).

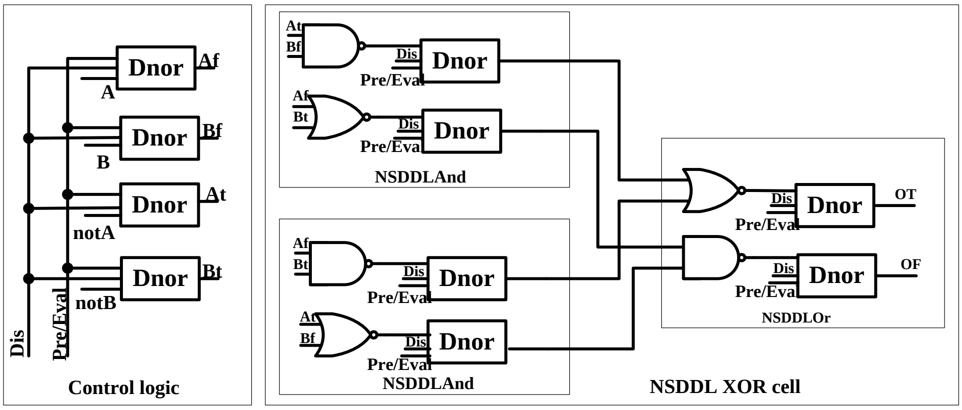


Fig. 6. Block scheme of NSDDL XOR SCA resistance cell

IV. RESULTS

This section will present comparative results of the simulations for standard and encrypted cells. For standard cells one can expect strong correlation between energy required for the particular transition and the 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 hide the cell's operation regarding I_{DD}. Therefore, they should provide a minimal correlation between the stimulus signals and the I_{DD} value. To quantify the resistivity to SCA we have adopted a measure based on the calculated integral of the consumed power over time (energy) [11-12].

$$E = V_{DD} \cdot \int_{0}^{T} i_{DD}(t) \cdot dt, \qquad (1)$$

In order to provide better insight into the behavior of every cell from the simulation results we have derived the following parameters: standard deviation (σ) normalized standard deviation in respect to average energy (E_{avg}) *(NSD)*. As a measure of the SCA resistance we have considered normalized standard deviation according to (2).

As for the INV and BUFF cells, only two transitions of input signal are possible. Therefore, it makes no sense to statistically process only two data values. That is why only relative average energy consumption difference is observed, and it is denoted with *δE* in Table I. One can notice that the uniformity of the cell's energy consumption is increased. This means that the presented cell has the significant SCA resistance. Also, when the *δE* parameter is observed, there is a 98.9 times increase in the SCA resistivity.

$$NSD = 100 \cdot \frac{\sigma}{E_{avg}} [\%]. \qquad (2)$$

TABLE I RESULTS OF STANDARD AND NSDDL INVERTER CELLS

A	notA	$E_{standard}$ [J]	E_{NSDDL} [J]
↑	↓	2.48136E-13	1.32838E-12
↓	↑	2.21651E-13	1.3299E-12
E_{avg} [J]		2.34894-13	1.3291E-12
δE [%]		11.275	0.114

Results obtained for standard AND, NAND, OR and

NOR, and NSDDL AND/NAND/OR/NOR cells are compared and presented in Table II. Also, the same table contain results for standard XOR and XNOR cells and the NSDDL XOR/XNOR cell.

It is interesting to track how the property defined as resistance to the SCA is being transferred from lower to higher hierarchical design level. With this in mind, we have performed a similar set of simulations for NSDDL XOR/XNOR as for the NSDDL AND/NAND/OR/NOR cell.

TABLE II RESULTS OF STANDARD AND NSDDL CELLS

	E_{avg} [J]	δE [%]	σ [fJ]	NSD [%]
AND	4.83E-13	210.15	405.4	83.91
NAND	4.11E-13	196.98	337.7	82.23
OR	3.62E-13	222.05	310.3	85.64
NOR	2.94E-13	202.67	243.1	82.59
NSDDL AND/NAND/OR/NOR	2.77E-12	2.81	24.31	0.87
XOR	3.90E-13	63.64	91.77	23.51
XNOR	3.25E-13	131.76	154.18	47.43
NSDDL XOR/XNOR	6.24E-12	3.53	56.58	0.907

Dynamic energy consumption, as mentioned before, is expressed through the integral of the I_{DD} over time during one cycle of the input signal change. This cycle is the same for standard cells as for NSDDL cells, in all three operational phases. As before, relative, average consumed energy difference, standard deviation and normalized standard deviation are taken as measures of the SCA resistance. Parameters are denoted with δE, σ and *NSD*, and are given in columns one, two and three, respectively.

The *NSD* parameter that is less than 1% (0.87%) for the AND/NAND/OR/NOR2 NSDDL cell remained almost the same. Although slightly increased to the value of 0.91%, it is still less than 1%, which qualifies this cell as the SCA resistant one. Actually, the *NSD* has increased for 4.6% in respect to the AND/NAND/OR/NOR NSDDL cell. The total improvement of the resistivity to SCA in comparison with standard cells overcomes 2500% for the XOR and 5000% for the XNOR cell.

Figures 7 and 8 show trends of the energy consumption for:

- the four unprotected standard cells (AND, NAND, OR, NOR) and encrypted NSDDL AND/NAND/OR/NOR cell (Fig.7)
- two unprotected standard cells (XOR and XNOR) and the encrypted NSDDL XOR/ XNOR cell (Fig. 8)

The input signal combinations are given in the horizontal axis labels, A and B, where the symbols "↑" and "↓" denote rising and falling transitions, respectively. The ordinate label denotes obtained energy levels for those cells.

