IEEE

Brno University of Technology
Student Branch

Sborník příspěvků studentské konference

Kohútka 2015

17. - 19. 8.

Název: Sborník příspěvků studentské konference Kohútka 2015

Editor: Ondřej Zach

Vydavatel: Vysoké učení technické v Brně
Fakulta elektrotechniky a komunikačních technologií

Rok vydání: 2015

Vydání: první

Organizační výbor konference:

Martin Kufa
Roman Mego
Miroslav Staněk
Ondřej Zach

ISBN 978-80-214-5239-8

Konferenci podpořili

Stříbrný sponzor

ROHDE&SCHWARZ

Bronzový sponzor

CST

Další partneři

IEEE
Czechoslovakia Section

ÚREL
ÚSTAV RADIOELEKTRONIKY

Úvodní slovo

Drahí priatelia,

po roku sme sa opäť vrátili na dnes už tradičnú konferenciu doktorandov a mladých vedcov, organizovanou študentskou sekciou IEEE VUT v Brne. Sme radi, že sa aj tento rok môžeme spoločne stretnúť v neformálnom kruhu, kde sa preberajú témy z nášho profesijného života, ale aj mimo neho.

Tento ročník konferencie sa koná v rekreačnom stredisku Kohútka, situovanom v pohorí Javorníky, ktoré je presne na hraniciach medzi Českom a Slovenskom. Aj vďaka tomu sa tento rok zúčastnili okrem študentov z českých univerzít aj študenti zo Slovenska. To našej konferencií dodáva oproti minulému ročníku medzinárodný nádych. Dúfame, že si vďaka tomu odnesiete od svojich kolegov nie len nové poznatky či už z Vašej oblasti pôsobenia, alebo mimo nej, ale aj reálny pohľad na prácu vedcov v zahraničí.

Takúto akciu by sme samozrejme nemohli zorganizovať bez finančnej podpory. Okrem Československej sekcie IEEE nás po prvýkrát podporili aj spoločnosti celosvetových mien, konkrétne Rohde & Schwarz a CST. Týmto by sme všetkým poďakovali za ich podporu a prejavenú dôveru.

Na záver by sme Vám chceli poďakovať za účasť na konferencií, pretože tým súčasne podporujete to, čím je samotné IEEE. Prajeme Vám príjemne strávený čas s Vašimi kolegami a priateľmi.

Za študentskú sekciu IEEE v Brne

Roman Mego

Obsah

Jakub Arm, Zdenek Bradac
Hardware Scheduler Monitor **4**
Sang Van Doan, Dalibor Biolek, Jiri Vavra, Zdenek Kolka
Emulation of Memristive System Based on Digital Processing with Using Digital Potentiometer **8**
Patrik Hubka, Jaroslav Lacik
Parametric Study of HMSIW Linearly Polarized U-Shape Slot Antenna **12**
Eva Klejmova
Comparison of Methods for AR Coefficients Estimation **15**
Marie Klimesova
Stability of Stochastic Differential Systems **18**
Martin Kotol
Estimation of Statistical Parameters of the Digital Signal Using Artificial Neural Networks **21**
David Krutilek
Computational Electromagnetics in Aerospace **25**
Zenon Kuder
Wireless Network Coding in LTE Relays **29**
Jan Kufa, Tomas Kratochvil
Impact of Predefined Quality Profiles in H.265 to Video Quality and Speed of Encoding **32**
Martin Kufa
Filtering Antennas: Introduction of Different Concepts **36**
Jiri Lambor
Study of Radiation Properties of SIW Horn Antenna at 60 GHz **41**
Ondrej Manas, Roman Marsalek, Ahmed Elsokary, Vaclav Valenta
Implementation of Cyclostationary Based Detection **44**
Martin Matis, Lubomir Dobos, Jan Papaj
Velocity Impact on Multi-Hop Communication in MANET Environment **48**
Roman Mego
Hi-Speed USB Communication with FPGA **52**
Michal Mrnka
Higher-Order Modes in Dielectric Resonator Antennas **56**
Martin Pospisil, Roman Marsalek
Digital Predistortion - from MATLAB Simulation to Fixed Point Implementation in USRP **61**
Jan Schneider, Jan Gamec, Daniel Novak, Alena Pietrikova
Quality of Solder Joints in Automotive Electronics after Accelerated Aging Tests **65**
Lubos Struharnansky, Matej Pacha
Combined System for Testing of AC Traction Motors **68**
Jan Velim, Zbynek Raida
The Influence of Objects on Surface Waves - the Determination of the Reference Obstacle **71**
Ondrej Zach
Current Video Coding - Solutions and Challenges **74**

Hardware Scheduler Monitor

Jakub Arm
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: xarmja00@stud.feec.vutbr.cz

Zdenek Bradac
Department of Control and Instrumentation
Brno University of Technology
Brno, Czech Republic
Email: bradac@feec.vutbr.cz

***Abstract*—This paper takes a brief look into fault tolerant techniques and presents a proposal of hardware hard real-time OS scheduler monitor. Purpose of the new proposal is to detect earlier software failures caused by unpredictable hardware faults or system overload. The proposal focuses more precisely on scheduler failures like a deadline miss and a delayed switching context time. Proposed monitor can be used as a standalone scheduler monitor or as a part of redundant system controller. The assumption is that the scheduler monitor detects hardware and RTOS kernel faults earliest. This should enhance RTOS system reliability and safety.**

I. Introduction

An operating system is an essential part of more complex electronic control system based on CPU. Its purpose is to provide an environment in which user programs are executed. This environment builds a layer over system resources and manages access to them. Because of higher requirements to control systems, the operating system has to be as efficient as possible, even if computation power is growing or can be connected together. The requirements of operating system is to access big computation power as fast as possible, moreover manage system resources to accomplish that. In the real-time world, good information on the right place in specified time is required. This is not guaranteed only by high computation power but by precise and deterministic controlling of tasks to be executed. The operating system has to decide which thread will be run and when. In specific areas like aerospace, military or medical, there are other significant requirements to the real-time operating system, i.e. reliability and safety. Some systems are supposed to run every time, so the high reliability is demanded. This is mainly due to costs of the system blackout. In every system can occur some fault. The high reliable system has to react to these faults and do some action, at least mask the fault. To do some action, the system has to detect this fault before it is too late and fault became a system failure. Some systems can operate even with the failure occurrence by lowering its capabilities. The reliability of the system consists of reliabilities of each part of the system like hardware, the operating system and the user application. The operating system is in the middle of the system, so it can catch the most system failures [1], hardware faults not causing immediate system crash as well as software faults come from the user application or from the operating system self. For the safe hard real-time operating system, there is a requirement for as high reliability as possible. The system has to minimize the failure risk before the allowable border. There are many levels which these risks can be reduced on and methods that can be used. Each method minimizes some source of faults. In these systems, all possible precautions against the failures have to be made. This makes up the functional safety according Safety Integrity Level. Systems interacting with human or system for humans can have higher requirements to safety of the system. This means when some failure occurs, the system has to do actions and ends up in fail-safe state to not be dangerous and deterministic. This feature of the system is called safety.

II. Fault sources

The sources of faults can be divided to hardware and software faults. Software faults can be caused by wrong construction of the system and by wrong program. Construction software faults are mainly caused by unsatisfactory EMC or ESD protection. Software faults can be caused by unspecified behaviour of microprocessors, by programmers and development environments, by faulty peripherals, or by unexpected human factor. Most of hardware faults can be avoided by careful and robust design of the system. Software failures can be avoided by modeling the system and performing analysis. But the system can be always affected by unexpected EMI radiation or by faults of hardware (broken transistors, broken lines, overpower, etc.) due to a fault in production.

III. Fault-tolerant system

"Anything that can go wrong, will go wrong." [2]. However, fault tolerant system has to deal with these faults to prevent failures, even for a price of small degradation. Crucial quantity of system functional safety is reliability defined by equations (1 and 2), where λ stands for the expected count of failures per time unit and $MTBF$ stands for mean time between failures. **Reliability** is therefore a probability of faultless system working in time unit. Reliability of important systems (e.g. in aerospace) has to be as high as possible. An impact of faults has to be eliminated. Following chapters describe ways and methods to minimize this impact of these faults or even eliminate them.

$$R(t) = e^{-\lambda_K t} \tag{1}$$

$$\lambda_K = \frac{1}{MTBF} \tag{2}$$

IV. Fault avoidance

Fault avoidance is a first set of methods on a way to minimize faults, i.e. comprehensive project description, using design templates or standard methods. These methods can be

done in a design phase of the project. **Comprehensive project description** leads also to time savings in an execution phase of this project. **Design templates** are also good when more teams are involved in the project to fast and better understanding of other team solutions.

Nowadays, creation of a **formal model** of a process is expanding like a method to model and then validate the process. This method tries to eliminate faults and verify some parameters of designed process. This method also verifies process features before it is physically created which saves programmers effort. But all faults can not be modeled. As a model there can be used fault-tree analysis model or model from UPPAAL.

Methods that can be used to avoid faults in an execution phase of the project are excessive cooling arrangement, generally excessive construction, using a signal shielding and other issues increasing electromagnetic immunity of a projected device. Basic using of these methods decreases rapidly the count of faults.

V. Fault masking

Fault masking is next level way to minimize faults based on hiding them. In case, a fault occurs in the system these methods try to mask to prevent failure creation.

Redundancy can be used when possible. More than one device do the same job and is parallel connected with the first device. The wiring can be redundant as well. In case of any fault on any device, the other devices set correct outputs. The incorrect output can not interfere with the correct one.

Triple modular redundant systems are based on redundancy. They have more separate parts doing the same, moreover, there is a voter which decides what is the correct output value using medians of all input values. Using more level median computation of values more times per time unit can output correct value in case the most of parts are working correctly.

When failures in data are small, there can be used technique called **self repairing code**. This method is used mostly in memories to recover small amount of a damaged data. This method requires some rule for data storage, e.g. linear code.

VI. Fault tolerance

Fault tolerant methods can be used when other methods can produce faults. The most used method is reconfiguration of a system. **Reconfiguration** unlike redundancy needs to have reconfiguration arbiter who decides when and how the system will be reconfigured. The arbiter replace faulty part of system with the correct one. In PLC systems, this is called *hot swap*.

VII. Fault detection

Fault detection is a first step in a fault tolerant system to eliminate failures. This system has to catch all unexpected or expected faults. Some faults are still caught in development phase mostly due to documentation, systematical architecture, used standard or tests.

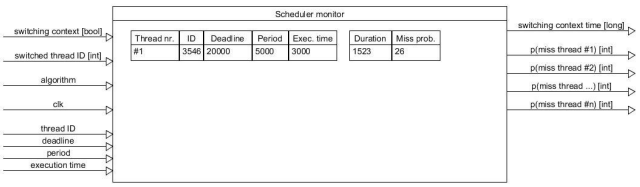


Fig. 1. Scheduler monitor realization

A. Watchdog

Watchdog is the first, the simplest and the most used way how to detect software hangs and unexpected delays. Watchdog is mostly done as a hardware peripherial controlled through memory registers. There can be software watchdog which is running in separate thread. The count of software watchdogs can be dynamically growing. On the other side, this type of watchdog depends on non-failure state of underlying system - OS kernel, CPU and other hardware. Mostly, the internal watchdog of microcontroller is used. The internal saves money, but can be affected by runaway code. An external WDT has a separate clock source that gives it better reliability [3]. If watchdog overruns, the monitored system is restarted. This increases reliability of the system, i.e. functional safety. But safety of the system are lowered by incorrect reset state settings. There are projects [4] implementing more watchdogs in one chip. But their count do not have to fit the requirements of the implemented system. The system can be dynamically wired, e.g. in FPGA, but every newly added watchdog uses the same amount of a chip space. Programmer has to have some system to feed these watchdogs also. There is always the famous question “What if the watchdog fails?” [2].

B. Operating system monitors

An operating system or more precisely the kernel has built in some monitor functions. For example Linux kernel contains tracing functions to get maximal the switching context time, the count of external interrupts, or function tracing. Method described in [5], is based on measuring system calls, OS signals, task schedulation timeout, or I/O throughput by software. These particular measurements are processed through neuron to make a decision if fault occurs. These supportive features are however dependent on faultless state of underlying kernel functions and monitored hardware platform.

C. Scheduler failures

Scheduler of the real-time operating system has to be deterministic. So that, the exact algorithm for predictable actions has to be implemented and a switching context time has to be bounded. Scheduler then clearly decides which thread to execute. In case of normal processor operation, the real-time OS controls precisely thread execution. In an other case like system overload (when the processor utilization factor is greater than one) due to incoming events or hardware faults, the RTOS can miss some deadlines [6] according to scheduler algorithm or the switching context time can be higher than expected. These scheduler failures are usually caught by watchdogs. Missed deadlines in a priority based RTOS are masked. However, watchdog detects these failures when expires his period. Meantime the RTOS can cause disaster.

Moreover, if the period is not set carefully. Watchdog period can not be set precisely because the switching context time may vary [7] or wait time for external resources may vary.

VIII. Proposal

The proposal focuses on monitor missed deadlines or even future missed deadlines of periodic threads, and on monitor the switching context time online. The proposed scheduler monitor will be done in hardware on the same chip as monitored system to meet these tendencies. Firstly, the monitor is placed the most far off the monitored system due to better reliability and smaller interference. Secondly, the monitor is placed near enough to monitor internal hidden state and to lower costs. As a best result of these requirements, microcontroller hardware peripherial was chosen. In development state, this will be realized on FPGA. There are projects focused on monitoring processor with another FPGA [8]. That means that MCU is connected via its peripherials to FPGA which monitors run state via IO port. This approach minimizes the influence of EMI radiation but on the other hand the state is mediated. So MCU internal state is hidden through IO barrier. The scheduler monitor has to be connected to scheduler to obtain real time information about switching thread execution. According to these information, it can compute real time the probability of missing deadline for each thread. These probabilities will be also used as an input to a failure classificator which will publish information to system flow controller which will perform actions like restart the system or set the safe state. This system will enhance reliability and safety of the monitored system. The binding to the monitored system has to be made. So the scheduler has to be rebuilt to provide information to the scheduler monitor. The monitor can be used for every type of scheduler algorithm.

A. Realization

The scheduler monitor [fig. 1] has to have a digital signal, realized by mapped memory, which active level means that scheduler is switching the context. According to this signal, the scheduler monitor measures the switching context time and publishes it in its memory space. After the scheduler switches the context, it will write the thread ID on the specific address of scheduler monitor. This causes update in measuring the duration of thread execution. According to thread parameters given at a thread creation, scheduler monitor calculates cyclically the probability of thread missed deadline. This probabilities are published in monitor memory space. The monitor gives only information public. It has to be connected at least to the classificator [fig. 2] which can react to the situation. The simplest classificator consists of comparators which decide the maximal tolerable value of probability and of the switching context time. The classificator can have feature to bring the system into fail-safe. After new monitored thread is created, its parameters has to be written to the monitor. By writing *thread_ID*, new monitored thread is added to monitor table. So the other parameters has to be written prior. If this is not done, the thread will not be monitored. If none thread is written to the scheduler monitor table, the monitor will only monitor the bounded switching context time and deadlines can be monitored by watchdogs.

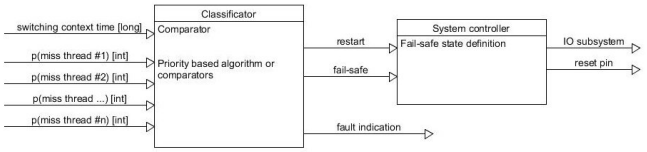


Fig. 2. Classificator realization

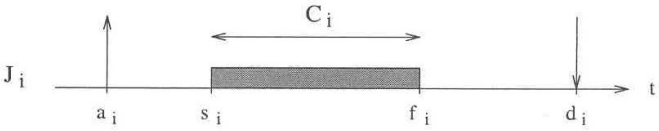


Fig. 3. Real-time task parameters [9]

B. Probability calculation

The assumption is that the thread is periodic and its execution time within each period is known prior. For our scheduler monitor, the probability of J_i thread missed deadline is calculated from real time values and thread parameters [fig. 3] according to formula (3),

$$p_i(f_i < d_i) = \begin{cases} C_i - C_{R_i} < d_i - t & \frac{C_i - C_{R_i}}{d_i - t} \\ C_i - C_{R_i} \geq d_i - t & 1 \end{cases} \tag{3}$$

where C_{R_i} means real executed time of thread within period and t means current time. This calculation needs deadline values to be updated after each period and of course real executed time to be updated real time. This calculation does not contain influence of scheduling algorithm, the count of other threads and their execution needs. It is precise for one thread on one CPU, otherwise it is only an approximation or simplification. But it is generally usable for all scheduling algorithms.

The monitor could calculates probabilities of periodic threads. So it could calculate the response time of aperiodic threads according to formula (4) [9]. It is assumed that periodic threads have higher priority in a mixed real-time system. This can be computed using other methods described in [10].

$$R_i = f_i - a_i \tag{4}$$

Second way is to calculate a schedulability from online thread parameters. The feasible schedule guarantees that all threads meets their constraints from the scheduler point of view. The feasible schedule is achieved when the schedulability (5) is lesser or equal U_S [6]. Each scheduling algorithm has specific utilization constant, e.g. for Earliest Deadline First algorithm it is one. So this approach assumes that the scheduler algorithm is known prior or can be dynamically switched. The result of this calculation is only a boolean value and it can not be calculated prior because it needs real thread execution time values.

$$\sum_{i=1}^{n} \frac{C_i}{T_i} \leq U_S \tag{5}$$

C. Further work

In our further work, we plan test proposed solution and expand it in advanced scheduler monitor which gives a model, e.g. in VHDL language converted from UPPAAL language, of the system and monitors thread execution according to this model. The thread missed deadline probability will be calculated from differences the real measured values against the model. This monitor will also measure thread execution time according to software flags to gain more precise period value. Calculation of the response time of aperiodic threads is going to be implemented. The probability calculation has to be optimized. The superior classificator algorithm also needs to be optimized. Then we plan to use this monitor as a part of processor redundancy controller.

IX. Conclusion

In this paper, fault tolerant methods are summarized. A new proposal of scheduler monitor is presented. This monitor detects some software failures caused by hardware faults on a real device earlier than other techniques (watchdogs, system monitors, etc.) excluding memory faults caught by Memory Management Unit. It focuses on monitoring periodic threads but can be also used for aperiodic threads. It serves as a supportive system monitor or can be used as a part of system redundancy controller. The proposed scheduler monitor focuses on measuring switching context time which detects thread execution faults earliest. Then detects monitored thread missed deadline before it is actually missed. The secondary feature is comparing real thread execution time to estimated time within each period. This scheduler monitor does not try to replace watchdog or other methods which are very useful to monitor software hangs but presents other possibility to detect faults earlier. Even when it is obvious that if watchdog period is not set to very carefully, the scheduler monitor thread execution time comparator can react earlier. The main future plan is to implement system model which will be compared to current system measured state.