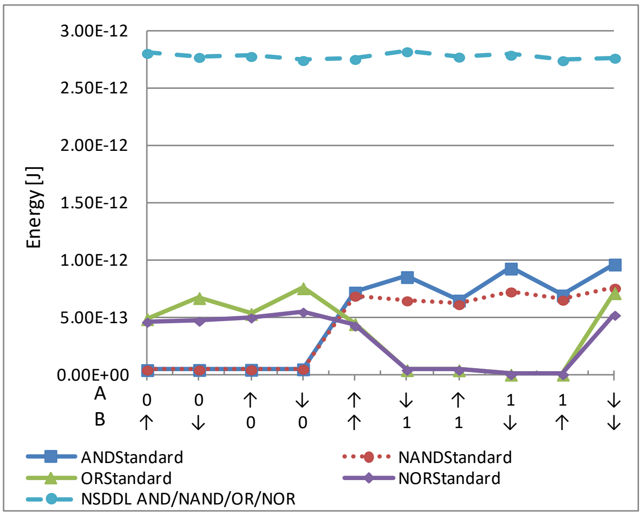


Fig. 7. Energy consumption during ten cycles of input signals change for the unprotected standard cells and the encrypted NSDDL AND/NAND/OR/NOR cell

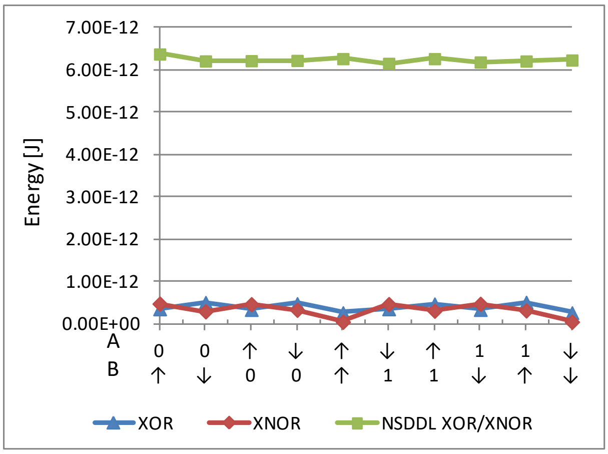


Fig. 8. Energy consumption during ten cycles of input signals change for the unprotected standard cells and the encrypted NSDDL XOR/XNOR cell

V. CONCLUSION

This paper presents the part of library logic cells resistant to side channel attacks (combinational logic) designed by the rules of the NSDDL method. This method is characterized by the duplicated implementation of hardware that generates true and false outputs. The false output has the same function as the inverted true output. The basic idea is to mask the correlation between the supply current and the activity of the cell. This is possible to obtain with doubled input signals Three-phase clock signal guarantees that all outputs will start from the high logic level during the pre-charging phase, and will end with the low logic level during the third phase. The cell performs the desired logic function in the second operating phase. Then the true output takes the desired output state and, while, simultaneously the false output changes with the opposite transition. The energies required for output transition under different combination of input signal were considered as a measure of the SCA resistance. The cell is resistant if all changes at its outputs require the same energy. As a measure of a cell SCA resistance we have considered the normalized standard deviation (NSD). All designed cells have showed a good resistivity to SCA when this parameter is observed.

ACKNOWLEDGEMENT

This work was partially funded by Serbian Ministry of Education, Science and Technological Development under contract No. TR32004.

REFERENCES

[1] Lomné, A. Dehaboui, P. Maurine, L. Torres, M. Robert, "Side Channel Attack"', in B. Badrignans, J. L. Danger, V. Fischer, G. Gogniat, L. Torres, "*Security Trends for FPGA*", Springer Netherlands 2011, pp. 47-72

[2] Stanojlovic M., Petkovic P., " Strategies against side-channel-attack" in *Proceedings of the Small Systems Simulation Symposium*, Niš, Serbia, pp. 86–89, 2010.

[3] Quan J., Bai G., "A new method to reduce the side-channel leakage caused by unbalanced capacitances of differential interconnections in dual rail logic styles", *Sixth International Conference on Information Technology*: New Generations, pp. 58-6, 2009.

[4] Tiri K., Verbauwhede I., "Place and Route for Secure Standard Cell Design", CARDIS'04, pp. 143-158, 2004.

[5] Velegalati R., "Securing Light Weight Cryptographic Implementations on FPGAs Using Dual Rail with Pre-Charge Logic", PhD Thesis, George Mason University, Fairfax, VA, 2009.

[6] Stanojlović M**.**, Milovanović, D. "Simulacija defekata osnovnog kola kontrolne logike NSDDL metoda", *Zbornik LV konferencije ETRAN,* Banja Vrućica, Bosna i Hercegovina, 06.06.-09.06., EL 4.5, 2011.

[7] Stanojlović M., "D flip-flop ćelija otporna na bočne napade analizom struje napajanja", *Zbornik LVII konferencije ETRAN*, Zlatibor, 03.06.-06.06., EL3.3, 2013.

[8] Stanojlović M**.**, Litovski V., Petković P., "Testiranje standardne AND ćelije otporne na bočne napade", *Zbornik LVIII konferencije ETRAN*, Vrnjačka Banja, 02.06.-05.06., EL2.5, 2014.

[9] Stanojlović Mirković M**.**, Litovski, V., Petković P., Milovanović, D., "Faults Simulations in XOR/XNOR Cell Resistant to Side Channel Attacks", *X Simposium on Industrial Electronics INDEL*, Banja Luka (Bosnia and Herzegovina), pp. 83-88, 6-8th November, 2014.

[10] Stanojlović M., Petković P., "Resistance of XOR/XNOR NSDDL Cell to Side Channel Attack", *Proceedings of Small Systems Simulation Symposium*, Niš, Serbia, 12.02.-14.02., pp. 141-144, 2012.

[11] Monteiro C., Takahashi Y., Sekine T., "Charge-sharing symmetric adiabatic logic in countermeasure against power analysis attacks at cell level", *Microelectronics Journal, Elsevier*, vol. 44, no. 6, pp. 496-503, 2013. doi.org/10.1016/j.mejo.2013.04.003

[12] Wang P., Zhang Y., Zhang X., "Design of two-phase SABL flip-flop for resistant DPA attacks", *Chinese Journal of Electronics*, vol. 22, no.4, pp.833-837, 2013.

Unified Equation Formulation for Electronic and Electrical Circuits Analysis

Marko A. Dimitrijević, Vančo B. Litovski and Ljupco Kocarev

Abstract - In an attempt to bring together simulation tools and particularly equation formulation methods of power electrical systems analysis and electronic circuit analysis we were studying and contrasted to each other the properties of Load Flow Analysis (LFA) and the Modified Nodal Analysis (MNA). Based on comparisons of their implementations presented in textbooks we came to the conclusion that MNA is favourable as compared to LFA from several points of view such as comprehension, simplicity and universality.

Keywords - Circuit analysis, Power systems simulation.

I. INTRODUCTION

After Tesla's introduction of alternating current and poly-phase systems at the end of the nineteenth and the beginning of the twentieth century, electrical engineering and electricity as such became and still is one of the most important vehicles of the development of human society. However, over time the discipline diverged in two directions: power electrical systems and electronics. For a long time, these separate research communities operated in independent silos. It was only recently that electronics and ICT became necessary for the modernization of power production, distribution, and consumption which has become known as the smart grid. On the other hand, it was only recently that electronics and ICT became one of the biggest consumers of electricity, and as a secondary consequence through the necessary power converters, begun to seriously threaten the quality of the delivered power [1].

Similar separation and re-convergence was to befall the design tools developed for these two trades, including simulation software. Today, modern power electrical system design literature covers subjects that are also electronics oriented [2]. However, simulation tools being developed as part of the power system design subsystems are, of course, not able to reproduce electronic components down to transistor level. On the other hand, simulation tools developed as part of electronic and ICT system design are not able to reproduce various phenomena specific to power generation, transmission and distribution systems. Given the demands that modelling the modern smart grid places on both domains, researchers should extend the capabilities of existing tools such that both worlds are covered [2]. This paper is an attempt to help achieving this goal.

In fact, this is an effort to go deeper into the subject of simulation by facing one against the other the equation formulation methods underlying the simulation programs coming from the two domains. In that way, the intention is to promote better mutual understanding between communities and facilitate convergence toward an omnipotent simulation tool.

As a vehicle of the comparison and the proposal, the examples from two modern text books of power electrical and smart grid systems are used. After repeating part of the example is given in one of them, its deficiencies are exposed and a solution, which is not only successful and effective but also easy to understand and to teach new generations, is proposed.

II. LOAD FLOW ANALYSIS

In the book [3] while introducing the theory of Load Flow Analysis (LFA), on page 396, the following "Example 7.2" is given (a similar, but conceptually identical, example is given in [4] on page 213):

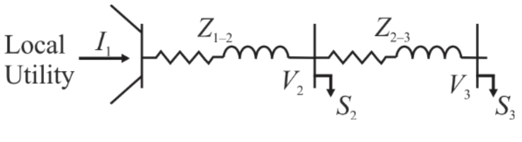


Fig. 1. Distribution feeder (picture 7.2 in [4])

"Example 7.2: Consider a distributed feeder presented in Fig. 7.2. Assume the following:

a. Feeder line impedances, that is Z_{1-2} and Z_{2-3} are known.

b. The active and reactive power consumed, that is S_2 and S_3, by loads are known.

c. The local power grid bus voltage V_1 is known and all data are in per unit."

A solution follows, which we have to repeat due to its

Marko Dimitrijević and Vančo Litovski are with the Faculty of Electronic Engineering, University of Niš – Aleksandra Medvedeva 14, Niš, Sebia E-mail: (marko.dimitrijevic, vanco.litovski) @elfak.ni.ac.rs

Ljupco Kocarev is also with the Macedonian Academy of Science and Art, Bul. Krste Misirkov 2, P.O. Box 428, Skopje, North Macedonia, E-mail: lkocarev@manu.edu.mk

fundamental importance to the knowledge delivered to the power electrical engineering community. Namely, it seems that the way of thinking expressed in this example is omnipresent and, in our opinion, difficult to fit into the fundamental laws of electrical circuit analysis. It is as follows:

“Solution: Let us write the Kirchhoff’s current law for each node (bus) of Fig. 7.2 and assume that the sum of the currents away from the bus is equal to zero. That is, for the buses 1-3, we have

$$\begin{aligned} (v_1 - v_2) y_{12} - I_1 &= 0 \\ (v_1 - v_2) y_{12} + (v_2 - v_3) y_{23} + I_2 &= 0 \\ (v_3 - v_4) y_{23} + I_3 &= 0 \end{aligned} \tag{1}$$

where $y_{12} = 1 / Z_{1-2}$, $y_{23} = 1 / Z_{2-3}$, and

$$I_1 = \left(\frac{S_1}{V_1} \right)^*, \quad I_2 = \left(\frac{S_2}{V_2} \right)^*, \quad I_3 = \left(\frac{S_3}{V_3} \right)^* \qquad \text{(2).”}$$

After substitution, the author comes forward with the following:

“The above can be written as

$$\begin{bmatrix} Y_{11} & Y_{12} & 0 \\ Y_{21} & Y_{22} & Y_{23} \\ 0 & Y_{32} & Y_{33} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}, \tag{3}$$

where $Y_{11} = y_{12}$, $Y_{12} = -y_{12}$, $Y_{21} = -y_{12}$, $Y_{22} = y_{12} + y_{23}$, $Y_{23} = -y_{23}$, $Y_{32} = -y_{32}$ and $Y_{33} = y_{23}$. The matrix equation (3) represents the bus admittance matrix; it is also the Y_{Bus} model for Example 7.2.”

Equation (3) in fact is not a final solution. It can’t be, and we will come back to that. To come to the final, the authors in [3] and [4] turn from voltages and currents into power by creating nonlinear equations describing a linear system. These, naturally, need special algorithms in order to be solved (Gauss-Seidel is advised first). That includes creating initial solutions and control of convergence. In [5] Newton-Raphson is further recommended due to its faster convergence.