Acknowledgment

The research was financially supported by Brno University of Technology and the European Regional Development Fund under project No. CZ.1.05/2.1.00/01.0014. The abovementioned funds and institutions facilitated efficient performance of the presented research and associated tasks. This work was supported also by the project “TA02010864 - Research and development of motorized ventilation for the human protection against chemical agents, dust and biological agents”, project “TA03020907 - REVYT - Recuperation of the lift loss energy for the lift idle consumption” and project “TA04021653 - Automatic Lift Inspection” granted by Technology Agency of the Czech Republic (TA ČR). Part of the work was supported by project “FR-TI4/642 - MISE - Employment of Modern Intelligent MEMS Sensors for Buildings Automation and Security” granted by Ministry of Industry and Trade of Czech Republic (MPO). Part of the work was carried out with the support of core facilities of CEITEC – Central European Institute of Technology under CEITEC – open access project, ID number LM2011020, funded by the Ministry of Education, Youth and Sports of the Czech Republic under the activity “Projects of major infrastructures for research, development and innovations”. Part of this paper was made possible by grant No. FEKT-S-14-2429 - “The research of new control methods, measurement procedures and intelligent instruments in automation”, and the related financial assistance was provided from the internal science fund of Brno University of Technology.

References

[1] J. Pardo, J. C. Campelo, and J. J. Serrano, “Reliability study of an embedded operating system for industrial applications,” in *GI Jahrestagung (1)’04*, 2004, pp. 83–88.

[2] “Murphy’s law.”

[3] E. Schlaepfer, “Comparison of internal and external watchdog timers,” *Maxim Integrated Products*, June 2008.

[4] M. Pohronská and T. Krajčovič, “Fpga implementation of multiple hardware watchdog timers for enhancing real-time systems security,” *http://www.cs.rochester.edu*, April 2011.

[5] G. Carrozza, M. Cinque, D. Cotroneo, and R. Natella, “Operating system support to detect application hangs,” 2008.

[6] R. Pelánek, “Periodic task scheduling,” *Masarykova Univerzita*, 2015.

[7] L. Otava, “Firmware pro robotické vozítko,” *Vysoké učení technické v Brně*, 2013.

[8] J. Strnadel, “Load-adaptive monitor-driven hardware for preventing embedded real-time systems from overloads caused by excessive interrupt rates,” in *Architecture of Computing Systems - ARCS 2013*, ser. Lecture Notes in Computer Science, ISSN 0302-9743, Vol.7767, vol. 2013, no. 7767. Springer Verlag, 2013, pp. 98–109.

[9] R. Pelánek, “Real time scheduling,” *Masarykova Univerzita*, 2015.

[10] M. Ryu and S.-J. Kim, “Deterministic and statistical deadline guarantees for a mixed set of periodic and aperiodic tasks,” in *In Proc. 9th Real-Time Computing Systems and Applications*, 2003, pp. 72–87.

[11] R. Ramezani and Y. Sedaghat, “An overview of fault tolerance techniques for real-time operating systems,” in *Computer and Knowledge Engineering (ICCKE), 2013 3th International eConference on*. IEEE, 2013.

[12] M. Bar, “The linux scheduler,” *http://www.linuxjournal.com*, 2000.

[13] “http://www.uppaal.org/.”

Emulation of Memristive System Based on Digital Processing with Using Digital Potentiometer

Sang Van Doan[1], Dalibor Biolek[1], Jiri Vavra[1], Zdenek Kolka[2]
[1] Department of Radar Technology
[1]Faculty of Military Technology, UNOB
[2] Dept. of Radioelectronics, BUT
[1,2] Brno, Czechia
[1] sangvan.doan@unob.cz, dalibor.biolek@unob.cz, jiri.vavra@unob.cz
[2] kolka@feec.vutbr.cz

***Abstract*—Memristor is the fourth elementary passive memory element, which was predicted in 1971 by professor Chua and manufactured by researchers at Hewlett-Packard Labs in 2008. In this study, we presented an approach for modelling and emulation of memristive system based on digital signal processing by using digital resistive potentiometer. Hybrid emulator of memristive system was designed as a universal application, which allows, in principle, the emulation of any memristive system just by changing the programed functions. Memulator has been successfully tested for specific memristive system, suggesting its utility in experimental verification of various memristive systems that should be considered to develop in the future.**

Keywords—Memristor; emulator; digital resistive potentiometer; memristive system

I. INTRODUCTION

In 1971, L. O. Chua reported his prediction of a novel passive element [1] which was then manufactured as a memristor by a research group from HP Labs in 2008 [2]. Therefore, apart from three known electrical elements which are resistor, capacitor and inductor, memristor is the fourth passive element which has two ports with variable resistance called memristance. Memristor could be involved in wide application such as digital, analogue and neuromorphic system, promising future advantages in memory technology, and its behavior has been described to be simulated by mathematical models [3, 4, 5]. SPICE, a program for simulation of electric and electronic circuits, is popularly used to build models of memristive systems [6, 7, 8]. In addition, Matlab and Simulink are also applied for simulation of mathematical models of memristive systems [9, 10]. Recently many researchers would like to simulate memristor using hardware together with emulation in computer [11, 12]. Here, we presented a simulator of memristor (in other words memulator), which was designed from idea of bipolar memristive system with threshold in order to simulate the behavior of physical memristor.

Digital emulation method is thought to be not compatible for dipole battery powered applications due to its complicated circuit design and power consumption feature. However, the digital circuit has advantages that it is variable and programmable. In contrast, analog method is known to use less power but lack variability. Therefore, a combination of a programmed single-chip processor and an analog circuit by an analog-digital interface would result in an effective emulation of memristor.

II. MEMULATOR DESIGN

Developed hybrid emulator (Fig. 1a) of memristive systems (memulator) consists of a main board (Fig. 1b) with the microprocessor STM32F303CC (MCU) and 16-bit AD7687 converter (ADC) and an analog interface board (Fig. 1c) which is equipped with a 10-bit digital potentiometer (DP).

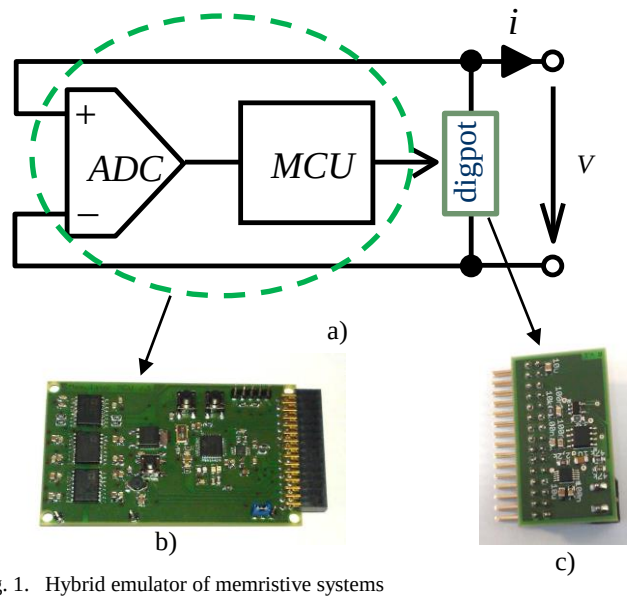


Fig. 1. Hybrid emulator of memristive systems

The emulator reads the voltage between terminals of DP using the ADC. The MCU uses the programmed function to process the data, then sends the data to the DP and simultaneously sends data to a computer for the simulation. Operation principle of memulator was verified by measurement of voltage and current flow through DP.

Memulator was designed for universal-purpose of modeling of a wide class of memristive systems. Here, it was used

specifically to emulate a voltage-controlled memristive system with a threshold, which is promising for modeling of building blocks of neuromorphic systems. This system is described by equations of

$$i = V/R_M, \quad (1)$$

$$dR_M/dt = f(V)w(R_M,V). \quad (2)$$

where, R_M is memristance emulated by the digital potentiometer. Time evolution of memristance is given by the differential equation (2). Its derivation depends on the product of two functions. The function $f(V)$ determines the rate change of memristance at a given voltage (shown in Fig. 2a), the window function $w(R_m, V)$ guarantees the overrun of its physical boundaries R_{on} and R_{off} (shown in Fig. 2b, c). Our formula used for the window w has the form of

$$w(R_m, V) = \underline{1}(V)\ \underline{1}\,(R_{off} - R_m) + \underline{1}\,(-V)\ \underline{1}\,(R_m - R_{on}) \quad (3)$$

where $\underline{1}(\ )$ indicates type of unit step function

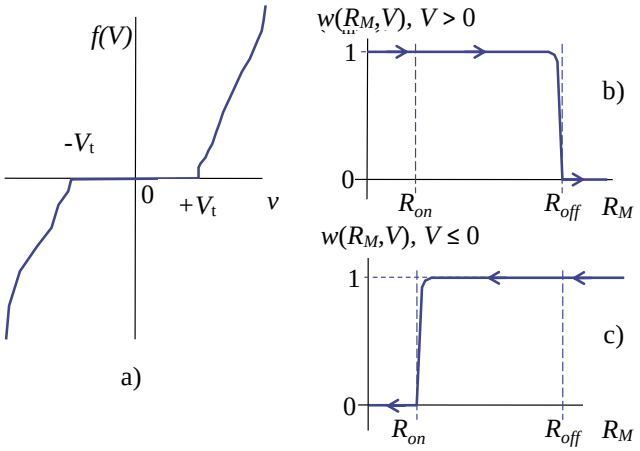


Fig. 2. Function $f(V)$ and $w(\ )$

The realized block diagram of emulator is shown in Fig. 3. The microprocessor (MCU) controls the resistance of the digital potentiometer via the SPI interface, depending on programmed equations (2) and (3). The voltage between points A and B has been read by 16-bit AD converter (ADC). In order to display parameters of voltage and memristance the MCU is connected with computer by USB interface.

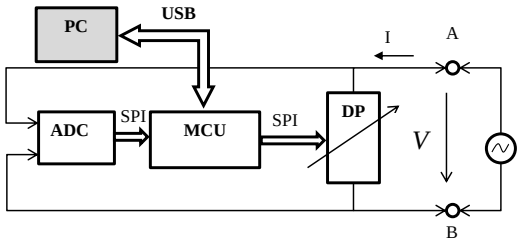


Fig. 3. Block diagram of realized emulator

Further, a program for the function of emulator includes:

```
a) cfg_str_uint(EMU0_RTM_AD0_RAW,
   HAL_SPI_Receive16_fast(&hspi2));
b) int32_t V;
c) if(cfg_get_uint(EMU0_RTM_AD0_RAW)>3
   2767)
d) {V=(65535-
   cfg_get_uint(EMU0_RTM_AD0_RAW))*4.5
   /32768;
e) if (V>=1) { if (R>1000) {R=R-
   10*V*cfg_get_uint(EMU_CNF_TSAMP);}}
   }
f) if(cfg_get_uint(EMU0_RTM_AD0_RAW)<=
   32767)
g) {V=(cfg_get_uint(EMU0_RTM_AD0_RAW))
   *4.5/32768;
h) if (V>1) { if (R<8000) {
   R=R+10*V*cfg_get_uint(EMU0_CNF_TSAM
   P);}}}
i) R=R;
j) cfg_str_uint(EMU0_RTM_R_RAW,
   R*1024/10000);
k) HAL_SPI_Transmit16_fast(&hspi3,
   cfg_get_uint(EMU0_RTM_R_RAW)+1024);
```

AD converter reads the voltage V_{in+} - V_{in-} and converts it into digital data following the Table. 1.

TABLE I. ANALOG-DIGITAL CONVERSION

Analog input voltage = V_{in+} - V_{in-} (Vref = 5V)	**Digital output codes [Hex.]**
+4.999847 V	7FFF
+152.6 μV	0001
0 V	0000
−152.6 μV	FFFF
−4.999847 V	8001
−5 V	8000

In (a) AD converter reads the voltage value. For negative voltages, digital codes are in interval (32767, 65535), For positive voltage, digital codes are in interval <0, 32767>. Codes for data conversion to voltage are in (d) and (g). In program numerical values are used for emulation R_{on} = 1kΩ, R_{off} = 8kΩ and V_t = 1V. Relevant conditions of type "if" corresponding to jumping functions in equation (3) are programmed in (e) and (h) together with the calculation of the resistance of the digital potentiometer. This resistance value is converted back to a digital data code in the (j) and is sent to the digital potentiometer in the command (k).

III. RESULT OF EMULATION

The memulator was connected to a generator of harmonic voltage of variable amplitude and frequency. The amplitude

was set to 4.5 V and the frequency was varied in the interval of 0.5 Hz to 1 kHz. An application program (Fig. 4) was created and used to collect "voltage" data that obtained from the AD converter and the "resistance" of data sent to the digital potentiometer.

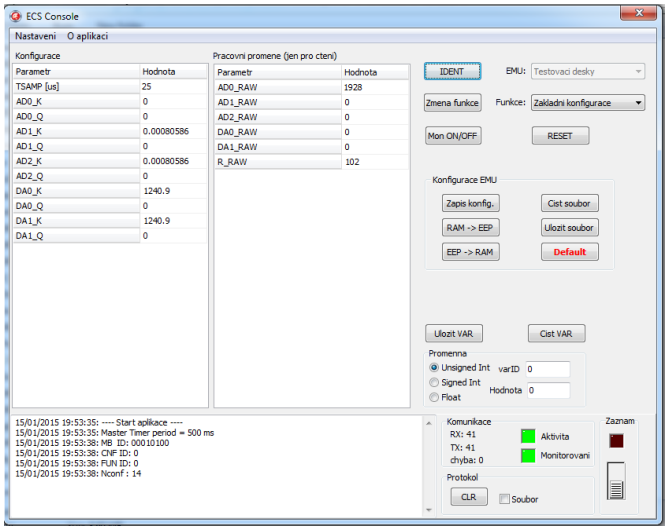


Fig. 4. Application program in computer for emulation of memristor

The selected examples of waveforms of voltage, current, memristance and V-I characteristics of several frequencies are shown in Fig. 5, 6, 7, 8, 9 and 10. The results are consistent with the theoretical knowledge about the behavior of memristive systems, in which V-I characteristics appear a typical hysteresis loop whose area is gradually disappearing at increasing frequencies.

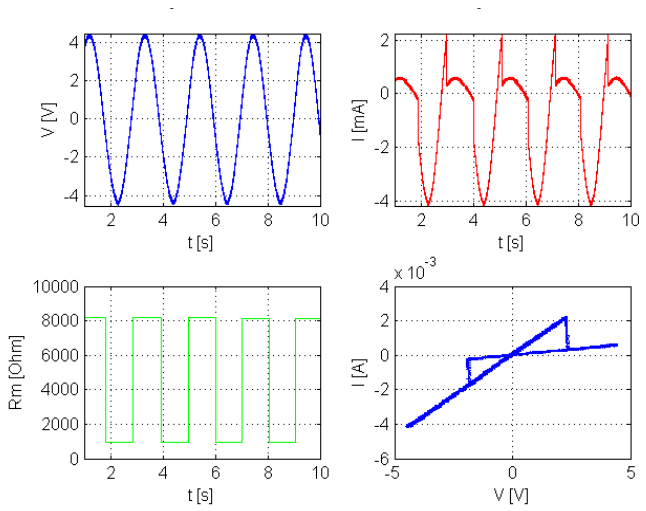


Fig. 5. Graphics of voltage, current, memristance and V-I characteristics for frequency f = 0.5 Hz

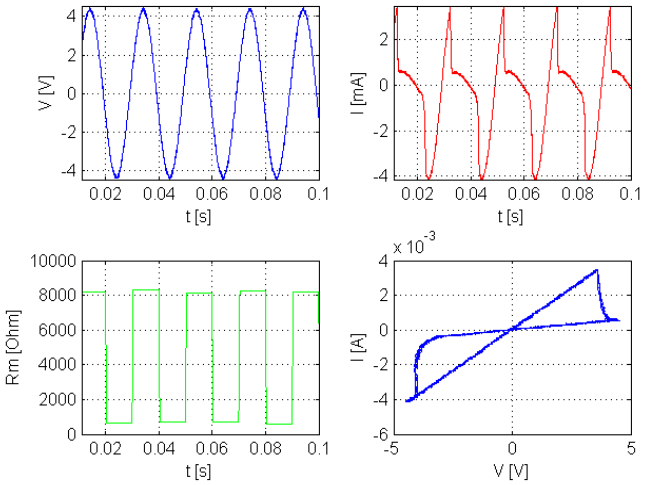


Fig. 6. Graphics of voltage, current, memristance and V-I characteristics for frequency f = 50 Hz

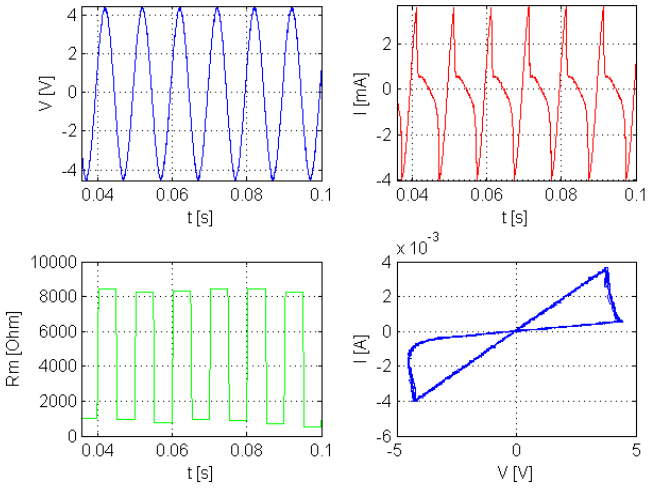


Fig. 7. Graphics of voltage, current, memristance and V-I characteristics for frequency f = 100 Hz

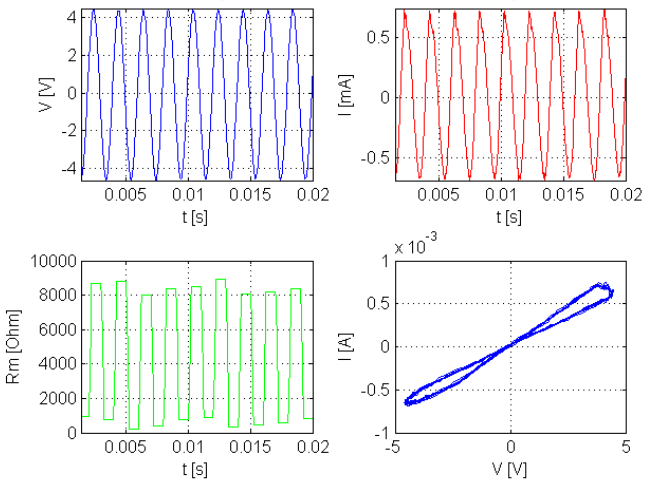


Fig. 8. Graphics of voltage, current, memristance and V-I characteristics for frequency f = 500 Hz

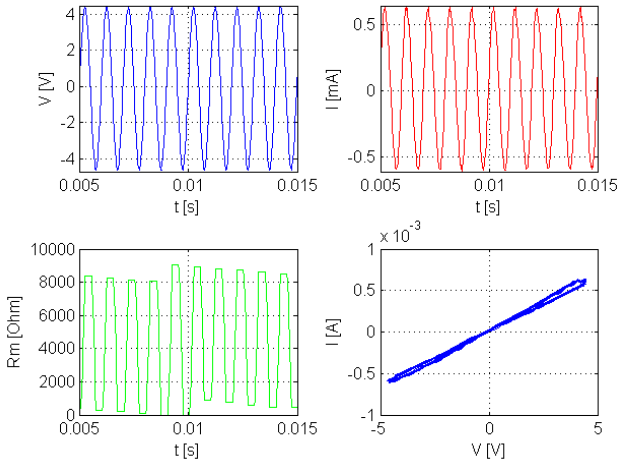


Fig. 9. Graphics of voltage, current, memristance and V-I characteristics for frequency f = 1 kHz

IV. Conclusion

Hybrid emulator of memristive system was designed as a universal application, which allows, in principle, the emulation of any memristive system just by changing the control program. Memulator has been successfully tested for specific memristive system, suggesting its utility in experimental verification of various memristive systems that should be considered to develop in the future.

Acknowledgment

This work has been partially supported by the development project K217 at Brno University of Technology.

References

[1] L. Chua, “Memristor-the missing circuit element,” Circuit Theory, IEEE Transactions on, vol. 18, no. 5, pp. 507 – 519, Sep 1971.

[2] D.B. Strukov, G.S. Snider, and D.R. Stewart, “The missing memristor found,” Nature, vol. 435, pp. 80–83, 5 2008.

[3] A.G. Radwan, M.A. Zidan, K.N. Salama, "HP Memristor mathematical model for periodic signals and DC," Circuits and Systems (MWSCAS), 2010 53rd IEEE International Midwest Symposium, pp. 861-864, 1-4 Aug. 2010.

[4] D. Biolek, Z. Biolek, V. Biolkova, Z. Kolka, "Some fingerprints of ideal memristors," Circuits and Systems (ISCAS), 2013 IEEE International Symposium, pp. 201-204, 19-23 May 2013.

[5] M.E. Fouda, M.A. Khatib, A.G. Radwan, "On the mathematical modeling of series and parallel memcapacitors," Microelectronics (ICM), 2013 25th International Conference, pp. 1-4, 15-18 Dec. 2013.

[6] D. Biolek, Z. Biolek, V. Biolkova, "SPICE modeling of memristive, memcapacitative and meminductive systems," Circuit Theory and Design, 2009. ECCTD 2009. European Conference, pp. 249-252, 23-27 Aug. 2009.

[7] D. Biolek, M.D Ventra, Y.V. Pershin, “Reliable SPICE of memristors, memcapacitors and meminductors,” Radioengineering, vol. 22, no. 4, pp. 945-967, December 2013.

[8] S. Kvatinsky, K. Talisveyberg, D. Fliter, A. Kolodny, U.C. Weiser, E.G. Friedman, "Models of memristors for SPICE simulations," IEEE 27th Convention of Electrical & Electronics Engineers in Israel (IEEEI), pp. 1-5, 14-17 Nov. 2012.

[9] K. Zaplatílek, “Memristor modeling in MATLAB® & Simulink®,” Proceeding of the European Computing Conference, pp. 62-67, 2011.

[10] H. Elgabra, I.A.H. Farhat, A.S.A. Hosani, D. Homouz, B. Mohammad, "Mathematical modeling of a memristor device," Innovations in Information Technology (IIT), 2012 International Conference, pp. 156-161, 18-20 March 2012.

[11] Z. Kolka, D. Biolek, V. Biolkova, “Hybrid modelling and emulation of mem-systems,” International journal of Numerical modelling: Electronic Networks, Devices and Fields, vol. 25, no. 3, pp. 216-225, 2012.

[12] H. Kim, M.P. Sah, Y. Changju, C. Seongik, L.O.Chua, "Memristor Emulator for Memristor Circuit Applications," Circuits and Systems I: Regular Papers, IEEE Transactions, vol.59, no.10, pp. 2422-2431, Oct. 2012.

Parametric Study of HMSIW Linearly Polarized U-Shape Slot Antenna

Patrik Hubka, Jaroslav Lacik
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xhubka00@stud.feec.vutbr.cz, lacik@feec.vutbr.cz

***Abstract*—This paper deals with a parametric study of half-mode substrate integrated waveguide (HMSIW) U-shape slot antenna. It is based on a U-shape slot radiator equipped by two shorts which is etched in the HMSIW top wall. The antenna operates in the frequency band of 10 GHz. The parametric study is provided to better understand behavior of the antenna and demonstrate its typical properties.**

Keywords—Half mode substrate integrated waveguide; slot antenna; parametric study

I. INTRODUCTION

In past years, several HMSIW antennas have been developed. In [4], the first HMSIW slot antenna has been introduced. It is a simple longitudinal radiator with linear polarization. However, the impedance bandwidth of the antenna is only 1.1 %. In [5], a linearly polarized cavity-backed antenna based on HMSIW has been introduced, it has wider impedance bandwidth compared to the antenna presented in [4]. The antenna achieves gain of 6.7 dBi. Another way for obtaining wider impedance bandwidth is to irradiate patch by an open edge of the HMSIW [6] or to exploit a dielectric resonator [7]. In case of the first approach, the antenna has 10 % impedance bandwidth and the peak gain 7.5 dBi. In case of the second approach, the antenna has impedance bandwidth about of 24% and gain 5.5 dBi. In [8], an HMSIW compact slot antenna has been reported. The antenna has Z-shape and reaches the impedance bandwidth of 3.5% and the gain 5.1 dBi.

In these days, a wearable antennas designed for ISM bands attract a lot of attention. A HMSIW textile dual-band antenna was reported in [9]. The measured impedance bandwidth of the antenna is 4.9 % and 5.1 % in 2.4 GHz and 5.8 GHz band, respectively. The measured gain of 4.1 dBi and 5.8 dBi is achieved in 2.4 GHz and 5.8 GHz band, respectively. In [10], a frequency-reconfigurable HMSIW linearly polarized slot antenna was introduced. The antenna operating frequency is controlled by a single varactor diode placed in an interdigital capacitor slot. The resonant frequency can be electrically controlled between 2.99 and 3.59 GHz with impedance bandwidth about 3.2%. However, the gain of proposed antenna is only from -4 to 1.8 dBi in the operating frequency band.

In [11], a HMSIW linearly polarized slot antenna was proposed. In this paper, the parametric study of that antenna is presented. The parametric study is done in order to better understand behavior of the antenna and to demonstrate its properties.

The presented research was supported by the Czech Grant Agency project no. P102/12/1274, by the Czech Ministry of Education in frame of National Sustainability Program under grant LO1401, and by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483. The research is the part of the COST Action IC1301 which is financially supported by the grant of the Czech Ministry of Education no. LD14057. For research, infrastructure of the SIX Center was used.

II. CONFIGURATION AND RESULTS OF ANTENNA

The geometry of the HMSIW U-shaped slot antenna [11] is shown in Fig. 1. The dielectric substrate ARLON Cuclad 217 with relative permittivity ε_r = 2.17, loss tangent tan(δ) = 0.0009 and height h = 1.524 mm is on its both sides covered by metal sheets. The slot is etched at the end of the waveguide at the distance L_{viaX} from its short end in the x-direction, and from the row of vias of HMSIW at the distance L_{viaY} in the y-direction. The slot is shorted by two shorts S_1 and S_2. The first short S_1 influences level of cross-polarization and second short S_2 influences impedance matching. The HMSIW operates in the fundamental mode $TE_{0,5;0}$. The antenna radiates linearly polarized wave with maximum radiation in the normal direction. The antenna is equipped by a HMSIW to GCPW transition.

The antenna was designed for the operating frequency of 10 GHz with the help of time domain solver of CST Microwave Studio 2013 (CST MWS) [11]. The resultant dimensions are summarized in Table I. The antenna was manufactured and measured. The photograph of the manufactured sample of the antenna is depicted in Fig. 2.

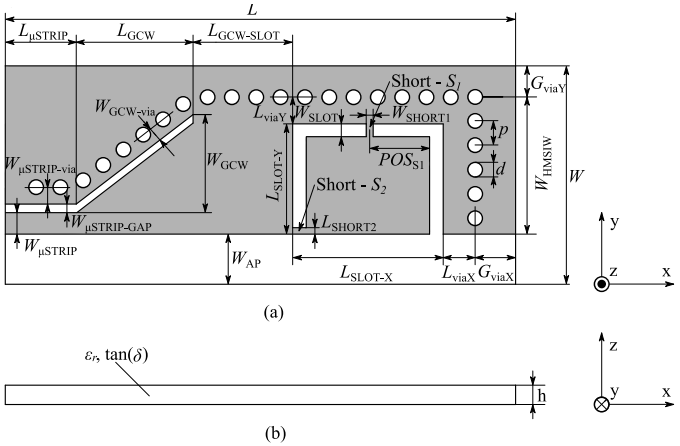


Fig. 1. Top (a) and side (b) view of HMSIW U-shape slot antenna.

TABLE I. DIMENSIONS OF HMSIW U-SHAPE SLOT ANTENNA

	[mm]		[mm]
L	31.2	W_{GCW}	5.2
$L_{\mu STRIP}$	6.4	W_{SLOT}	1.1
L_{GCW}	3.8	W_{SHORT1}	0.5
$L_{GCW\text{-}SLOT}$	3.8	G_{viaX}	3.0
$L_{SLOT\text{-}X}$	12.4	G_{viaY}	3.0
$L_{SLOT\text{-}Y}$	7.1	W_{HMSIW}	9.0
L_{SHORT2}	0.4	$W_{\mu STRIP}$	1.9
L_{viaX}	1.7	W_{AP}	3.0
L_{viaY}	1.9	POS_{S1}	5.0
W	14.4	h	1.524
$W_{GCW\text{-}via}$	1.5	d	1.4
$W_{\mu STRIP\text{-}via}$	1.9	p	2.5
$W_{\mu STRIP\text{-}GAP}$	0.7		

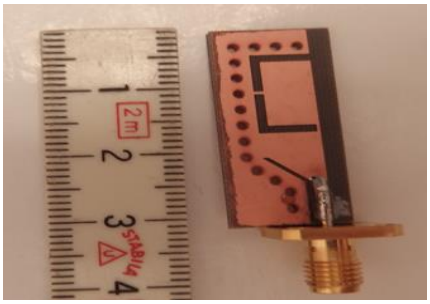


Fig. 2. Fabricated sample of HMSIW U-shape slot antenna.

Detail description of simulated and measured results is presented in [11]. Here graphs of simulated and measured results are shortly presented. The simulated and measured impedance bandwidth (Fig. 3) for the reflection coefficient less than –10 dB of the antenna is 3.0% and 3.2%, respectively.

The antenna radiates linearly polarized wave in the broadside direction. The simulated and measured radiation patterns are depicted in Fig. 4. The measured gain of the antenna is 6 dBi.

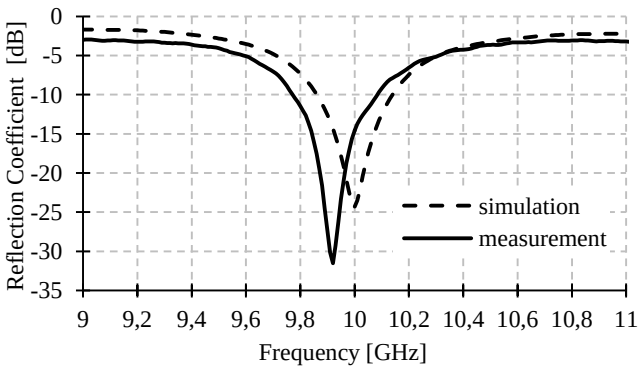
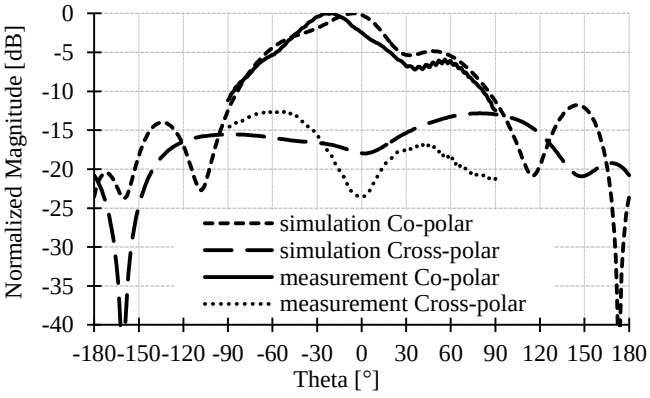
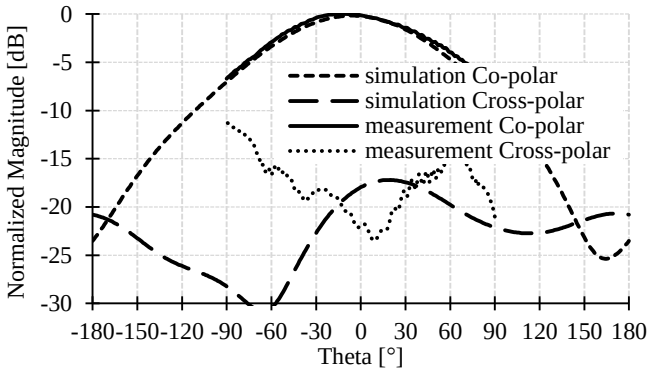


Fig. 3. Reflection coefficient of HMSIW U-shape slot antenna.



(a)



(b)

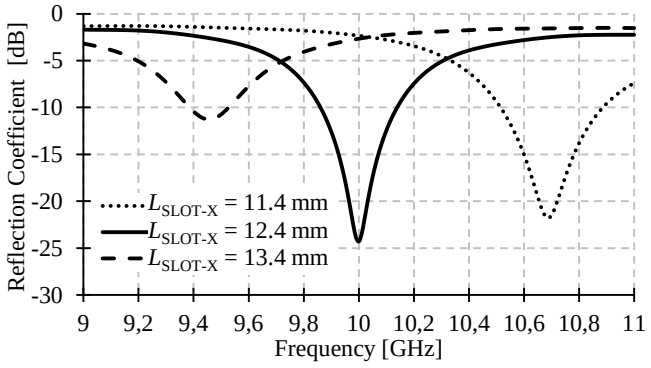
Fig. 4. Normalized radiation pattern of HMSIW U-shape slot antenna at 10 GHz: (a) E-plane (xz-plane), (b) H-plane (yz-plane).

III. PARAMETRIC STUDY

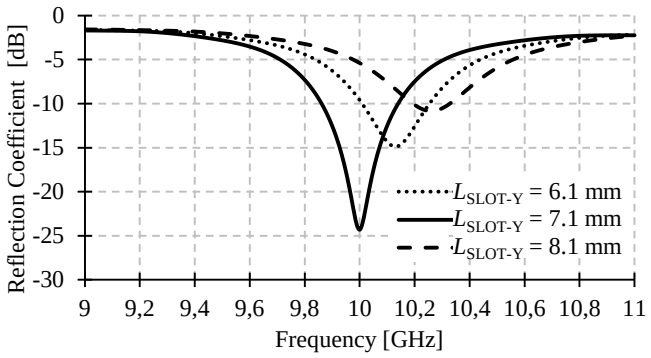
The parametric study is carried out with the help CST Microwave Studio 2013 and demonstrates the effect of antenna geometrical parameters on the reflection coefficient, co- and cross-polarization level. During this study all parameters given by Table I. are kept and only one parameter is changed.

In the Fig. 5(a) and (b) the effect of the length of the slot $L_{SLOT\text{-}X}$ in x-direction and the length of the slot $L_{SLOT\text{-}Y}$ in y-direction are depicted. The length of the slot in x-direction influences mainly the operating frequency whereas the length of the slot in y-direction influences mainly the impedance matching of the antenna. The variation of width of slot W_{slot} is depicted in Fig. 6. We can see that it strongly influences the operating frequency band of the antenna, The effect of the position of the short S_1 is depicted in Fig. 7. Obviously, it slightly influences operating frequency and impedance matching. The effect of the height of dielectric substrate is depicted in Fig. 8. As the height of the substrate decreases, the operating frequency slightly shifts towards higher frequencies and the impedance bandwidth is narrower. Considering the above facts we can conclude that the operating frequency and impedance matching of the antenna depends on the length of the slot of the antenna in the x- and y-directions, and on the width of the slot. By combining these parameters, the antenna can be tuned to desirable operating frequency band.

The parameters used in parametric study have only minimal influences on co- and cross-polarization.



(a)



(b)

Fig. 5. Effect of length L_{SLOT-X} of slot in x-direction (a) and the length L_{SLOT-Y} of slot in y-direction on reflection coefficient.

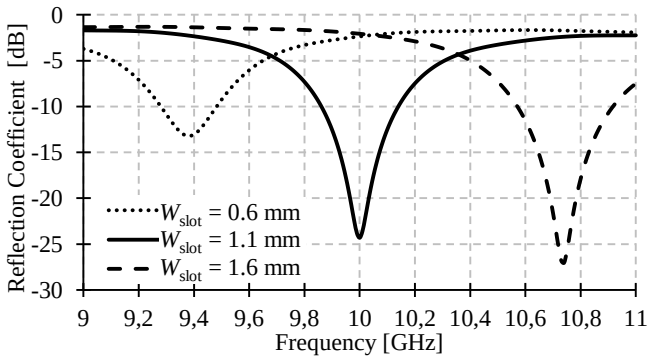


Fig. 6. Effect of width of slot W_{slot} on reflection coefficient.

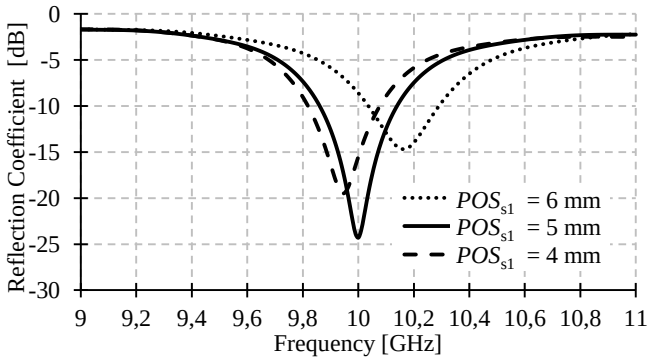


Fig. 7. Effect of position of short S_1 on reflection coefficient.

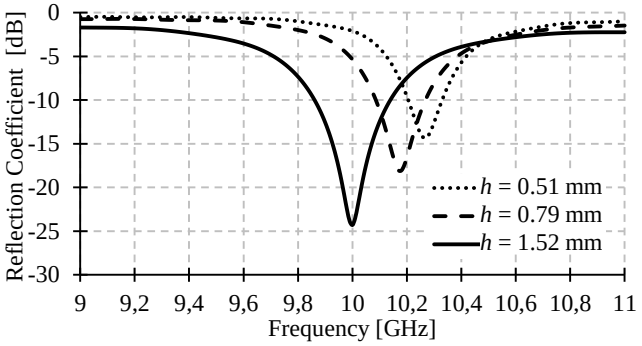


Fig. 8. Effect of width of dielectric substrate on reflection coefficient.