III. MODIFIED NODAL ANALYSIS

Nodal analysis (NA) is based on the Kirchhoff’s current law and as it is already explained above the node equations is expressing the following: The sum of currents leaving the node is equal to zero. The node voltages are unknown while the currents are first expressed as functions of node voltages using the constitutive equations of the circuit elements connected to a node. Frequency domain analysis is facing a problem when ideal voltage source is connected between nodes since the branch current cannot be expressed as a function of the branch voltage (ideal voltage source has no Norton equivalent). Time domain analysis, in addition, has problem to express the voltage equation of the inductance. Both problems are solved when Modified Nodal Analysis (MNA) is implemented [6,7]. The trick is that the branch current of the voltage element (ideal voltage source or inductor) is introduced as a new variable enabling the proper node equations to be assembled. In that way the number of variables is incremented by 1. In addition, the system of circuit equations is extended by a new equation related to the branch of the voltage element. Fig. 1 depicts an ideal voltage source connected between nodes *j* and *k*, and the “stamp” or the contribution of such an element to the system of equations describing the circuit.

Implementation of MNA is not domain dependent. This means that one uses the same concept for (power-) frequency and the time domain analysis [8], while the system to be analyzed may be linear or nonlinear. To our knowledge there are serious attempts within the power electrical engineering community to adopt MNA and implement it for system simulation [5,8-13]. That is not the case for the LFA as it is thought at universities and hence our intervention.

Node	v_j	v_k	i_{new}
j			1
k			-1
New equation	1	-1	E

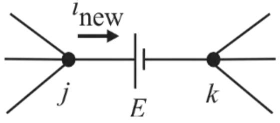


Fig. 2. An ideal voltage source connected between nodes *j* and *k*, and the “stamp” or the contribution of such an element to the system of equations describing the circuit.

IV. LOAD FLOW ANALYSIS VERSUS MODIFIED NODAL ANALYSIS

We will start here to express our full disagreement with developments expressed in second section. To begin with, the node voltage V_1 is already known and one should not expect that the solution of (3) would produce the same value as the original. Then, node 1 (bus bar 1) has no internal impedance which is easily recognized from the fact that S_1 is missing in Fig. 2. That may look as unimportant mistake but in fact is a fundamental error misleading the whole analysis. As already mentioned ideal voltage source has no Norton equivalent meaning that I_1 cannot be related to V_1. The bus bar current is unknown until the end of the analysis since it depends on the load. Theoretically, for the same bus bar voltage, it may vary from zero to infinity. If, however, the ideal voltage source was to be kept in the circuit it would be impossible to formulate the nodal equations as above. One would need to use MNA by which I_1 would be introduced as unknown circuit variable and the order of the system of equations would be raised by 1.

Further, equation (2) is also misleading. Namely, a set of unknown variables ($V_i, i=1,\ 2,\ 3$) is used to create a set of other unknown variables ($E_i, i=1,\ 2,\ 3$) while both stay in the same system. According to the author, to get the "Y_{Bus} model" one should create a system of three equations with six unknowns. That is difficult to believe, and as already commented it leads nowhere. Furthermore, if the node voltages change why would the load power stay constant? Namely, the truth is $\mathbf{S}$, as given in the figure, are *nominal powers* which, for the given impedances are valid only for the *nominal voltages*. The load impedance is the only invariant (since steady load is presumed). In our opinion the loads were to be noted as nominal ($\mathbf{S}_N$) being valid only for nominal voltages ($\mathbf{V}_N$) and used to find the admittances only.

To put the same comment in other words, if $\mathbf{I}$ is unknown (since $\mathbf{V}$ is unknown), why are they on the right-hand side of the equation? Or, how many complex valued unknowns are in (2) and (3), three, six or nine?

Unfortunately, the story does not end here. Namely, looking at the final "Y_{Bus} model" and switching off the source (putting $I_1 = 0$) one gets

$$\begin{bmatrix} Y_{11} & Y_{12} & 0 \\ Y_{21} & Y_{22} & Y_{23} \\ 0 & Y_{32} & Y_{33} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} 0 \\ I_2 \\ I_3 \end{bmatrix} \tag{4}$$

or

$$\mathbf{Y} \cdot \mathbf{V} = \begin{bmatrix} 0 \\ I_2 \\ I_3 \end{bmatrix}, \tag{5}$$

whose solution is

$$\mathbf{V} = \mathbf{Y}^{-1} \cdot \begin{bmatrix} 0 \\ I_2 \\ I_3 \end{bmatrix} \neq \mathbf{0}. \tag{6}$$

Can the grid produce node voltages even if there is no power source active in it?

Here we come to one of the fundamental mistakes done when loads are represented in LFA (and not only in it). Namely, instead of impedances (which are in fact the loads) one uses current sources. However, a current source is a *source of energy* even if it is negative by value while the impedance is a *consumer of energy*. Interchanging those leads to wrong solutions (like, having voltages and currents in absence of real sources in the system). Furthermore, looking in circuit theoretical mode, the impedance's current changes linearly with the change of its voltage. The current of a current source is constant and independent of the voltages drop on it (looking at it as if it is impedance, it becomes nonlinear conductance with hyperbolic decay). Finally, by use of a constant current source we ignore any change of the currents in all loads. So, since we know the currents ($\mathbf{I}$) and the loads ($\mathbf{S}$), there is no need for any analysis. The voltages may be obtained from (2) backwards.

To conclude, both node voltages and load currents (and consequently load powers) are unknown at the beginning of the analysis and the concept expressed in [3,4] is misleading.

One is not to forget that the books we are speaking about ([3] had two and [4] eight editions) are intended to be read by novices (students).

Before proceed to implementation of the MNA as an ultimate solution of the equation formulation problem we will try to accommodate to the NA. Of course, we will suppose that $\mathbf{V}_N$ and $\mathbf{S}_N$ are known which means we know the load admittances Y_{L2} and Y_{L3}.

Supposing I_1 is known, one is to substitute the admittances of the loads and proceed with the formulation in the form

$$\begin{bmatrix} Y_{11} & Y_{12} & 0 \\ Y_{21} & Y_{22} & Y_{23} \\ 0 & Y_{32} & Y_{33} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ 0 \\ 0 \end{bmatrix}, \tag{7}$$

where $Y_{11} = y_{12}$, $Y_{12} = Y_{21} = -y_{12}$, $Y_{22} = y_{12} + y_{23} + Y_{L2}$, $Y_{23} = -y_{23}$, $Y_{32} = -y_{32}$ and $Y_{33} = y_{23} + Y_{L3}$.

Again, Y_{L2} and Y_{L3} are the admittances of the loads which are passive and do not generate energy (hence, must be on the left-hand side of the equation).

Since V_1 (not I_1) is known, by simple manipulations one gets

$$\begin{bmatrix} -1 & Y_{12} & 0 \\ 0 & Y_{22} & Y_{23} \\ 0 & Y_{32} & Y_{33} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} -Y_{11}V_1 \\ -Y_{21}V_1 \\ 0 \end{bmatrix}. \tag{8}$$

This seems an acceptable final solution. It resembles the case when the admittance of the first line, $Y_{11} = 1/Z_{1-2}$, is used as internal admittance of the voltage source V_1. This (handmade) trick is, unfortunately, not applicable when more than one line is connected to the bus bar.

In other words, we would prefer a universal method enabling in the same time for automatic (computer generated) equation formulation. Here comes the MNA. If MNA was to be applied instead (7) one would have:

$$\begin{bmatrix} Y_{11} & Y_{12} & 0 & -1 \\ Y_{21} & Y_{22} & Y_{23} & 0 \\ 0 & Y_{32} & Y_{33} & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ I_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ E_g \end{bmatrix} \qquad (9)$$

Here a voltage source (E_g) is connected to node 1. This system complies with the equation formulation used in the SPICE program [12,14] meaning that if SPICE description were created one would produce the solution in no time. Note there are versions of SPICE available for free.

To finalize, one is not to forget that when all node (or branch) voltages are known we may calculate *everything*: currents and powers, meaning there is no need to switch to powers before the node voltages are found.

V. CONCLUSION

To conclude, it is our opinion that it is time to demystify the power system analysis (including PFA) as taught until now, since despite of a very long period of use of such systems is still obscure and ineffective. Using a simple, but as general as necessary, example we showed that MNA should be taught to our student as a universal means of equation formulation for circuits and systems simulation.

ACKNOWLEDGEMENT

This research was partly funded by The Ministry of Education, Science and Technical Development of Republic of Serbia under contract No. TR32004.

REFERENCES

[1] Dimitrijević, M., Milojković, J., Bojanić, S., Nieto-Taladriz, O. and Litovski, V. (2013), "ICT and power: new challenges and solutions", *International Journal of Reasoning-based Intelligent Systems,* Vol. 5, No. 1, pp. 32-41., doi:10.1504/IJRIS.2013.0.

[2] Bellosta, L.C.M, et al. (1998), "Simulation of Surges on Power Lines Using SPICE and EMTP: A Comparative Study", in *Proceedings of the 9th Mediterranean Electrotechnical Conference in Tel-Aviv, Israel, 1998*, Institute of Electrical and Electronics Engineers, pp. 202-206, doi: 10.1109/MELCON.1998.692371.

[3] Keyhani, A. (2016), *Design of Smart Power Grid Renewable Energy Systems,* Wiley-IEEE Press, New York, USA, second edition.

[4] Weedy, B.M., Cory, B.J., Jenkins, N., Ekanayake, J.B. and Strbac, G. (2012), *Electric Power Systems*, John Wiley & Sons Ltd, New Jersey, USA, Fifth Edition.

[5] Vergura, S., Liserre, M. and Vacca, F. (2007), "Adjoint Network Theory to Analyse the Power Converters with Respect to their Line-side Behaviour", in Lattarulo, F. (Ed.), *Electromagnetic compatibility in power systems*, Elsevier Science, Amsterdam, Netherlands.

[6] Ho, C.W., Ruehli, A. and Brennan, P. (1975), "The Modified Nodal Approach to Network Analysis", in *Proceedings 1974 Int. Symp. on Circuits and Systems in San Francisco, USA, 1975*, Vol. 22, No. 6, Institute of Electrical and Electronics Engineers, pp. 504-509, doi: 10.1109/TCS.1975.1084079.

[7] Litovski, V. and Zwolinski, M. (1997), *VLSI Circuit Simulation and Optimization*, Chapman and Hall, London, United Kingdom.

[8] Lattarulo, F. (2007), *Electromagnetic compatibility in power systems*, Elsevier, Amsterdam, Netherlands.

[9] Hsu, Y.C. (2010), "Design, Modeling and Validation of Permanent Magnet Generators for Wind Turbines", Department of System and Naval Mechatronic Engineering, Taiwan.

[10] Hsu, Y.C., Hsieh, M.F. and McMahon, R.A. (2009), "A General Design Method for Electric Machines Using Magnetic Circuit Model Considering the Flux Saturation Problem", in *Proceedings of the Eighth Int. Conf. on Power Electronics and Drive Systems in Taipei*, *Taiwan*, *2009*, Institute of Electrical and Electronics Engineers, pp. 625-630, doi: 10.1109/PEDS.2009.5385860.

[11] Kato, S., Hoshi, N. and Oguchi, K. (2004), "Brushless slip-power recovery system simulation using modified nodal approach", *IEEJ Trans. IA*, Vol. 124, No. 12, pp. 1252-1260, doi: 10.1541/ieejias.124.1252.

[12] Hsieh, M. and Hsu, Y. (2012), "A Generalized Magnetic Circuit Modeling Approach for Design of Surface Permanent Magnet Machines", *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 2, pp. 779 - 792, doi: 10.1109/TIE.2011.2161251.

[13] Thomas, R., Lahaye, D.J.P., Vuik, C. and Van Der Sluis, L. (2013), "Transients in Power Systems: A Literature Survey", Reports of the Department of Applied Mathematical Analysis, Delft University of Technology, Delft, Report 13-12, uuid: f30d85e1-3349-4a52-8e74-53bddca50d91.