IV. Conclusion

In this paper, a parametric study of HMSIW linearly polarized U-shape slot antenna has been presented. Based on the parametric study, the antenna can be retuned at desirable band for desired applications e.g. body area network, energy harvesting or wireless power transmission.

Future work will be focused on detail description of presented antenna. Design guide will be given.

References

[1] Y. Ding and K. Wu, "A 4 × 4 Ridge Substrate Integrated Waveguide (RSIW) Slot Array Antenna", IEEE Antennas and Wireless Propagation Letters, 2008, vol. 8, p. 561 – 564.

[2] B. Sanz Izquierdo, P. R. Young, N. Grigoropoulos, J. C. Batchelor and R. J. Langley, "Slot Antenna on C Type Compact Substrate Integrated Waveguide", 2005 European microwave Conference, 2005, vol. 1, p. 1 – 4.

[3] Y. Ding and K. Wu, "T-Type Folded Substrate Integrated Waveguide (TFSIW) Slot Array Antenna", IEEE Transactions on Antennas and Propagation, 2010, vol. 58, no. 5, p. 1792 – 1795.

[4] L. Wu, "Half Mode Substrate Integrated Waveguide Slot Antenna", IEEE Antennas and Propagation Society International Symposium, 2009, p. 1 – 4.

[5] S.A. Razavi and M.H. Neshati, "Development of a Linearly Polarized Cavity-Backed Antenna Using HMSIW Technique", IEEE Antennas and Wireless Propagation Letters. 2012, vol. 11, pp. 1307 – 1310.

[6] H. Dashati, M.H Neshati, "Development of Low-Profile Patch and Semi-Circular SIW Cavity Hybrid Antennas", IEEE Transactions on Antennas and Propagation. 2014, vol. 62, no.9, pp. 4481–4488.

[7] Q. Lai, Ch. Fumeaux, W. Hong, and R. Vahldieck, "60 GHz Aperture-Coupled Dielectric Resonator Antennas Fed by a Half-Mode Substrate Integrated Waveguide", IEEE Transactions on Antennas and Propagation, 2010, vol. 58, no. 6, p. 1856 – 1864.

[8] B. Rana, S.K. Parui,"Half-Mode Substrate Integrated Waveguide Fed Compact Slot Antenna", 2013, 2013 IEEE Applied Electromagnetics Conference (AEMC), pp. 1–2.

[9] S. Agneessens, H.Rogier, "Compact Half Diamond Dual-Band Textile HMSIW On-Body Antenna", IEEE Transactions on Antennas and Propagation. 2014, vol. 62, no. 5, pp. 2374–2381.

[10] S. Sam, S. Lim,"Compact Frequency-Reconfigurable Half-Mode Substrate-Integrated Waveguide Antenna", IEEE Antennas and Wireless Propagation Letters. 2013, vol. 12, pp. 951–954.

[11] P. Hubka, J. Lacik, "Linearly Polarized HMSIW U-Slot Antenna", Microwave and Radio Electronics Week MAREW 2015. 2015, pp. 23-26.

Comparison of Methods for AR Coefficients Estimation

Eva Klejmova
Department of Radio Electronics
Brno University of Technology
Technicka 3082/12, 616 00 Brno, Czech Republic
Email: klejmova@phd.feec.vutbr.cz

***Abstract*—The paper aim is to give recommendation for work with methods used for estimation of coefficients of autoregressive process. We investigate Burg and Yule-Walker method using Monte Carlo simulations. Performance of methods is evaluated by precision of estimation for signals of short and long length. The influence of lag order is also examined and compared with Akaike information criterion. The results are presented in graphical form and briefly discussed. Taking these results into account, Yule-Walker method shows better performance in case of long length signals and in case of overvalued lag order. Burg method provide better results in case of short length signal and undervalued lag order.**

***Keywords*—AR process, Yule-Walker method, Burg method, spectral analysis, lag order, AIC**

I. Introduction

Autoregressive process as a method of time-frequency analysis is applied in many fields of science. It is a useful instrument for signal description in many different application areas such as engineering, biology and medicine, or economics [1], [4]. Important factor for best possible estimation is the performance of selected method for AR coefficients estimation. Choosing of inappropriate method, depending on the characteristics of the input signal, can result to inadequate spectrum representation [5].

On the basis of simulation, we analyze the behavior of Yule-Walker and Burg method via Monte Carlo simulations. We examine the advantages and disadvantages, and formulate recommendations for its usage. We also investigate influence of the lag order.

II. Methodical background

A. Autoregressive (AR) process

AR process is a method that uses a parametric approach and creates a model generating an input signal. It can be used for spectral estimation based on the evidence that any linear process can be approximated by AR process of adequate order. To compute spectrum, analyzed signal $s(n)$ is regarded as the output of a linear filter influenced by white noise w with variance σ_w^2. The output spectrum can then be described as [2]

$$S_s(f) = \left|H\left(e^{j2\pi fT}\right)\right|^2 \sigma_w^2, \tag{1}$$

where $H\left(e^{j2\pi fT}\right)$ is a linear time variant filter.

The estimation of time-frequency model is then done according to formula [5]:

$$\widehat{S}_s(f) = \frac{\widehat{\sigma}_w^2}{\left|1 + \sum_{k=1}^{p} \widehat{a_p}(k) e^{-j2\pi fk}\right|^2}, \tag{2}$$

where $\widehat{a_p}(k)$ are estimates of the AR parameters, p is the lag order and σ_w^2 is white noise variance.

The key for spectrum estimation as accurate as possible is to obtain AR model parameters. Good results provides ordinary least squares estimation (OLS). However this method leads to multiple transposition and multiplying of large matrices making OLS very complex for computation especially for large value of lag order. To prevent this Burg method and Yule-Walker method are commonly used [5]. Yule-Walker method is similar to OLS but it estimates the autocorrelation from the data which is used to solve AR model parameters [5]. Burg method is based on least-squares lattice method and use minimization of the forward and backward errors in linear predictors [5].

Another key aspect for AR modeling is selection of adequate lag order p. Selection of low level order cause excessive smoothing of the spectrum. Furthermore, selecting the level of p too high leads to significant highlighting of non-significant spectral coefficients. Several information criteria can be used to ensure optimal selection. One of them, most commonly used and not computationally demanding is Akaike information criterion (AIC) [5]:

$$AIC(p) = ln\widehat{\sigma}_{wp}^2 + \frac{2p}{N}, \tag{3}$$

where σ_{wp}^2 is estimated variance of linear prediction error, p is the order and N is length of the signal. The order is selected as optimal when AIC reaches minimum.

B. Monte Carlo Analysis

The term Monte Carlo includes group of computer based statistical methods that can be applied to different scientific areas requiring statistical testing (e.g. statistical physics, molecular simulation,...). The term itself comes from name of the city Monte Carlo, known for number of casinos that represent symbol of random number generator. This suggests that the Monte Carlo experiment tries to replicate original data by generating large number of random realization of a stochastic process. Resulting simulated process is approximated by

averaging these sub realizations in some way. Monte Carlo methods usually contain following steps [3]:

1) Define required model
2) Generate random inputs
3) Perform a deterministic computation on the inputs
4) Analyze the result (using summary statistic)

Methods used to obtain AR model parameters contain white noise and therefor show different statistical properties when large variety of random processes is available. In this case Monte Carlo analysis is suitable tool to assess performance and quality of applied methods.

III. Application

A. Simulation

For Monte Carlo simulation autoregressive process of order 20 was used. This order was chosen to provide sufficient complexity of signal (e.g. several close spectral components) with reasonable computational requirements. Selected AR coefficients are shown in Table I.

TABLE I. AR coefficients

a_1	1	a_8	0.16296	a_{15}	-0.23397
a_2	-0.99109	a_9	0.68828	a_{16}	0.04167
a_3	0.45403	a_{10}	-0.06771	a_{17}	0.21317
a_4	0.01037	a_{11}	-0.58740	a_{18}	0.05886
a_5	0.83333	a_{12}	-0.04689	a_{19}	-0.08189
a_6	-0.06972	a_{13}	0.36074	a_{20}	0.07825
a_7	-0.67935	a_{14}	0.00583	a_{21}	0.16283

Using these coefficients signal in time domain was created according to equation 1. We selected two signal lengths to examine influence of sample size. Taking into account the AR order of 20, sample size 100 was selected for short signal and sample size of 1000 for long signal. The Monte Carlo analysis was done for three number of simulations, $N = 10, 100, 1000$. For each simulation Burg and Yule-Walker method was applied to obtain estimated coefficients and spectrum estimation. To asses performance of both methods in case of selecting non optimal lag order, the computation was done for lag order in range of 3 to 60 in case of short signal and in range of 3 to 200 in case of long signal. Mean square error (MSE) was calculated using theoretical and estimated spectrum representation for each process. This was done according following formula [5].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{Y}_i - Y_i\right)^2, \tag{4}$$

where $\widehat{Y}_i$ is value of theoretical spectral coefficient Y_i is its estimated counterpart. The lower the resulting number is the more accurate is the estimation.

B. Results

Values of MSE for both methods for optimal lag order are shown in Table II in dB. It is obvious that for larger number of simulations the estimation of spectrum is more precise. Better results are also for the signal of longer length. The MSE is smaller in case of Yule-Walker method in in all considered cases.

TABLE II. Values of MSE in dB

	10 simulations		100 simulations		1000 simulations	
signal length	short	long	short	long	short	long
Yule-Walker method	15,69	6,01	13,87	5,67	12,93	5,38
Burg method	22,27	10,79	20,27	10,33	19,46	10,22

In Figures 1- 4 lag order was set to optimal value of 20. The frequency of x-axis is relative to sampling frequency (value of $0,5$ in figures corresponds to $0,5f_s$). In Figure 1 and 2 is theoretical and estimated spectrum of short signal for 10 and 1000 simulations. We can see that larger number of simulation leads to more precise estimation and smoothness of the graph. However neither of examined methods was able to adequately describe rapid changes of spectral peaks. We can also see that the Burg method tends to be more accurate in capturing the shape of the graph.

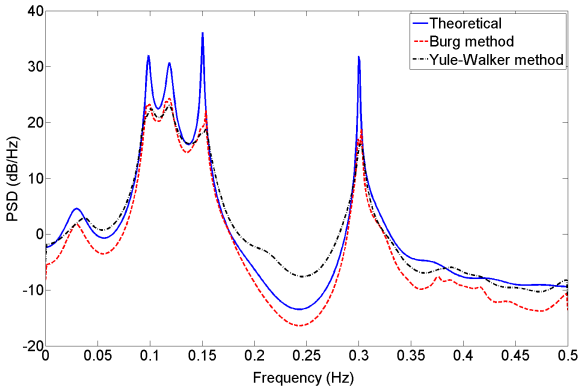


Fig. 1. Theoretical and estimated spectrum of short signal, 10 simulations

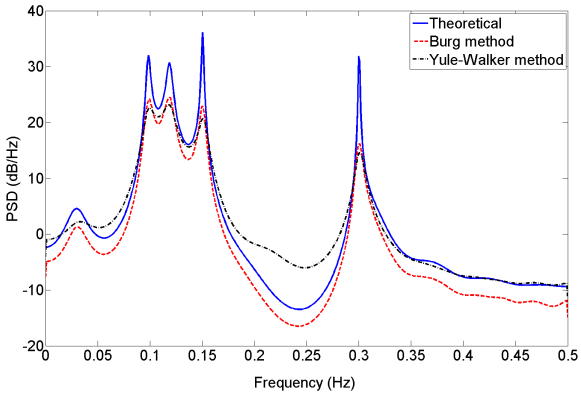


Fig. 2. Theoretical and estimated spectrum of short signal, 1000 simulations

In Figure 3 and 4 is theoretical and estimated spectrum of long signal for 10 and 1000 simulations. We can see that larger number of simulation leads to more precise estimation and smoothness of the graph. We can also see that the difference for both methods is not so significant as for short the signal. Both methods were able to capture all significant spectral peaks.

In Figure 5 and 6 we can see influence of selected lag order on precision of coefficient estimation. For signal of length 1000 the minimum for AIC and for both methods corresponds to the order of original AR process. We can see that the

trend decrease rapidly to optimal lag order and than linearly increases. In case of the short signal, the minimum lays on lower value that the optimal order. The trend also decrease rapidly to minimum. Then the value remains almost constant for Yule-Walker method and increases exponentially for AIC and Burg.

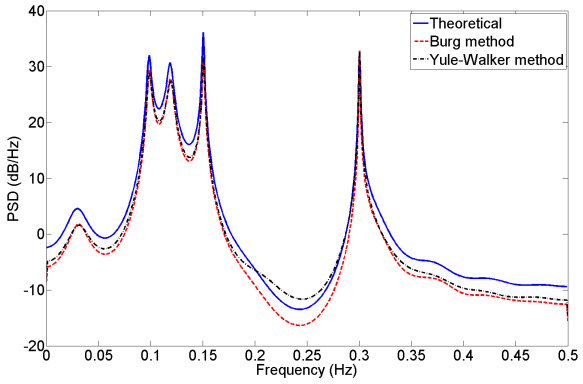


Fig. 3. Theoretical and estimated spectrum of long signal, 10 simulations

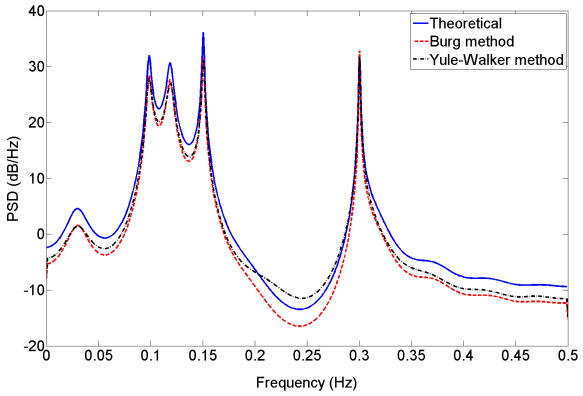


Fig. 4. Theoretical and estimated spectrum of long signal, 1000 simulations

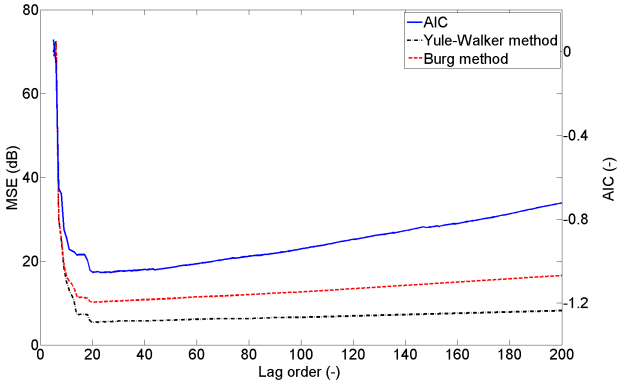


Fig. 5. MSE and AIC of estimated spectrum for different lag order, long signal, 1000 simulations

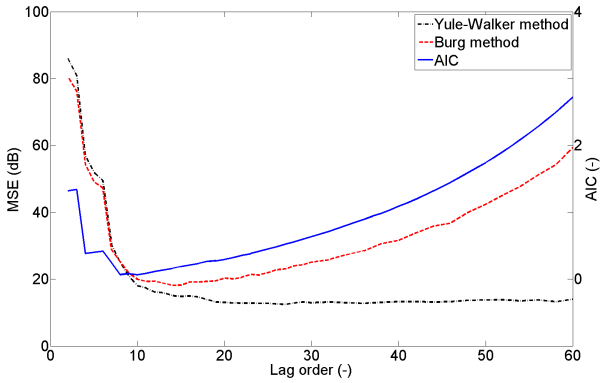


Fig. 6. MSE and AIC of estimated spectrum for different lag order, short signal, 1000 simulations

IV. Conclusion

In this paper we compared Yule-Walker and Burg method for power spectrum estimation. These methods were used on simulated AR signal with order 20 and their performance was compared based on precision of estimation of spectrum. Yule-Walker method shows better results for long signal and performs significantly better for overvalued lag order. Burg method shows better peak detection for short signal. Furthermore this method performs better in case of undervalued lag order, for overvalued lag order the MSE is significantly larger than Yule-Walker. Therefor we recommend using Yule-Walker method for long signals where we can afford to choose higher lag order and Burg method for short signals where we tend to choose lower lag order. Furthermore we found that in case of long signal, minimum of AIC and minimum of MSE corresponds with lag order of original AR process for both methods. For short signal the minimum of AIC and the minimum of MSE approximately corresponds.

Acknowledgment

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401, the BUT Internal Grant Agency under project no. FEKT-S-14-2281. For research, infrastructure of the SIX Center was used.

References

[1] Blumenstein, J., Pomenkova, J., Marsalek, R.: *Comparative Study Of Time-Frequency Analysis Approaches With Application To Economic Indicators*. In 26th European Conference on Modelling and Simulation ECMS 2012. Germany: 2012. pp. 291-297

[2] Ján, J.:*Číslicová filtrace, analýza a restaurace signálu*. VUTIUM Brno, 2002.

[3] Robert, C. P.; Casella, G.: *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer, 2004.

[4] Sebesta, V., Marsalek, R., Pomenkova, J.: *The Modified Empirical Mode Decomposition Method For Analysing The Cyclical Behavior Of Time Series*. In 27th European Conference on Modelling and Simulation ECMS 2013. Norway: 2013. pp. 288-292

[5] Proakis, J. G., Rader, Ch. M, Ling, F. L., Nikias, Ch. L., Moonen, M. and Proudler, J. K.: *Algorithms for Statistical Signal Processing*. Prentice Hall, 2002.

Stability of Stochastic Differential Systems

Marie Klimesova
Department of Mathematics, Faculty of Electrical Engineering and Communication
Brno University of Technology
Technická 2848/8, Žabovřesky, 61600, Brno, Czech republic
Email: xklime01@stud.feec.vutbr.cz

***Abstract*—This paper surveys the elementary theory of stability of solution of stochastic differential equations (SDEs) and systems. It can be used in population models, epidemic and genetic models in medicine and biology, meteorology models, in physical science, for analysis in economy, financial mathematics, etc. The article starts with a review of the stochastic theory. Then, conditions are deduced for the asymptotic mean square stability of the zero solution of stochastic equation with one-dimensional Brownian motion and system with two-dimensional Brownian motion. It is used a Lyapunov function. The method of Lyapunov functions for the analysis of behavior of SDEs provides some very useful information in the study of stability properties for concrete stochastic dynamical systems, conditions of existence the stationary solutions of SDEs and related problems.**

***Keywords*—Brownian motion, stochastic differential equation, Lyapunov function, stochastic Lyapunov function, stability, stochastic stability.**

I. Introduction

Stochastic modeling has come to play an important role in many branches of science and industry where more and more people have encountered stochastic differential equations. Stochastic model can be used to solve problem which evinces by accident, noise, etc. Definition of probability spaces, stochastic process, stochastic differential equation and an existence and uniqueness of solution of these equations, were mentioned in Student EEICT 2014 [14]. It was taken from B. Øksendal [12] and E. Kolářová [8]. In this paper we focus on the description of the stochastic stability. The stability theory was introduced by R. Z. Khasminskii [7]. The basic principles of various types of stochastic systems are described by X.Mao [9]. In the paper we derived sufficient conditions for general system of the zero solution of the stochastic differential equation using Lyapunov function.