[14] Nagel, L. W. (1975), "SPICE2: A Computer Program to Simulate Semiconductor Circuits", EECS Dept. Uni. of California, Berkeley, Tech. Report No. UCB/ERL M520

Proceedings of the 8th Small Systems Simulation Symposium 2020, Niš, Serbia, 12th-14th February 2020

Development of C# Application for Neural Network Based Precipitation Data Mining

Miljan Jeremić, Milan Gocić, Miljana Milić, Jelena Milojković

Abstract – The impact of climate changes, and especially its negative effects to human society, represent potentially devastating problems of today and the future that cannot be easily confronted, defined, and solved. In order to successfully combat climate changes' negative effects, it is necessary to assess their actual beginning in time as well as its duration and the final effects, but it is also difficult to determine the intensity and points of observation. Due to the extensive damage caused over a long period of time, it is necessary to find reliable and standardized indicators in space and time that would define phenomena such as rainfall, droughts and floods. The paper presents an application in C# programming language that, using data from multiple measuring stations in Niš, Serbia recorded in a database, and a machine learning technique, such as Data Mining, creates a neural network that models and analyzes a rainfall index for a period of 20 years.

Keywords – Climate changes, Data mining, neural networks, machine learning.

I. INTRODUCTION

Learning is a process that humans perform almost continuously. Since the ability to learn is a basic characteristic of intelligent beings it is not surprising that machine learning represents the central area of research in artificial intelligence. Artificial systems that are capable of learning, evidently improve their performance, while the intelligent biological systems increase the probability of their survival and extension of the species.

Machine learning [1, 2] is an area that studies the processes that underline learning in humans as well as in artificial systems. In order to cover all relevant aspects, it relies on a number of other disciplines, including artificial intelligence, probability and statistics, information theory, psychology, neurobiology, and control theory

Inductive learning or inductive empirical learning has the greatest importance in the field of computing and artificial intelligence. The idea of this type of learning is to learn from the available examples or to learn from the experience of others. Considering the object of learning, the most general learning methodology is the functional learning, i.ee input-output mappings.

Measuring the quality of a trained system is one of the key problems in machine learning. It can easily happen that the chosen hypothesis represents the most reliable one for a given training set, but also very unreliable on new examples that did not participate in the training i.e. the test set. The ability to generalize or to correctly deal with instances that were not the part of the training set data, is a measure of the quality of any machine learning algorithm. The synthesis is based on two sets, a training set and a test set. The test set data should not be used in the training stage. A common statistical measure of system performance is the mean square error that is calculated using the system's output values and target values.

In practical applications there are many different machine learning models, and their structure depends on the information contained in the training sets, how the training process is controlled, and whether the training system can affect the surrounding.

The aim of this paper is to describe the entire creation process of a Data Mining architecture for a system that, based on a neural network type machine learning model, shows the visual result of rainfall estimation in the territory of the city of Niš during one year. It begins with data acquisition from the relational database, and ends with the visualization of the machine learning algorithm's conclusions. In the next section few most important machine learning algorithms and techniques will be explained briefly. Then, the basic information about the initial precipitation data will be listed. It is followed by the applied Data Mining methodology [3], and its results. The paper ends with some concluding remarks.

II. MACHINE LEARNING ALGORITHMS AND TECHNIQUES

There are many available data science techniques today, such as linear regression, logical regression, SVM (Support Vector Machine), Random Forest algorithm, K-means clustering, neural networks and convolutional neural networks, and they are most commonly applied on data in the form of an Excel CSV file or in JSON format.

One of the best-known techniques of machine training are neural networks and decision trees [4, 5].

A. Decision making systems

M. Jeremić and M. Gocić are with the Facultyof Civil Engineering and Architecture, University of Niš, Aleksandra Medvedeva 14, Niš, Serbia E-mail miljan.jeremic@gmail.com, milan.gocic@gaf.ni.ac.rs.

M. Milić and J. Milojković are with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, Niš, Serbia E-mail: miljana.milic@elfak.ni.ac.rs, jelena.milojkovic@elfak.ni.ac.rs.

There are three large groups of machine learning techniques [2]: supervised learning, unsupervised learning and reinforcement learning.

In the supervised learning training set is in the form $\{x_i, f(x_i)\}$, where $f(x_i)$ denotes a target value for the instance x_i. The task of the training is to find the approximation for the function f, while the measure of performance is the quality of the approximation for points that are outside of the training set.

For the unsupervised learning, the training set contains only the input instances $\{x_i\}$ A typical problem for the unsupervised learning is the clustering problem, i.e. the classification of the available data into a smaller number of groups. Since the clustering is done on the basis of data similarities, it means that one cluster contains data with similar properties. The reinforcement learning is more complex learning concept in compare with the supervised and the unsupervised one. It originates from the control theory in which a dynamic environment can be described with a condition, action and a reward. Within this type of learning technique, it is necessary to learn how to map the situations into actions, while achieving the maximal reward. The training algorithm is not familiar with types of actions that should be performed for the particular situation. An example of such a learning process is learning how to play chess.

B. Neural networks

Neural networks [8] are designed to mimic the structure of neurons in the human brain. Each artificial neuron should be connected to other neurons within the system in a proper way. Neural networks consist of layers of neurons. Data is transferred from one layer of neurons to the next - inner layer. Eventually, the signals within the network arrive at the output layer, where the network presents its best troubleshooting assumptions.

The examples of neural networks; applications could be found in a series of different industries, and telecommunications and media as well. They could be used for text or language translations, frauds detections etc. Also, their usage is common in digital filters, function approximations, sample detections, forecasting in many complex processes. One large area that exploits neural networks is gaming

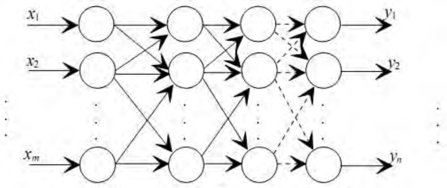


Fig 1. A neural network with its connections

The advantages of neural networks are that they can be used to model complex systems, they are more robust and give more accurate predictions than linear models. The disadvantage of this technique is that it is not suitable for hypothesis testing and is therefore used in combination with statistical techniques [6].