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition 1: Let $B_t = (B_1(t), ..., B_m(t))$ be m-dimensional Brownian motion and $b : [0,T] \times R^n \to R^n$, $\sigma : [0,T] \times R^n \to R^{n\times m}$ be measurable functions. Then the process $X_t = (X_1(t), ..., X_m(t))$, $t \in [0,T]$ is the solution of the stochastic differential equation

$$dX_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t, \tag{1}$$

$b(t, X_t) \in R$, $\sigma(t, X_t)W_t \in R$. After the integration of equation (1) we give the solution of the SDE

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s. \tag{2}$$

Assume that for every initial value $X_t(0) = X_0 \in R^n$, there exists a unique global solution which is denoted by $X(t; t_0, X_0)$. So equation (1) has the solution $X_t(0) \equiv 0$ corresponding to the initial value $X_t(0) = 0$. This solution is called the **trivial solution** or equilibrium position.

II. Stability of Stochastic Differential Equations

In 1892, A.M. Lyapunov developed a methods for determining stability without solving the equation. We are used the second Lyapunov method: Let K denote the family of all continuous nondecreasing functions $\mu : R_+ \to R_+$ such that $\mu(0) = 0$ and $\mu(r) > 0$ if $r > 0$. For $h > 0$, let $S_h = \{x \in R^n : |x| < h\}$. A continuous function $V(x,t)$ defined on $S_h \times [t_0, \infty)$ is said to be **positive-definite** (in the sense of Lyapunov) if $V(0,t) \equiv 0$ and, for some $\mu \in K$,

$$V(x,t) \geq \mu(|x|) \quad \text{for all } (x,t) \in S_h \times [t_0, \infty).$$

A function $V(x,t)$ is said to be **negative-definite** if $-V$ is positive-definite. A continuous non-negative function $V(x,t)$ is said to be **decrescent** (i.e. to have an arbitrarily small upper bound) if for some $\mu \in K$,

$$V(x,t) \leq \mu(|x|) \quad \text{for all } (x,t) \in S_h \times [t_0, \infty).$$

A function $V(x,t)$ defined on $R^n \times [t_0, \infty)$ is said to be **radially unbounded** if

$$\lim_{|x|\to\infty} \left(\inf_{t\geq t_0} V(x,t) \right) = \infty.$$

Let $C^{1,1}(S_h \times [t_0, \infty), R_+)$ denote the family of all continuous functions $V(x,t)$ from $S_h \times [t_0, \infty)$ to R_+ with continuous first partial derivatives with respect to every component of x and to t. Then $v(t) = V(t, X_t)$ represents a function of t with the derivative

$$\begin{aligned} \dot{v}(t) &= V_t(t, X_t) + V_x(t, X_t)b(t, X_t) \\ &= \frac{\partial V}{\partial t}(t, X_t) + \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}(t, X_t)b_i(t, X_t). \end{aligned}$$

If $\dot{v}(t) \leq 0$, then $v(t)$ will not increase so the distance of X_t from the equilibrium point measured by $V(t, X_t)$ does not increase. If $\dot{v}(t) < 0$, then $v(t)$ will decrease to zero so the distance will decrease to zero, that is $X_t \to 0$.

Theorem 1: (Lyapunov theorem) If there exists a positive-definite function $V(x,t) \in C^{1,1}(S_h \times [t_0, \infty), R_+)$ such that

$$\dot{V}(x,t) := V_t(t, X_t) + V_x(t, X_t)b(t, X_t) \leq 0$$

for all $(x,t) \in S_h \times [t_0, \infty)$, then the trivial solution is stable. If there exists a positive-definite decrescent function $V(x,t) \in C^{1,1}(S_h \times [t_0,\infty), R_+)$ such that $\dot{V}(x,t)$ is negative-definite, then the trivial solution is asymptotically stable.

Definition 2: The trivial solution of stochastic equation (1) is said to be

(i) stochastically **stable** or **stable in probability** if for every pair of $\epsilon \in (0,1)$ and $r > 0$, there exists $\delta = \delta(\epsilon, r, t_0) > 0$ such that $\mathrm{P}\{|x(t,t_0,x_0)| < r\} \geq 1-\epsilon$ for all $t \geq t_0$, whenever $|x_0| < \delta$. Otherwise, it is said to be stochastically **unstable**.

(ii) stochastically **asymptotically stable** if it is stochastically stable and, moreover, for every $\epsilon \in (0,1)$, there exists $\delta_0 = \delta_0(\epsilon, t_0) > 0$ such that $\mathrm{P}\{\lim_{t\to\infty} x(t,t_0,x_0) = 0\} \geq 1-\epsilon$ whenever $|x_0| < \delta_0$.

(iii) stochastically **asymptotically stable in the large** if it is stochastically stable and, moreover, for all $x_0 \in R^n$ $\mathrm{P}\{\lim_{t\to\infty} x(t,t_0,x_0) = 0\} = 1$.

Suppose one would like to let the initial value be a random variable. It should also be pointed out that when $\sigma^{(x,t)} = 0$, these definitions reduce to the corresponding deterministic ones. We now extend the Lyapunov Theorem (1) to the stochastic case. Let $0 < h \leq \infty$. Denote by $C^{2,1}(S_h \times R_+, R_+)$ the family of all nonnegative functions $V(x,t)$ defined on $S_h \times R_+$ such that they are continuously twice differentiable in x and once in t. Define the differential operator L associated with equation (1) by

$$\begin{aligned} L &= \frac{\partial}{\partial t} + \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(t, X_t) b_i(x,t) \\ &+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} \left[\sigma(x,t)\sigma^T(x,t)\right]_{ij}. \end{aligned}$$

The inequality $\dot{V}(x,t) \leq 0$ will be replaced by $LV(x,t) \leq 0$ in order to get the stochastic stability assertions.

Theorem 2: If there exists a positive-definite

(i) function $V(x,t) \in C^{2,1}(S_h \times [t_0,\infty), R_+)$ such that $LV(x,t) \leq 0$ for all $(x,t) \in S_h \times [t_0,\infty)$, then the trivial solution of equation (1) is stochastically **stable**.

(ii) decrescent function $V(x,t) \in C^{2,1}(S_h \times [t_0,\infty), R_+)$ such that $LV(x,t)$ is negative-definite, then the trivial solution of equation (1) is stochastically **asymptotically stable**.

(iii) decrescent radially unbounded function $V(x,t) \in C^{2,1}(R^n \times [t_0,\infty), R_+)$ such that $LV(x,t)$ is negative-definite, then the trivial solution of equation (1) is stochastically **asymptotically stable in the large.**

Proof: [9], pp. 111.

III. Main results

We have a homogenous linear stochatic differencial equation

$$dX_t = A(X_t)dt + GdB_t, \tag{3}$$

where $X_t = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $G = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, $B_t = \begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}$.

Definition 3: Lyapunov quadratic function V is given

$$V(X_t) = X_t^T Q\, X_t, \tag{4}$$

where $Q = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$ is a symmetric positive-definite matrix, i.e. $\alpha > 0, \alpha^2 - \beta^2 > 0$.

We compute derivation of Lyapunov function of equation (3)

$$\begin{aligned} dV(X_t) &= V(X_t + dX_t) - V(X_t) = (X_t^T + (AX_t)^T dt \\ &+ (GdB_t)^T)Q(X_t + AX_t dt + GdB_t) - X_t^T Q X_t \\ &= X_t^T Q X_t + X_t^T Q A X_t dt + X_t^T Q G dB_t \\ &+ (AX_t)^T dt Q X_t + (AX_t)^T dt Q A X_t dt \\ &+ (AX_t)^T dt Q G dB_t + (GdB_t)^T Q X_t \\ &+ (GdB_t)^T Q A X_t dt + (GdB_t)^T Q G dB_t \\ &- X_t^T Q X_t = X_t^T Q X_t + X_t^T Q A X_t dt \\ &+ X_t^T Q G dB_t + X_t^T A^T dt Q X_t \\ &+ X_t^T A^T dt Q A X_t dt + X_t^T A^T dt Q G dB_t \\ &+ dB_t^T G^T Q X_t + dB_t^T G^T Q A X_t dt \\ &+ dB_t^T G^T Q G dB_t - X_t^T Q X_t. \end{aligned}$$

We use the rules:

$$dt \cdot dt = dt \cdot dB_1(t) = dt \cdot dB_2(t) = dB_1(t) \cdot dB_2(t) = 0,$$
$$dB_1(t) \cdot dB_1(t) = dB_2(t) \cdot dB_2(t) = dt.$$

After modyfying we get

$$\begin{aligned} dV(X_t) &= X_t^T Q A X_t dt + X_t^T Q G dB_t + X_t^T A^T dt Q X_t \\ &+ dB_t^T G^T Q X_t + dB_t^T G^T Q G dB_t. \end{aligned}$$

In matrix form

$$\begin{aligned} &\mathrm{dV}\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} \\ &= \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix} \\ &+ \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}^T \begin{pmatrix} a & b \\ c & d \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} \\ &+ \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ &\cdot \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \end{aligned}$$

We determine

$$\begin{aligned} &\begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = M \\ &= \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}. \end{aligned}$$

Then we have

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix}^T \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}$$
$$= \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}^T \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}$$
$$= m_1 dB_1(t) dB_1(t) + m_3 dB_2(t) dB_1(t) + m_2 dB_1(t) B_2(t)$$
$$+ m_4 dB_2(t) B_2(t) = m_1 dt + m_4 dt = tr(M) dt,$$

where $tr(M)$ is trace of square matrix M.

We get

$$\begin{aligned} dV(X_t) &= 2\left[(a\alpha + c\beta)X_1^2(t) + (d\alpha + b\beta)X_2^2(t)\right. \\ &+ ((b+c)\alpha + (a+d)\beta)X_1X_2(t) \\ &+ \left.(2\beta(hf+eg) + \alpha(e^2+f^2+g^2+h^2))\right] dt \\ &+ 2\left[(e\alpha + g\beta)X_1(t) + (g\alpha + e\beta)X_2(t)\right] dB_1(t) \\ &+ 2\left[(f\alpha + h\beta)X_1(t) + (h\alpha + f\beta)X_2(t)\right] dB_2(t). \end{aligned}$$

We apply expectation $E\{dV(X_t)\}$

$$\begin{aligned} E\{dV(X_t)\} &= 2\left[(a\alpha + c\beta)X_1^2(t) + (d\alpha + b\beta)X_2^2(t)\right. \\ &+ ((b+c)\alpha + (a+d)\beta)X_1(t)X_2(t) \\ &+ \left.(2\beta(hf+eg) + \alpha(e^2+f^2+g^2+h^2))\right] dt \\ &= LV dt. \end{aligned}$$

For $Q = I$ we get

$$\begin{aligned} LV &= 2\left[aX_1^2(t) + dX_2^2(t) + (c+b)X_1(t)X_2(t)\right. \\ &+ \left. e^2+f^2+g^2+h^2\right]. \end{aligned} \quad (5)$$

A. Examples

Example 1: We have stochastic differential equation in the form

$$\begin{aligned} d\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \end{aligned} \quad (6)$$

We determine stability of solution using derivation of Lyapunov function

$$\begin{aligned} dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= -2X_1(t)dB_1(t) + 2X_2(t)dB_2(t) \\ &+ 4dt, \\ E\left\{dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}\right\} &= 4dt = LV dt. \end{aligned}$$

Function $LV = 4 > 0$ is positive-definite. Trivial solution of system (6) is unstable.

Example 2: We have stochastic differential equation in the form

$$\begin{aligned} d\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dB_1(t) \\ dB_2(t) \end{pmatrix}. \end{aligned} \quad (7)$$

We determine stability of solution

$$\begin{aligned} dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= 4(X_1(t)^2(t) + X_2(t)^2(t) \\ &+ X_1(t)X_2(t) + 1)dt \\ &+ 2X_1(t)dB_1(t) + 2X_2(t)dB_2(t), \\ E\left\{dV\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}\right\} &= 4(X_1^2(t) + X_2^2(t) \\ &+ X_1(t)X_2(t) + 1)dt = LV dt. \end{aligned}$$

Function is negative-definite for $LV < 0$, i.e.

$$\begin{aligned} 0 &> 4(X_1^2(t) + X_2^2(t) + X_1(t)X_2(t) + 1), \\ |X_1(t) + X_2(t)| &< \sqrt{X_1(t)X_2(t) - 1}, \end{aligned}$$

for $X_1(t)X_2(t) \geq 1$, then trivial solution of system (7) is stable.

IV. Conclusion

In this paper it was defined stability and stochastic stability of the stochastic differential equations. It was computed specific examples by using Lyapunov theorem. Such type of equations can be used also in biomedical engineering, in meteorology, epidemic modeling, predicting economics, etc. For example, stochastic equation of the type

$$dX_t = \left(a + bX_t + cX_t^2\right) dt + \left(kX_t + mX_t^2\right) dB_t$$

can describe a the growth of tumors under immune surveillance and chemotherapy.

Acknowledgement

This research was supported by Grant FEKT-S-11-2-921 of Faculty of Electrical Engineering and Communication, BUT.

References

[1] BAŠTINEC, J.; DZHALLADOVA, I.:*Sufficient conditions for stability of solutions of systems of nonlinear differential equations with right-hand side depending on Markov's process.* In 7. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. 2011. p. 23 - 29. ISBN 978-80-7231-815-5.

[2] DIBLÍK, J., KHUSAINOV, D.Y., BAŠTINEC, J., RYVOLOVÁ, A.: *Exponential stability and estimation of solutions of linear differential systems with constant delay of neutral type.* In 6. konference o matematice a fyzice na vysokých školách technických s mezinárodní účastí. Brno, UNOB Brno. 2009. p. 139 - 146. ISBN 978-80-7231-667-0.

[3] DITLEVSEN, S., BATZEL, J., BACHAR, M.: *Stochastic biomathematical models*, Heidelberg: Springer, 2013, 206 p.

[4] DURRETT, R.: *Probability: theory and examples*, 3. ed. Belmont, CA: Thomson Brooks/Cole, 2005, 497 s. ISBN 05-344-2441-4.

[5] DZHALLADOVA, I.A.: *Optimization of stochastic systems*, Kiev, KNEU Press, 2005. ISBN 966-574-774-6.

[6] DZHALLADOVA, I.; BAŠTINEC, J.; DIBLÍK, J.; KHUSAINOV, D.: *Estimates of exponential stability for solutions of stochastic control systems with delay.* Abstract and Applied Analysis. 2011. 2011(1). p. 1 - 14. ISSN 1085-3375. (IF=1,318).

[7] KHASMINSKII, R.: *Stochastic stability of differential equations.* New York: Springer Berlin Heidelberg, 2011, 358 s. ISBN 978-3-642-23279-4.

[8] KOLÁŘOVÁ, E.: *Stochastické diferenciální rovnice v elektrotechnice.* Thesis. Brno: BUT, 2006, 26 s. ISBN 80-214-3330-2.

Estimation of Statistical Parameters of the Digital Signal Using Artificial Neural Networks

Martin Kotol
Brno University of Technology, Technická 12, 616 00
Brno, Czech republic
kotol.martin@phd.feec.vutbr.cz

***Abstract*— This paper deals with the artificial neural networks (NN), which are used to estimate the statistical parameters of the digital signal. We will compare the results of training NN with Radial Basic Function and Feedforward neural network. In the case of Feedforward networks, several activation functions will be used. Then the results of training of all neural networks will be compared.**

Keywords*— *artificial neural network, Feedforward, Radial Basic Function, statistical parameters, digital signal

I. INTRODUCTION

This paper will compare Feedforward neural network and Radial basis function neural network as estimators of the statistical parameters of the digital signal. At first, principles of neural networks will be described. Next, specific types as Feedforward (FF) neural networks and Radial Basic Function (RBF) NN will be described in detail together with typical structure and function of these neural networks. Description of training signals used for artificial neural network follows. We will further examine how to adjust these these signals to a suitable shape for training the network and creating of training patterns. Finally, we will compare the results of training the neural networks from viewpoint of accuracy of estimation of statistical parameters and the time needed for their training.

II. GENERAL NEURAL NETWORK

This chapter is focused on the general description of the neuron and its following implementation in the neural network. The next section will be focused to a general description of neural networks. Artificial neuron is a simple computational unit containing: input function, activation function and output function. The Input function is designed to take input data and to send the data to the Activation function, where the data are transformed accordingly by the activation function. Activation function contain a set threshold which decides whether the neuron is active or not. When the output of the activation function is lower than the set threshold, the neuron is set as inactive and vice versa. The Output function then distributes the data from the neuron.

Artificial neurons are connected in layers which create the whole neural network (NN). Individual connections between the neurons have a different weight. The weight is set as a random value and during the training is this value being adjusted in order to obtain a result with the smallest error. General neural network is depicted in Fig.1. It consists of three basic layers: input, hidden and output layers.

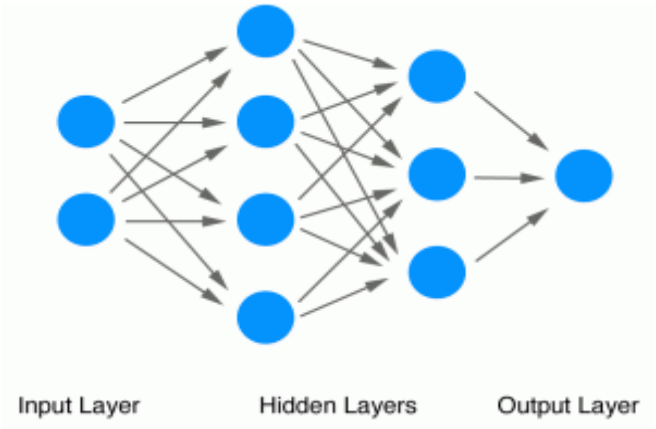


Fig. 1. General neural network

As shown in Fig.1, the neuron in the *n*-th layer of the network is connected to all neurons of the next layer. Neurons on the *n*-th layer are interconnected.

The first layer is the input layer of NN that is designed to take an input data and distribute it to the next layer of the network. The last layer is the output layer of the network, which distribute the signal from the network. All layers which are between the input and output layers are hidden layers and these layers perform further processing (transformation) of the data according to predetermined functions. These transformation functions are given by NN type and may vary according to the type of the activation function.

III. USED NEURAL NETWORK

A. Feedforward neural network

This section will deal with feedforward neural network with backpropagation of errors. This network was chosen because it is one of the most commonly used in applications of NN. FF network topology is fixed at the beginning, NNs with one, two or three hidden layers are used. Further increasing of the number of hidden layers does not bring a large increase in the accuracy of the network compared with complexity of design of such a network. The general scheme of FF network is in Fig.2. x_1 - x_n are input neurons. These input neurons are connected with the hidden layer, which may have *m* hidden neurons and an output layer may contain up to *n* neurons.

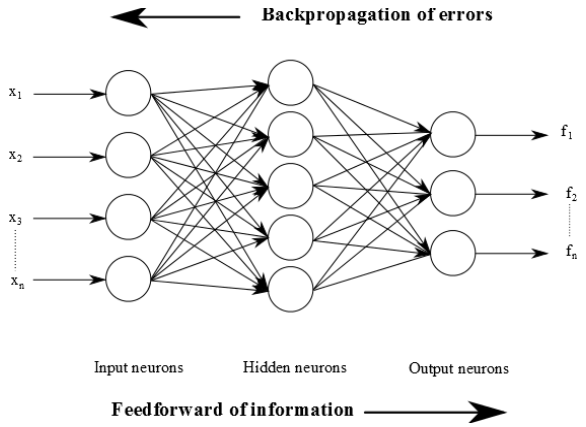


Fig. 2. Feedforward neural network

Feedforward neural network uses the forward propagation of information. This means that the information is distributed only in one direction without feedback. At the beginning of training the weights between neurons are set randomly. After the information is processed through the network, the error at the output of the network is. This output error is sent to the beginning of network and according to this error, the weight between the neurons are adjusted and another cycle of training follows The training is completed when the output error reaches defined value or the maximum number of allowed training of cycle is exceeded.

B. Radial basic function

Radial basis networks are very universal approximation tools, this also the reason why they were used in this study. The network topology of RBF is different from the previous one. RBF network has not fixed topology, input, hidden and output layers are, however, used. The topology of a general RBF network can be seen in Fig.3. x_1 - x_6 are the inputs of input neurons.

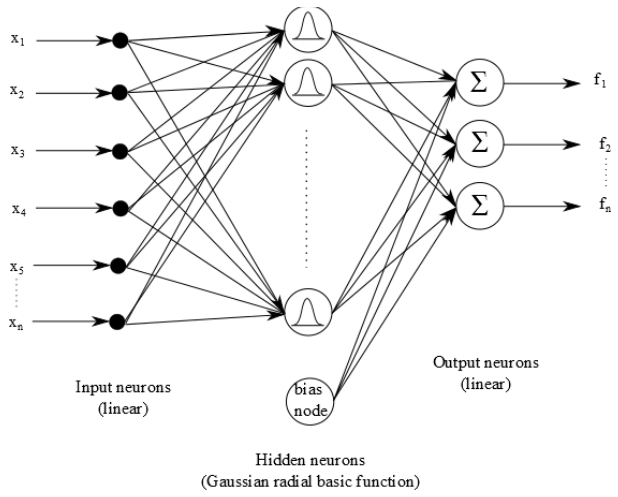


Fig. 3. Radial basic function neural network

The network contains linear input neurons, the number of input neurons is *k*. The input neurons are connected with the hidden neurons which contain radial basic activation function.

These hidden neurons are connected with the output neurons which sum the contributions of all hidden neurons.

The training of the network is performed as follows: Before the training, the number of input and output neurons is set. Number of hidden neurons is set to one. The training signal then passes through the entire network. Then the output error of network is calculated . If the error is greater than the defined limit, one hidden neuron is added and another cycle of training is performed. This is repeated until the output error of the network drops below defined threshold or the number of hidden neurons reaches maximum defined value.

IV. THE SIGNALS USED TO COMPARE NEURAL NETWORK

This chapter describes the signal used for comparison of mentioned NNS. In order to estimate the statistical parameters, a randomly generated digital signal of the length of 200 samples was created. AWGN was added to this signal.

For training purposes, we created an input matrix. It contained the same 100 digital signals and each with a different noise. From these signals, the statistical parameters were calculated and from these statistical parameters, the output matrix for training was created.

V. RESULTS

In this chapter shows and compares the results of training NN and the accuracy of estimation of statistical parameters of signal using artificial neural networks. For training, two types of NNs were used, FF and RBF network. In the case of FF networks following activation functions were used: Bayesian regularization, Levenberg-Marquardt and Adaptive learning rate.

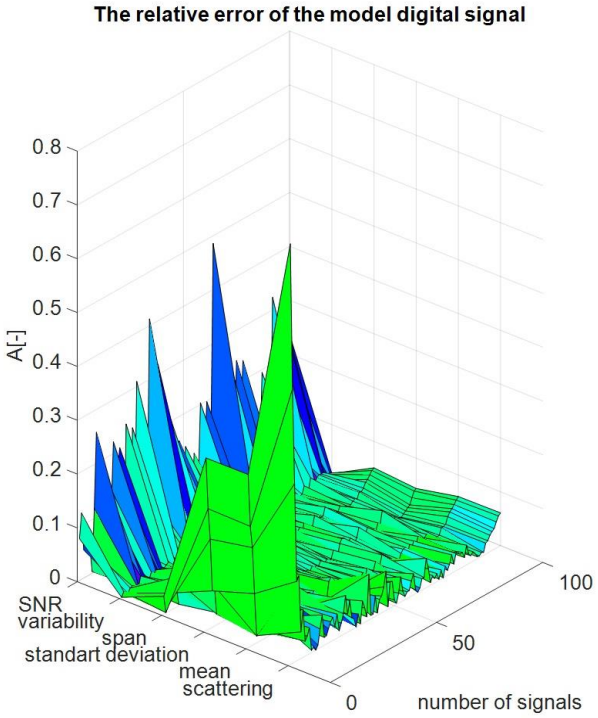


Fig. 4. Relative error for activation function Bayesian regularization

The Fig.4 shows the relative error of estimating the statistical parameters with Bayesian regularization activation functions. The figure shows that the largest error of estimation is for SNR. This is due to that SNR has the greatest dynamics of the estimated parameters.

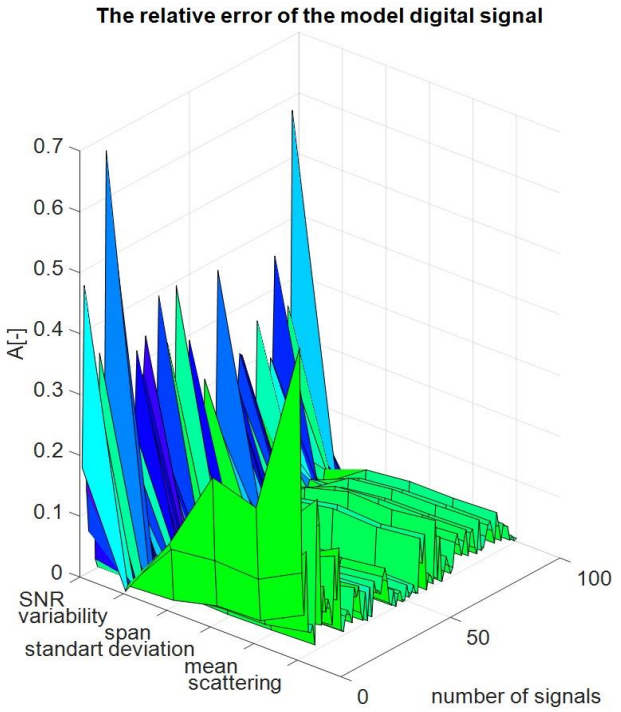


Fig. 5. Relative error for activation function Levenberg-Marquardt

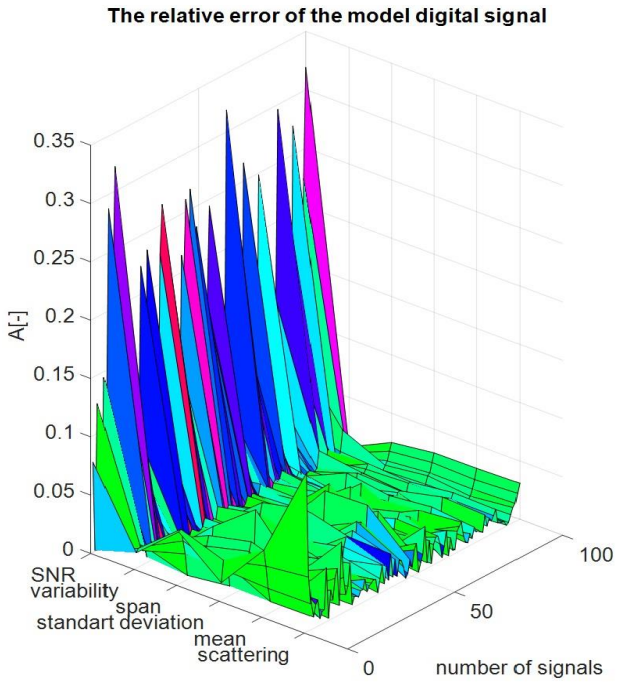


Fig. 6. Relative error for activation funkction Adaptive learning rate

Figure 5 shows the relative error of estimation of the statistical parameters with activation functions Levenberg-Marquardt. While using this activation function, the SNR has again the highest error..

Fig. 6 shows the relative error of estimation of statistical parameters for the neural network type FF with activation function adaptive learning rate. The figure shows that FF network with activation function Adaptive learning rate estimated statistical parameters with the smallest relative error from all activation functions used FF neural networks. For all FF networks, the same network topology was used for better comparison results of training.

Activation function	Required accuracy	Time training [mm:ss]	Accuracy achieved	Number of iterations
Bayesian regularization	**1,00E-09**	**33:15**	**1,00E-09**	**306**
Levenberg-Marquardt	**1,00E-09**	**10:25**	**1,00E-09**	**474**
Adaptive learning rate	**1,00E-09**	**6:23**	**4,11E-05**	**1,00E+05**
RBF	**1,00E-09**	**15**	**1,00E-09**	**99**

Fig. 7. Comparison of the results of NN training networks

Fig.7 shows that the least time for training needed the RBF network. FF network with activation function Adaptive learning rate has not reached the required accuracy, because it reached the maximum number of iterations during training

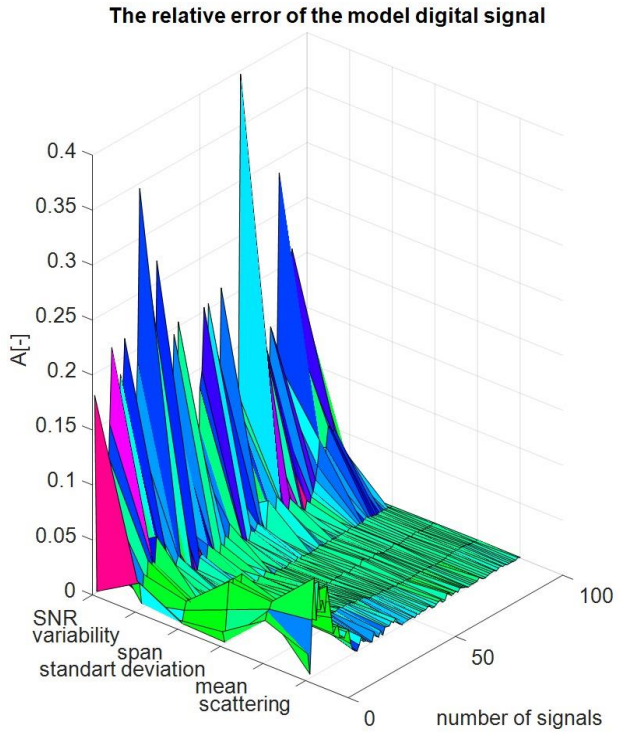


Fig. 8. Relative error for net Radial basic function

The Figure 8 indicates the relative error of estimation of the statistical parameters with RBF network. As can be seen, the error of the estimation is similar to the case of FF network with Adaptive learning rate activation function.

VI. Conclusion

The paper generally described the principle of neural network next, the NNs for estimation of statistical parameters of generated digital signal were presented. We described how to create input and output matrices for training of the neural networks. Finally, we the results of estimation of the statistical parameters of the digital signal using trained neural networks were presented. The results show that the best accuracy can be obtained using FF network with Adaptive learning rate activation function. The fastest NN in terms of training was RBF network, the FF network with Bayesian regularity had the highest time demands. The largest estimate error was observed for the parameter SNR because SNR has great dynamic range which is problem for estimation.

[1] D. Kriesel, A Brief Instroduction o Neural Networks, University of Born in Germany,2005.

[2] Š. Jiří and N. Roman,Teoretické otázky neuronových sítí, pp.389.

[3] I.S. Jacobs and C.P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4] RAIDA, Z, Modeling EM structures in the neural network toolbox of MATLAB, IEEE Antennas and Propagation Magazine, 2003, vol. 44, no. 6, p. 46-67.

[5] Jingwen Tian; Meijuan Gao; Fan Zhang, "Network Intrusion Detection Method Based on Radial Basic Function Neural Network," E-Business and Information System Security, 2009. EBISS '09. International Conference on , vol., no., pp.1,4, 23-24 May 2009 doi:10.1109/EBISS.2009.5138016

Computational Electromagnetics in Aerospace

HIRF and Lightning

David Krutílek
Department of Radio Electronic
Brno University of technology
Technicka 3082/12, 616 00 Brno, Czech Republic
xkruti01@stud.feec.vutbr.cz

***Abstract* — This paper aims to summarize the existing pieces of knowledge about computational electromagnetics (CEM) methods in the field of aircraft industry. Research that has been carried out so far can be divided into two groups according to the external effects affecting the aircraft during its common operation. These are effects of high intensity fields (HIRF) and effects caused by lightning stroke.**

Keywords—Aircraft; CEM; Simulation; HIRF; (In)direct effect of lightning, EMC.

I. Introduction

Computational electromagnetism (CEM) is a branch established at the end of 1960's when the first applications for task solving in the field of electromagnetism with the aid of finite element method occurred. Retrospectively, computer technology was at rudimentary level and thus the computational costs were astronomical. As the computer technology was improving and spreading, the price of hardware decreased and new, more or less effective methods were presented. These methods led to the fusion of computations and measurements and brought certain financial benefits for reduction of initial production costs. State of the art and CEM utilization in the aviation industry can be divided into two main groups based on the external effects affecting the aircraft during its common operation:

- Effects of fields with high intensity (high intensity radiated fields, HIRF)

The main sources of HIRF can be radio and television transmission towers, radio-technical devices used at the airport, radars, etc. In the field of aircraft industry, there used to be explorative and developmental projects ARTEMIS [1] (Analytical Research of Threats in ElectroMagnetically Integrated Systems; 2007 - 2010) and HIRF-SE [2] (High Intensity Radiated Field – Synthetic Environment; 2008 – 2013). These projects used to be concerned with problematics of HIRF and their aim was to replace costly tests with computations within certification process on EMC.

- Influences of effects caused by lightning stroke

The aircraft is not in touch with the stroke of lightning discharge so often as it is in case of flight through HIRF. There has not been any explorative and developmental project, which would be concerned with the effects of lightning stroke into various means of transport or buildings, up to now. This topic is discussed within conference SIPDA [3] (O Simpósio Internacional de Proteção contra Descargas Atmosféricas - The International Symposium on Lightning Protection), or for instance also within ICLP [4] (International Conference on Lightning Protection).

II. Use of CEM

A. Computations of effects caused by HIRF

CEM has been used for investigation of electromagnetic field interaction with the aircraft since 70's in the 20th century [5-7]. At that time, there was no real possibility to achieve the real results with the aid of available computational methods. CEM was in those days primarily used for undemanding, elementary tasks.

The abilities of CEM, one of which is to achieve the coincident results by means of different computational methods based on time (time domain, TD) and frequency (frequency domain, FD) approaches, are described in article [8], where modified numerical metal model of the aircraft EV-55 [9] is calculated. Generally valid division of applicability of these approaches for small aircraft simulations depending on the area of solved problem and its frequency can be found in the table below. Every technique from time or frequency domain shows its strengths and weaknesses that form boundaries of computational technique applicability.

The strong point of time domain technique is a possibility to analyze broadband problems, transitional and time non-linear processes or behavior of reflected field in close surroundings of subject. In frequency domain, it is advantageous to analyze narrowband problems, models with frequency dependent material characteristics and tasks with extremely long time responses, such as resonators or lightning current analysis.

Frequency band	External surface fields	Internal volume fields		Cable currents
10 kHz – 10 MHz	FD	FD		FD
10 MHz – 100 MHz	TD	FD	TD	TD
100 MHz – 3 GHz	TD	TD		TD

Fig. 1 Application domain of the FD and TD [8].

The following logical step is to verify the numerical methods on real metal aircraft [10]. Measuring of small metal aircraft VUT-100 Cobra [11] took place in semi-anechoic chamber. In the figure 2 is illustrated the aircraft VUT 100 during its measurements and the simulation of highlighted points discovers field intensity (FI), surface current (SC) and antenna position (AP).

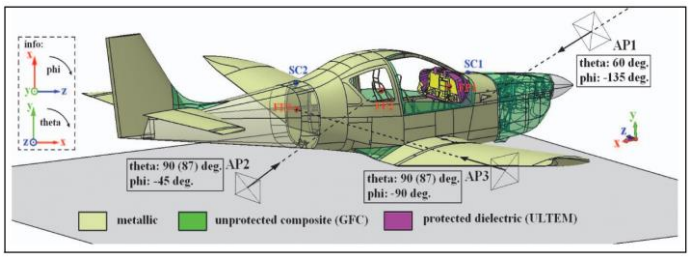


Fig. 2 VUT-100 Cobra position under test [10].

Computations were implemented in frequency (CONCEPT II [12]) and time (CST MWS [13]) solver. Selected comparison of computed and experimentally obtained data can be seen in Figures 3 and 4:

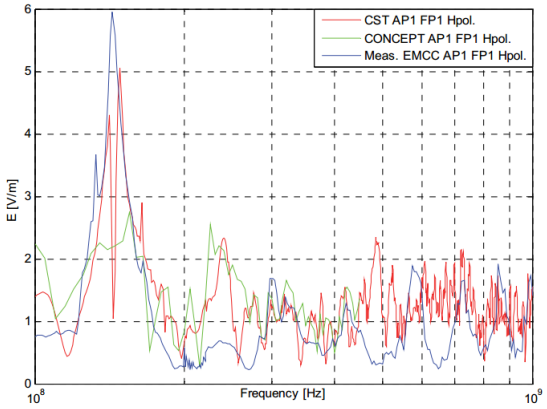


Fig. 3 Field intensity in FP1 [10].

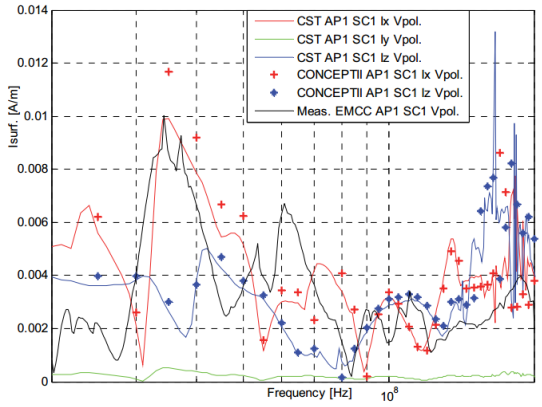


Fig. 4 Surface current in SC1 [10].

The diagrams show relatively good agreement of simulated and measured data in spite of a very simplified aircraft model, which consists of metal parts only.

The next research on the same aircraft model has been concerned with transients on cable harness, which were calculated by means of method of moments [14]. Illumination by antennas is based on situation that can be seen in the Fig. 2. The explored points on cable harness are illustrated in the Fig. 5.