There are many types of artificial neural networks topologies today [7], but the most commonly used is a feedforward back propagation multilayer perceptron. Based on a set of known inputs and outputs for the real problem, training of the network is performed by applying the known inputs of the modeled problem to the inputs of the network, and observing the difference between the known output of the problem and the output of the neural network. The aim of the training is to minimize this error. Artificial Neural Network mimics the functioning of the human brain, where the data processing is performed by connecting a large number of neurons. A typical neural network consists of multiple layers, one input, one or more hidden layers, and one output layer. The number of hidden layers depends on the complexity of the problem being modeled, but for common classes of problems one or two appear to be enough. Figure 1 shows a neural network consisting of three layers: the input, hidden and the output layer.

Neural networks accept multiple inputs in parallel and process the information received in a distributed manner. The information stored in the neural network is distributed across multiple computing units, which is opposite to the conventional information storage in memories where each specific information is stored in its memory space. The property of distributed information storage as well as the redundancy are the most important benefits [8].

A neural network has the ability to learn and adopt, which makes them applicable for processing uncompleted datasets in the unknown or unlearned environment. The network can generalize and conclude how to process data that were not present in its training set. Considering their structure, neural networks represent multivariable systems, which makes them useful for modeling, identification and control of multivariable processes.

C. Data Mining

Data Mining can be defined as the procedure of examining large databases in order to generate some new information about these data and their relations. It is a technique used to discover the hidden, valid, and potentially useful regularities and patterns amongst the data in large datasets. It employs machine learning, statistics, AI and database technology and is therefore a complex multi-disciplinary procedure that can be useful in solving everyday problems using data.

Very often, classic database reports do not provide data required for the end users, or decision makers. Some advanced techniques then need to be employed in order to extract more specific information or even business predictions from the available data.

The idea is to replace the decision makers with an appropriate software that can predict for a client [2]. In order to perform such analyzes the knowledge or support of a particular set of algorithms is required. It should be mentioned that Data Mining also uses some statistical methods for data analysis.

There are few Data Mining algorithms that should be studied first in order to select the one that is most suitable for a particular data analysis problem

Some of the most often Data Mining algorithms are: classification algorithms, regression algorithms (linear or logic regression dedicated for one or two steps ahead forecasting), segmentation/cauterization algorithms (objects are separated into sets with some common properties), association algorithms (dedicated for the correlation search among data) and sequence analysis algorithms (they help in search for the user's path through the web, or the order of inserting items into the cart, etc.) .

Classification is the most common Data Mining methodology, whose goal is to accurately predict the class of the target object that has the unknown class label [9-10]

III. DATA AND THE FIELD OF STUDY

In this section, a step by step procedure of the Data Mining system development will be given. The selected machine learning approach that supports our Data Mining problem are Neural Networks. An *SQL Server Data Tools* toolset, *Visual Studio,* and a *C#* programing language are used for this purpose [11]. The problem to be solved must be defined and presented through Data Mining analysis.

The procedure is exemplified on the analysis of the precipitation value for the city of Niš, using the available application and its neural network algorithm in order to generate a graphical representation of the detected hidden regularities in the available data.

At the beginning of the procedure, one needs to determine what kind of data set we will be applied as the input of the Data Mining algorithm. The data for the selected Data Mining algorithm need to be prepared for further analysis and decision making. Data necessary for the Data Mining analysis are usually preprocessed first. It should be defined what are the input data or variables, and what should be the outputs i.e. for example the forecasts. It is usually not possible to get a direct access to a data storage, a relational database, or an Excel spreadsheet and expect some output to be generated from that vast amount of written data.

The data from different sources should be selected, cleaned, transformed, formatted, anonymized, and constructed (if required).

Data cleaning is a process where noisy data are removed from the series, while the missing data are generated and filled.

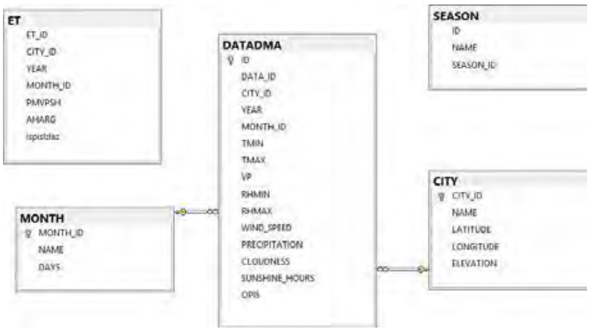


Fig. 2. Meteorological database relational model

The intention here is to analyze the output values for the PRECIPITATION variable recorded and stored in an SQL Server database. The relational model of the meteorological database is shown in Figure 2.

Some data transformations could be applied on data to make it useful for Data Mining. Some of them are: smoothing, aggregation, generalization, normalization or attribute construction.

For the requirements of this application, a query has been generated in the SQL Server database, where a new column is formed that tests each value of the PRECIPITATION and classifies it into three categories: small, medium and large. This table is denoted with DM and it will record all the values of the required variables applied at the input. The second table, DM1, will be created after the SQL Server query is executed. Results of this query will be written in a Predicted column. The table has 348 records in total. This will be a source of data for further Data Mining analysis.

After creating the appropriate source of the required data storage, it is now possible to begin with the development of the application in a specific environment. Here, we have used a Visual Studio, version 2015. For a given project and the available data, it was the most suitable to use Data Tools toolset.

IV. METHODOLOGY

Within the *Data Mining* algorithms, the input data are processed and the corresponding values are generated at the particular outputs. After all necessary preparations over the data, different mathematical models could be applied to determine patterns amongst data. For our problem neural network-based classification modelling technique was selected. Other available techniques are decision trees genetic algorithms and similar.

Within the Data Tools environment, the Business Intelligence tab is used to create the new project. Then the *Analysis Services Multidimensional* and *Data Mining Analysis* were used. These toolboxes as well as the Data Mining mode must be installed and enabled within the SQL Server.

In any Data Mining project, data sources and their views are basic objects used to define tables and queries for Data Mining structures. In turn, a structure defines one or more tables or columns that are used as input attributes, keys, and prediction outputs for the structure.