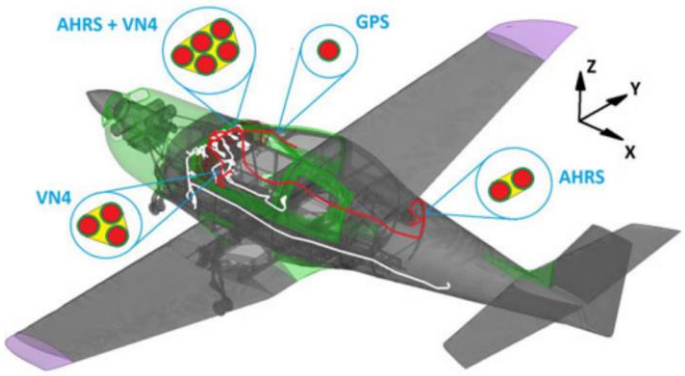


Fig. 5 Cable harness in VUT-100 Cobra [14].

Selected results are compared in the Fig. 6. The results proved worse coincidence than in case of previous surface current exploration and electric field intensity exploration. Computed results represent approximate but sufficient degree of transients on cable harness.

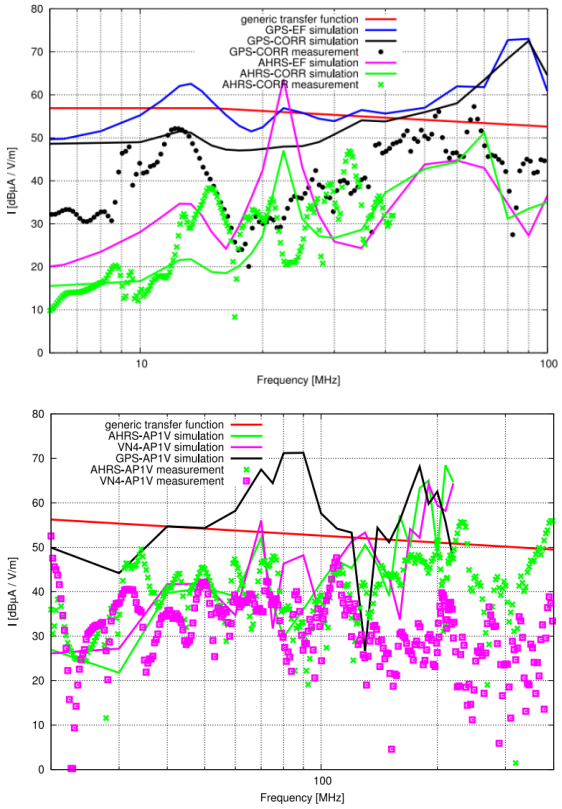


Fig. 6 Transients on cable harness of VUT-100 Cobra [14].

Additional pieces of information related to problematics of transients on cable harness provide literature [15] and [16].

B. Indirect effects of lightning stroke

Simulation of direct effects of lightning affecting the aircraft is very complicated. There has not been any article published so far, that would describe the direct effect on the complex model of the aircraft. In most cases, only elementary components are solved, such as joints or composite structures. Relatively complicated topic is discussed in the article, that describes minimizing of metal demountable joint sparkling on composite structure during lightning current. Such situation can be seen in the Fig. 7, which illustrates sectional view of the joint and possible varieties of lightning stroke with sparkling.

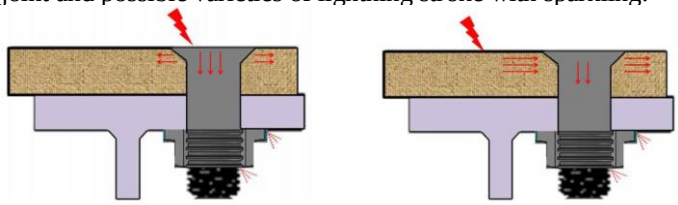


Fig. 7 Cut of joint CPRF/metal [17].

If the joint would be placed in fuel tanks, sparkling could have disastrous impact. The lightning stroke and its subsequent sparkling could light up the gas inside the tank, which would cause an explosion. In the Figure below, there is a result of experiment with sparkling on such joint.

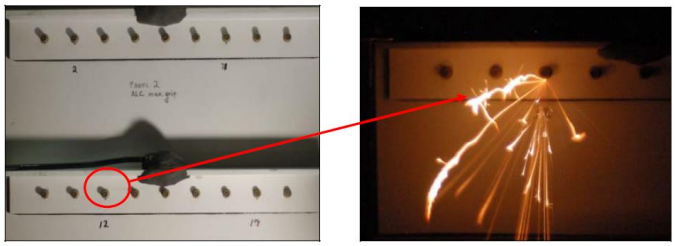


Fig. 8 Real sparkle in joint – test case [17].

Sparkling can be minimized by appropriate modification and using of suitable materials with lower contact resistance between metal joint and composite material CFRP. The proposal itself was simulated and verified by computations. The sectional view of the joint and more strained areas can be seen in the Fig. 9.

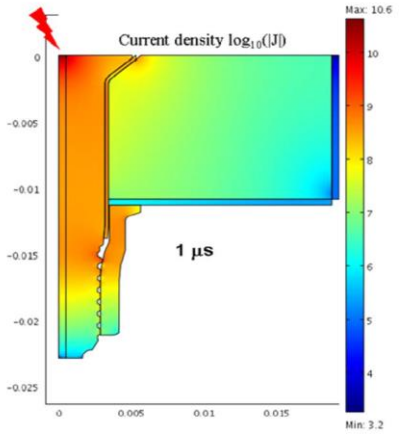


Fig. 9 Numerical model of joint [16].

Another literature is focused on resistance of composite materials during lightning stroke. It is an experimental measurement of mechanical damages or penetrative tensions on composite materials. Unfortunately, the research is concerned only with carbon- or glass-fibre composite structures, but it doesn't discuss possibilities of protective grating, painting or admixtures, which are normally used for higher conductivity and better current distribution.

C. Direct effects of lightning stroke

Simulation of indirect effects of lightning is focused on investigation of transients on cable harness that are caused by lightning current flowing on metal surface of the aircraft. Especially huge all-metal transport aircrafts are examined [18].

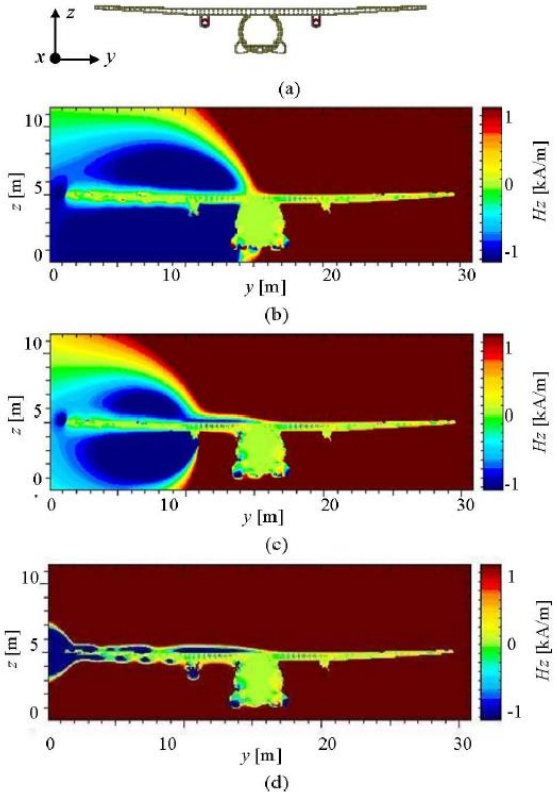


Fig. 10 Magnetic field (H_z) by lightning stroke in to big metal AC [18].

The results of induced tensions on cable harness are not verified with an experiment. The verification as such is in this case difficult to realize not only for financial reasons (the lightning stroke damages the aircraft in most cases) but also because of the realization itself (lightning stroke is simulated during the flight). The results are primarily intended for designers to imagine what is happening inside the aircraft, what transients can be induced in cable harness during the lightning stroke, how should be the construction of the aircraft modified and when it is necessary to change the cable constellation in case the results are not acceptable.

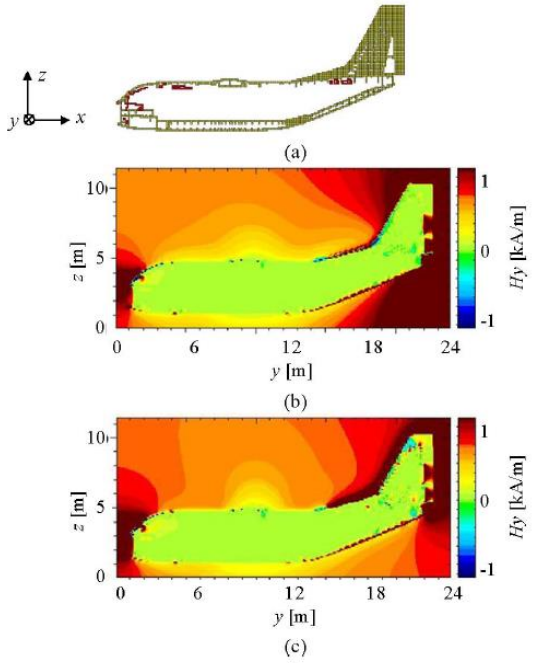


Fig. 11 Magnetic field (H_z) by lightning stroke in to big metal AC [18].

III. Conclusion

Numerical analyses of electromagnetic field are inseparable part of the aircraft development and its protection against lightning stroke and HIRF effects. This is also proven by public interest of multinational companies taking part in European projects, such as HIRF-SE etc. The analyses help us to mark weak points on the aircraft form the lightning protection point of view and such areas can be removed before production or testing of the aircraft. It is supposed, that resistance of the aircraft against HIRF and lightning stroke will be proved by means of simulations and all the expensive certification tests and measurements won't have to be done anymore.

Acknowledgment

Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

References

[1] ARTEMIS – Analytical Research of Risks in Electromagnetically Integrated Systems [online]. 2007 [cit. 2015-4-22]. Description — International projects. Available at WWW: <http://www.evektor.cz/en/international-projects.aspx>.

[2] HIRF SE High Intensity Radiated Field Synthetic Environment [online]. 2008 [cit.2013-8-15]. Description — HIRF SE. Available at WWW: <http://www.hirf-se.eu> or <http://ec.europa.eu/research/transport/projects/items/hirf_se_en.htm>.

[3] SIPDA [online]. 2015 [cit.2015-4-22]. Description — The International Symposium on Lightning Protection. Available at WWW: <http://www.usp.br/sipda/?q=en/home>.

[4] ICLP [online]. 2015 [cit.2015-4-22]. Description — The International Conference on Lightning Protection. Available at WWW: <http://www.iclp-centre.org>.

[5] Baum, C. E., Reminiscences of High-Power Electromagnetics, Electromagnetic Compatibility, IEEE Transactions on , vol.49, no.2, pp.211,218, May 2007.

[6] Baum, C. E., How to Think About EMP Interaction, Proceedings of the 1974 Spring. FULMEN Meeting, Kirtland AFB, April 1974.

[7] Baum, C.E., Liu, T.K., Tesche, F.M., On the Analysis of General Multiconductor Transmission Line Networks Interactions Notes, Interaction Notes 350, Air Force Weapons Laboratory, Albuquerque, NM, January 1978.

[8] Alvarez, J., Angulo, L., Bandinelli, M., Bruns, H., Francavilla, M., Garcia, S., Guidi, R., Gutierrez, G., Jones, C., Kunze, M., Martinaud, J., Munteanu, I., Panitz, M., Parmantier, J., Pirinoli, P., Reznicek, Z., Salin, G., Schroder, A., Tobola, P., Vipiana, F., "HIRF interaction with metallic aircrafts. A comparison between TD and FD methods," Electromagnetic Compatibility (EMC EUROPE), 2012 International Symposium on , vol., no., pp.1,6, 17-21 Sept. 2012.

[9] Airplane EV-55 [Online]. 2015[cit. 2015-4-22] Aviable on WWW: <http://www.evektor.cz/en/ev-55-outback>.

[10] Řezníček, Z., Tobola, P., Rasek, G. A., Loos, S.E., TD and FD simulations of internal EM environment in small aircraft and experimental test comparison, Antennas and Propagation (EUCAP), 2012 6th European Conference on , vol., no., pp.2261,2265, 26-30 Mar, 26-30 March 2012.

[11] Airplane VUT100 [Online]. 2015[cit. 2015-4-22] Aviable on WWW: <http://www.evektor.cz/en/vut100-cobra>.

[12] Technische Universitat Hamburg-Harburg. CONCEPT-II. (2015). [Online]. Available: http://www.tet.tu-harburg.de/concept.

[13] CST Microwave Studio Users' Guide. Darmstadt: 2015.

[14] Guido A. Rasek, Arne Schroder, Pavel Tobola, Zdeněk Rezníček, Steffen E. Loos, Thorsten Tischler, Heinz-Dietrich Bruns, HIRF Transfer Function Observations: Notes on Results Versus Requirements and Certification Approach, Electromagnetic Compatibility, IEEE Transactions on, Vol. 57, No. 2, April 2015.

[15] Parmantier, J-P., Ridel, M., Bertuol, S., Junqua, I., Giraudon, C., Girard, C., Terrade, F., Moreau, J-P., Modelling of HIRF coupling on Complex Cable Architectures, International Conference on Electromagnetics in Advanced Applications (ICEAA), 12-16 Sept., 2011, p. 219 – 222, ISBN: 978-1-61284-976-8.

[16] S. Bertuol, M. Ridel, J.-P. Parmantier, L. Pisu, M. Bozzetti, A. Francavilla, M. Righero, F. Vipiana, S. Arianos, and G. Vecchi, "Field to transmission line coupling: A test case within HIRF SE," in Proc. Int. Symp. Electromagn. Compat., 2012, pp. 1–6.

[17] Mulazimoglu H., Haylock L., "Recent Developments in Techniques to Minimize Lightning Current Arcing between Fasteners and Composite Structure " aviable on: https://www.alcoa.com/fastening_systems_and_rings/aerospace/en/pdf/hasim_icosle_2011.pdf.

[18] Min ZHANG, Zhiyong HUANG, "Transient Current Burst Analysis induced in Cable Harness due to Direct Lightning Strike on Aircraft", 2010 Asia-Pacific International Symposium on Electromagnetic Compatibility, April 12-16 2010, pp. 1197-1200. .

Wireless Network Coding in LTE Relays

Zenon Kuder
Department of Radio Electronics
Brno University of Technology
Email: xkuder00@stud.feec.vutbr.cz

***Abstract*—Wireless Network Coding is a technique that leverages the inherent broadcast nature of the wireless channel to enhance the speed and/or reliability of relayed communication by using the shared information among the receiving and transmitting node. Since relay nodes were introduced in LTE-Advanced, it is natural to consider the possibility of enhancing the relaying performance by employing the wireless network coding.**

I. Introduction

Increasing user demand for throughput in wireless networks and decreasing available spectrum has led to the fast-paced emergence of new wireless technologies that allow system throughputs several orders of magnitude higher than in the conventional and ubiquitous systems of the seconds generation, such as GSM.

One of the most successful ways to increase the spectral efficiency is Multiple Input–Multiple Output (MIMO) spatial multiplexing, where several streams of independent data can be transmitted simultaneously using the same bandwidth.

However, the gains of spatial multiplexing depend on the minimum of the numbers of antennas on both the receiver and transmitter and low correlation between the paths between the transmitting and receiving antennas. Therefore in practice, for the small form factor devices such as mobile phones, it is not viable to have more than 2–4 antennas.

On the other hand, the base stations can accomodate many antennas and complex, high-performance hardware. This assymetry has been inherent to radio communication standards since the very beginning, from the choice of lower frequency for uplink in GSM Frequency-Divison Duplexing (FDD) to the choice of Single-Carrier FDMA (SC-FDMA) for LTE uplink to achieve lower Peak-to-Average Power Ratio (PAPR).

One of the techniques to overcome this assymetry is multi-user MIMO (MU-MIMO). In uplink, several mobile stations with low antenna number can transmit simultaneously and their signals are separated thanks to the high number of antennas at the base station and the resulting high number of degrees of freedom. In downlink, on the other hand, the high number of antennas at the base station can be used to beamform different signals to users in different directions or with highly uncorrelated complex path coefficient.

Further increase in the antenna count has led to the recent establishment of the concept *Massive MIMO* for systems with large amounts of antennas at the base station, where pilot contamination has the dominant impact on the radio system, while channel estimation, interference and noise are negligible [1].

To overcome the limitation of antenna number at the mobile stations and also their low spatial separation, several mobile stations can cooperate and relay their received signals to each other. This technique is known as Virtual MIMO (V-MIMO) [2]. As shown in Figure 1, in this cooperation the nodes serve as relays and therefore there can be gain in the network performance hidden in the relay technology.

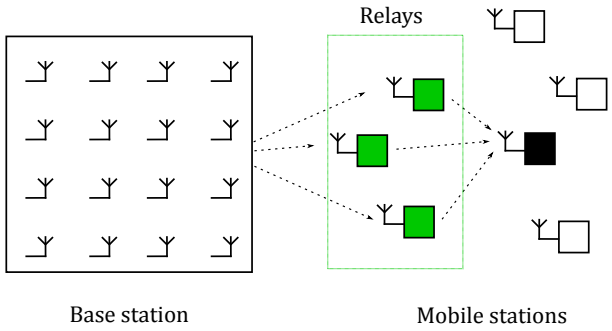


Fig. 1. Virtual MIMO.

While relays are an old concept, there are novel ideas that could improve their efficiency. One of them is Wireless Network Coding (WNC) [3], where the knowledge of the transmitted signal at each node is used to optimize the number of steps needed for a two-way symmetrical data exchange.

This article presents an overview of the WNC techniques.

II. System Model

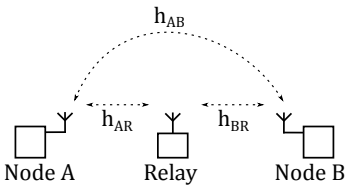


Fig. 2. Relay model.

For the purpose of this article, let's assume that two nodes, A and B communicate via a relay node R, as illustrated in Figure 2. Channel reciprocity is assumed, e.g. Time-Divison Duplexing (TDD) can be used, and therefore the communication between the nodes goes via wireless channels $h_{AR} = h_{RA}$ and $h_{BR} = h_{RB}$. Additionally, as the wireless channel is

inherently broadcasting, there is also a communication directly between A and B via channel $h_{AB} = h_{BA}$, whose parameters are insufficient for successful and/or satisfactory communication. It can, however, be used to enhance the parameters of the relayed link. In the simplest scenario, Additive White Gaussian Noise (AWGN) with variance σ^2 is assumed to model both the noise and interference.

Node A and B are about to transmit their data packets of length N $\mathbf{m_{AB}}$ and $\mathbf{m_{BA}}$ to each other, where $\mathbf{m_{xy}}$ is a string of bits from node x to node y, $x, y \in \{A, B, R\}$.

When two nodes communicate via a relay, in general four steps are used:

1) $A \xrightarrow{\mathbf{m_{AB}}} R$: Node A sends its message $\mathbf{m_{AB}}$ to R,
2) $B \xrightarrow{\mathbf{m_{BA}}} R$,
3) $R \xrightarrow{\mathbf{m_{BA}}} A$,
4) $R \xrightarrow{\mathbf{m_{AB}}} B$.