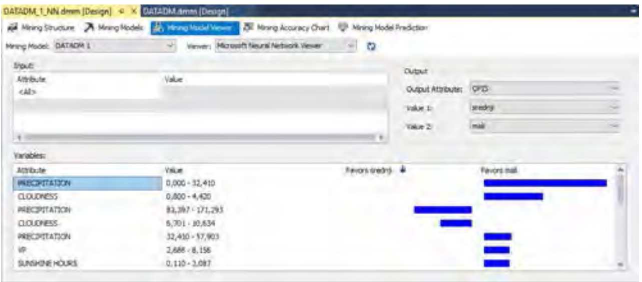
Now our particular mining structure can be created, with the defined source of data the desired mining algorithm. In this case the structure's source of data is the database with the values for meteorological parameters, which is shown in Figure 3. After this step, a connection between the mining structure and the database should be established.

Fig. 3. Creation of the *Mining* model and the corresponding neural network withing the Analysis Services

As mentioned before, the created model has the direct access to the database in order to select the required table. The next step is to create a logical data view, or Data Source View. As a Mining structure the Decision Trees algorithm is selected, which is used to predict or classify one or more discrete variables. In this example, this is the PRECIPITATION variable.

When the algorithm processes the data, it saves some of the data for later testing. In the next step, the percentage of the data that will represent the test set should be selected. In this case this is 20%.

Finally, the Mining Model Viewer tab shows the results of neural network-based Data Mining. This as well as the generated precipitation decision tree is shown in Figures 4 and 5, respectively.



Mining Model: DATADM 1
Viewer: Microsoft Neural Network Viewer

Attribute	Value
<All>	

Output Attribute: CPD
Value 1: srednji
Value 2: mali

Attribute	Value	Favors srednji	Favors mali
PRECIPITATION	0,000 - 32,410		
CLOUDNESS	0,800 - 4,420		
PRECIPITATION	83,397 - 171,293		
CLOUDNESS	6,701 - 10,634		
PRECIPITATION	32,410 - 57,903		
VP	2,688 - 8,156		
SUNSHINE HOURS	0,110 - 3,087		

Fig 4. The report window generated by the C# program

The generated data model is uploaded to Analyses Services within the DM database. It should be noted that within the SQL Server one can find a model in the form of the decision tree or a neural network, the same as one shown in the C # application. A prediction query can also be created in this part of the database. Now some reports can be created that are results of data processed through our data model.

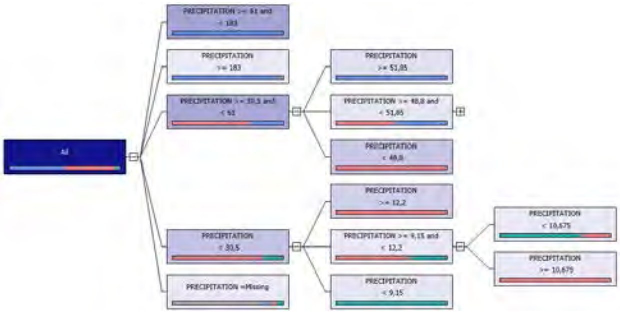


Fig 5. The decision trees for the PRECIPITATION variable

V. RESULTS AND DISCUSSION

Queries are generated with a single mouse click, without prior knowledge of the query. The query itself should eventually consider a prediction function to be applied, and make a prediction based on one or more input columns. After running the query, a probability of the occurrence for a particular output value is obtained based on a previously created data model.

In this paper the meteorological data analysis for the period from 1990. to 2018. stored in the SQL Server database is performed. When the PRECIPITATION is taken as an observation parameter, the data Mining model was created over these data, that applied a neural network-based analysis structure. The algorithm offered the following conclusions about the given PRECIPITATION data: the expected average value for a month should be greater than 32.411mm; the most probable future value would be between 83 and 172mm, which can be classified as the medium precipitation value.

VI. CONCLUSION

The most important benefits of the Data Mining procedure are: the ability to create useful knowledge-based information for the companies, and act accordingly and timely. It is a cheap and efficient methodology that facilitates reasonable and automated forecasting and decision-making, by quickly discovering the hidden correlations amongst data

Creating such an application that works with real data, and generates a neural network, provides the necessary background for the needs of drought analysis and assessment in order to mitigate its unwanted consequences.

REFERENCES

[1] Tan, P. N., and Steinbach, M., Kumar, V., Introduction to Data Mining, Pearson Education, 2006.
[2] Milosavljevic, M. *Artificial Intelligence (in Serbian)*, Singidunum University, Belgrade, 2015.

[3] Chiwara, A. A., and Gupta, H., *Data Mining, Concepts and Techniques Association Rule Mining*, State University of New York, CSE 634, Chapter 8, 2006.

[4] Bhargava, N., Sharma, G., Bhargava, R., and Mathuria, M. "Decision tree analysis on j48 algorithm for data mining", *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(6), 2013.

[5] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., and Zhou, Z. H. "Top 10 algorithms in data mining", *Knowledge and information systems*, *14*(1), pp. 1-37, 2008.

[6] Kalinić, Z. S., and Marinković, V., "Određivanje relativnog uticaja pojedinih faktora na prihvatanje mobilne trgovine primenom neuronskih mreža", *Poslovna ekonomija*, 10(2), pp. 206-223, 2016.

[7] Lu, H., Setiono, R., and Liu, H.," Effective data mining using neural networks", *IEEE transactions on knowledge and data engineering*, *8*(6), pp. 957-961, 1996.

[8] Craven, M. W., and Shavlik, J. W. "Using neural networks for data mining" *Future generation computer* systems, 13(2-3), pp. 211-229 1997.

[9] Kesavaraj, G. and Sukumaran, S. "A Study on Classification Techniques in Data Mining" *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT),* pp. 1-7. IEEE, 2013.

[10] Begüm Ç. and Deniz Ü., "Comparison of Data Mining Classification Algorithms Determining the Default Risk", *cientific Programming*, *2019*.

[11] Watson, B., C# 4.0 kako do rešenja, SAMS, Mikro knjiga, 2011.

List of Authors