Node A and B have now successfully exchanged their messages in four timeslots: $t_{basic} = 4T_{ts}$.

III. Throughput Improvement Methods for Relays

A. Decode-and-Forward (DF)

To enhance the total throughput, it is possible to decrease the number of required timeslots to 3 by utilizing the fact that both node A and B know their transmitted messages and the fact that the relay signal is broadcast to both A and B in both steps 3 and 4.

1) $A \xrightarrow{\mathbf{m_{AB}}} R$: Node A sends its message $\mathbf{m_{AB}}$ to R,
2) $B \xrightarrow{\mathbf{m_{BA}}} R$,
3) $R \xrightarrow{\mathbf{m_{RA}}=\mathbf{m_{RB}}} A, B$,

while $\mathbf{m_{RA}} = \mathbf{m_{RB}} = \mathbf{m_{AB}} \oplus \mathbf{m_{BA}}$, where $\oplus$ is the exclusive OR (XOR) operation.

As the relay decodes the data and encodes the combined message to be transmitted back, it is easy to see that assuming correct decoding and the knowledge of transmitted messages $\mathbf{m_{AB}}$ and $\mathbf{m_{BA}}$ for node A and B the decoding of the desired message is again a simple XOR operation:

For node A:

$$\mathbf{m_{AB}} \oplus \mathbf{m_{RA}} = \cancel{\mathbf{m_{AB}}} \oplus \cancel{\mathbf{m_{AB}}} \oplus \mathbf{m_{BA}} = \mathbf{m_{BA}} \quad (1)$$

For node B:

$$\mathbf{m_{BA}} \oplus \mathbf{m_{RB}} = \cancel{\mathbf{m_{BA}}} \oplus \mathbf{m_{AB}} \oplus \cancel{\mathbf{m_{BA}}} = \mathbf{m_{AB}} \quad (2)$$

By using DF, the data exchange between A and B can be done in 3 timeslots, i.e. $t_{DF} = 3T_{ts}$. In contrast with the basic relaying, where $2N$ bits are transmitted in $4T_{ts}$, the gain is

$$G_{DF} = \frac{\frac{2N}{3T_{ts}} - \frac{2N}{4T_{ts}}}{\frac{2N}{4T_{ts}}} = \frac{1}{3}. \quad (3)$$

B. Amplify-and-Forward (AF)

Further improvement can be accomplished by combining the first two steps, i.e. having both A and B transmit their message at the same time. However, without additional complexity this would be a source of intereference at the relay. Keeping in mind the fact that both nodes know the signals they transmitted, it can be seen that the relay does not have to decode the signals, but can instead simply amplify them. The nodes A and B then bear the complexity of separating the desired signal from the "echo" of their transmitted signal.

As the relay now amplifies the sum of the received signals in the signal domain, this technique is called Amplify-and-Forward and can be described by the following equations.

Signals $\mathbf{x_{AB}}$, $\mathbf{x_{BA}}$ are complex baseband representations of the packets $\mathbf{m_{AB}}$, $\mathbf{m_{BA}}$.

$$\mathbf{y_R} = \sqrt{\varepsilon} h_{AR} \mathbf{x_{AB}} + \sqrt{\varepsilon} h_{BR} \mathbf{x_{BA}} + \mathbf{n_R}, \quad (4)$$

where $\mathbf{n_R}$ is the complex AWGN noise vector at the relay R with the variance σ^2.

The summed signal $\mathbf{y_R}$ received at R is amplified by a factor of β to the level of transmit power to create the transmit signal $\mathbf{x_R}$ to be forwarded from the relay.

$$\mathbf{x_R} = \beta \mathbf{y_R} = \beta(\sqrt{\varepsilon} h_{AR} \mathbf{x_{AB}} + \sqrt{\varepsilon} h_{BR} \mathbf{x_{BA}} + \mathbf{n_R}). \quad (5)$$

The coefficient β is defined as: [3]

$$\beta = \sqrt{\frac{\varepsilon}{\varepsilon \cdot h_{AR}^2 + \varepsilon \cdot h_{RB}^2 + \sigma^2}}, \quad (6)$$

where ε is the average transmitted energy per symbol.

The signals received at A and B can be described as

$$\mathbf{y_A} = h_{AR} \mathbf{x_R} + \mathbf{n_A}, \quad (7)$$

$$\mathbf{y_B} = h_{BR} \mathbf{x_R} + \mathbf{n_B}. \quad (8)$$

Assuming both A and B have the knowledge of β and their path loss to R:

$$\begin{aligned} \mathbf{y_A} &= h_{AR} \mathbf{x_R} + \mathbf{n_A} \\ &= h_{AR} \beta \mathbf{y_R} + \mathbf{n_A} \\ &= h_{AR} \beta (\sqrt{\varepsilon} h_{AR} \mathbf{x_{AB}} + \sqrt{\varepsilon} h_{BR} \mathbf{x_{BA}} + \mathbf{n_R}) + \mathbf{n_A} \\ &= \underbrace{\beta h_{AR}^2}_{\text{known*}} \underbrace{\sqrt{\varepsilon} \mathbf{x_{AB}}}_{\text{known}} + \underbrace{\beta h_{AR}}_{\text{known*}} \underbrace{\sqrt{\varepsilon} h_{BR} \mathbf{x_{BA}}}_{\text{unknown/desired}} + \underbrace{\beta h_{AR} \mathbf{n_R} + \mathbf{n_A}}_{\text{noise}}. \end{aligned} \quad (9)$$

By subtracting the known influence of the transmitted signal $\beta h_{AR}^2 \sqrt{\varepsilon} x_{AB}$, the desired signal can be processed by a traditional receiver in the form $\mathbf{r_A}$:

$$\mathbf{r_A} = \beta h_{AR} h_{BR} \sqrt{\varepsilon} \mathbf{x_{BA}} + \beta h_{AR} \mathbf{n_R} + \mathbf{n_A}. \quad (10)$$

In a similar manner, node B can receive its signal r_B:

$$\mathbf{r_B} = \beta h_{AR} h_{BR} \sqrt{\varepsilon} \mathbf{x_{AB}} + \beta h_{BR} \mathbf{n_R} + \mathbf{n_B}. \quad (11)$$

Note that the values marked as *known**, i.e. β, h_{AR}, h_{BR} have to be communicated between the nodes in the control channel or in another similar way.

C. Using the Direct Link Between A and B

Additionally the nodes can leverage the weak signals of the direct links $A \rightarrow B$ and $B \rightarrow A$ in case of using a more advanced signal detection method to increase the extrinsic bit information available for the signal detection.

IV. Potential Usage

As mentioned, relaying is one of the upcoming improvements to Long Term Evolution (LTE) and also it can be used among the mobile terminals in the arrangement of V-MIMO. By improving the relaying performace, less wireless resources (e.g. resource blocks in case of LTE) can be used and the general network performance can be increased.

V. Conclusion

This article summarized some of the recent advances in the relaying and MIMO technologies; particularly the usage of relays in the V-MIMO and the potential of network coding and two-step amplification to increase the throughput of relays, and therefore the performance of the V-MIMO system.

References

[1] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[2] Y. Zhou, F. Adachi, K.-K. Wong, X.-G. Xia, D. Toumpakaris, H. Steendam, W.-P. Zhu, and L.-L. Yang, "Guest Editorial: Virtual MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 1977–1980, Oct. 2013.

[3] P. Popovski and H. Yomo, "Bi-directional Amplification of Throughput in a Wireless Multi-Hop Network," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 2, May 2006, pp. 588–593.

Impact of Predefined Quality Profiles in H.265 to Video Quality and Speed of Encoding

Jan Kufa, Tomas Kratochvil
Department of Radio Electronics, SIX Research Center
Brno University of Technology
Brno, Czech Republic
xkufaj00@stud.feec.vutbr.cz, kratot@feec.vutbr.cz

***Abstract*— This paper presents a coding performance of the coding standard H.265/HEVC with different predefined quality profiles. Quality of encoded video and time needed to encode were measured. For our comparison, the latest available version of codec for encoding was used. Implementation x265 was used. Comparison is provided by set of objective quality metrics. Full HD resolution of videos for our test was used.**

***Keywords*— H.265, HEVC, x265, video coding, objectiv quality metrics, Full HDTV, speed of encoding video, PSNR, SSIM, VQM**

I. Introduction

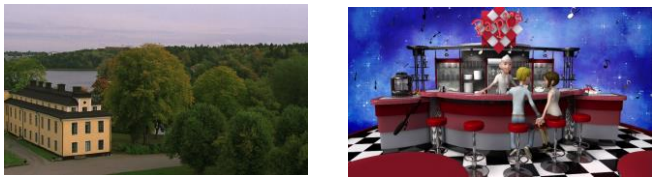
Nowadays, the trend is to deliver still higher and higher quality of video services to customers. For higher quality it is necessary to use better encoders because the bandwidth of the transmission line is limited. Video applications on the internet are more expanded day by day and they are consuming much more of available bandwidth than in the past. Used codec H.265 is known as HEVC (High Efficiency Video Coding). Encoder H.265 is used for broadcasting of the Ultra HDTV for IPTV Netflix.

Newest codecs are more complex and need more time to encode the video. Real time encoding is necessary for some systems like DVB-T and are not necessary for another services like YouTube for example.

Why to examine predefined settings? If the compression of video is five hours instead of one hour, it would be good to be sure that it is profitable and image quality is better and the time spent on the compression is not wasted. I have never seen such a comparison while other tests are common, for example video encoder comparison tests, and it is difficult to discover something new in this area.

The paper is organized as follows. Description of the used encoder and videos is briefly described in Section II. Used objective quality metrics are outlined in Section III. Results obtained from objective metrics are evaluated, compared and discussed in Section IV. Finally, Section V. concludes the paper.

This paper was supported by the project no. CZ.1.07/2.3.00/20.0007 and by the BUT project no. FEKT-S-14-2177. The described research was performed in laboratories supported by the SIX project; no. CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation. Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.



a) Life *b) Tree*

Fig. 1. Thumbnails of videos used in the comparison.

II. Encoded Videos and Codec

A. Videos used for the comparison

For a set of tested videos it was necessary to choose different kinds of videos (see Fig. 1). The research was performed on two uncompressed video clips in Y4M format. This format is defined as RAW, that means that each frame is stored as sequence of pixels encoded in the YUV color space.

TABLE I. PARAMETERS OF VIDEOS USED IN COMPARISON

Name	Frames per second [Fps]	Frames [-]	Time [s]
Life	30	825	27,5
Tree	50	500	10

Two videos with different spatial and temporal indexes were chosen, different frames per second and one is synthetic video. If the quality of video samples should be comparable, the samples must have the same value of bit rate. There were defined only bitrate and profile without any other modification and without tune options. Variance of the bitrate set was between -2 and 3 %. This is enough small error for our measurement.

B. Codec H.265 HEVC and implementations

Reference implementation HM for encoding video by H.265 was not chosen because it is extremely slow in the encoding. The better solution is to use x265 implementation of HEVC codec because this encoder has a much higher speed of encoding and impact on quality is negligible [2]. Implementation x265 is a H.265/HEVC video encoder application library, designed to encode video or images into an H.265/HEVC encoded bit stream. x265 is the most widespread HEVC encoder. This implementation is still in research and the new releases of this software are available on reference [3]. The 64-bit version of this encoder was chosen.

TABLE III. PARAMETERS OF USED ENCODER HEVC

Implementation	x265
Built date	24.1.2015
Compiler	GCC 4.6.3
Encoder version	1.4+424-2b93cf2a5ac8

Parameters like a bitrate, output name and quality were set by MATLAB script used to effective and automated encoding. Parameters for implementation are in TABLE II. Parameters of predefined quality profiles of used encoder HEVC are in TABLE III.

Encoder is able to measure time needed to encode video by itself. This is big advantage for our measurement. Table below describes parameters which are changing with predefined profiles.

TABLE III. PARAMETERS OF QUALITY FOR ENCODER x265

Profile	1	2	3	4	5
	Ultrafast	Superfast	Veryfast	Faster	Fast
	6	7	8	9	10
	Medium	Slow	Slower	Veryslow	Placebo

Profile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
ctu	32	32	32	64	64	64	64	64	64	64
bframes	4	4	4	4	4	4	4	8	8	8
b-adapt	0	0	0	0	2	2	2	2	2	2
rc-lookahead	10	10	15	15	15	20	25	30	40	60
scenecut	0	40	40	40	40	40	40	40	40	40
refs	1	1	1	1	3	3	3	3	5	5
me	dia	hex	hex	hex	hex	hex	star	star	star	star
merange	25	44	57	57	57	57	57	57	57	92
subme	0	1	1	2	2	2	3	3	4	5
rect	0	0	0	0	0	0	1	1	1	1
amp	0	0	0	0	0	0	0	1	1	1
max-merge	2	2	2	2	2	2	3	3	4	5
early-skip	1	1	1	1	0	0	0	0	0	0
fast-intra	1	1	1	1	1	0	0	0	0	0
b-intra	0	0	0	0	0	0	0	1	1	1
sao	0	0	1	1	1	1	1	1	1	1
signhide	0	1	1	1	1	1	1	1	1	1
weightp	0	0	1	1	1	1	1	1	1	1
weightb	0	0	0	0	0	0	0	1	1	1
aq-mode	0	0	1	1	1	1	1	1	1	1
cuTree	0	0	0	0	1	1	1	1	1	1
rdLevel	2	2	2	2	2	3	4	6	6	6
tu-intra	1	1	1	1	1	1	1	2	3	4
tu-inter	1	1	1	1	1	1	1	2	3	4

Parameters of used computer are in TABLE IV. It is up to date average powerful home computer.

TABLE IVV. PARAMETERS OF USED COMPUTER

CPU	**Core/Threat**	**Frequency**	**RAM**
i5-35550	4/4	3.3 GHz	8 GB 1600 MHz

III. Objective Quality Metrics

In our research there were used only full reference objective metrics to evaluate videos. It was chosen only few most used metrics, three metrics and each of them represents their group of objective metrics. Why was not used the SSIM and PSNR measuring toolbox built-in x265 encoder in the stage of compression? Because x265 is mainly encoder, not a measuring program. The measuring algorithms implemented in the encoder can be optimized for speed and can be simplified. Or can have better measured value than the real value because than they look much more coding efficiency.

Metrics from external programs for measurement quality was used. There are listed and described objective metrics used in our test [4]:

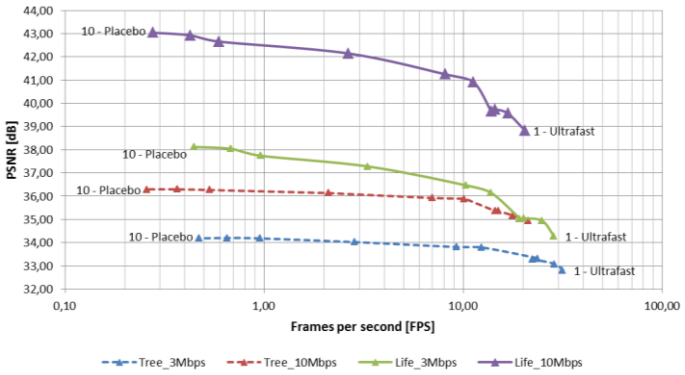
- PSNR (Peak Signal-to-Noise Ratio) - higher value in dB means higher video quality and it is usually between 25 and 50 dB [5].
- SSIM (Structural Similarity) - higher value of the index means higher quality. SSIM can take values from 0 to 1.
- VQM (Video Quality Metric) - it consists of linear combination of the seven parameters and final value is computed using the equation. VQM has the best similarity with subjective tests from mentioned quality metrics. Details of VQM are described in the article [6].

IV. Results and Evaluation

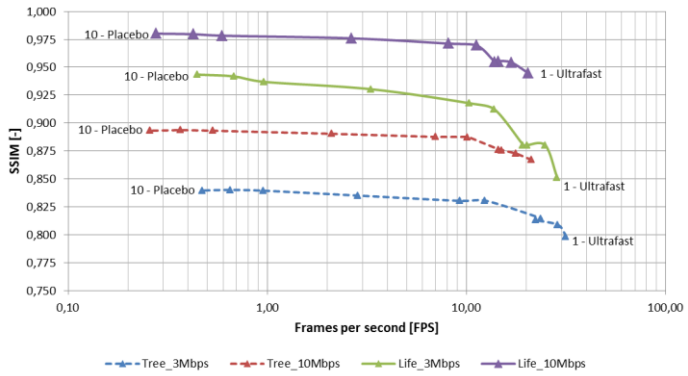
The lower the bitrate, the more important is the choice of better quality profile used. In other words, the lower bitrate which was set, the higher profile should be used for compression. Increase of image quality for quality profile better than Medium for SSIM for video *Tree* with bitrate 10 Mbps is negligible.

Medium profile was chosen by the x265 creators as the default profile. It could be said that this is not the best choice. The usage of slow profile provides considerable increase of quality at the expense of compression slow down measured with all the objective metrics. Therefore, I recommend using this profile as the default because it is the best compromise between the image quality and the time needed to compression. The difference in image quality between the profile Ultrafast and placebo is shown in Fig. 3.

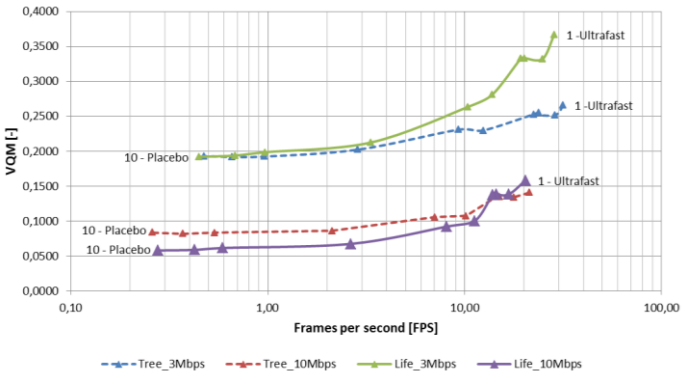
Faster profiles have small problems with utilizing all cores in Intel Core i5, speed can be also reduced by another components of computer because encoding speed is very fast. In case of Ultrafast maybe half of the possible time is used to compute. For longer videos differences in loading videos will not be so high. On horizontal axis in Fig. 2 is value of frames per second because it is independent on length of video. In the left side of graphs starts the best quality profiles and gradually to the right side are the faster profiles.



a) PSNR



b) SSIM



c) VQM

Fig. 2. Results of speed of encoding and image quality using objective metrics.

V. Conclusion

Placebo quality profile, as one might expect, did not provide quality improvement and taking into account the extremely long encoding time it should be used only by people who do not take care for time. Bearing in consideration all the results Slow profile is the best one in my own opinion. In case of higher quality profile, the quality increase is not too high to compensate the expenditure of time.

The bitrate increase of the output compressed videos prolongs the time of compression. Three times higher bitrate needs about twice long time to encode video.

b) Profile Placebo

b) Profile Ultrafast

Fig. 3. Video Life with bitrate 3 Mbps and zoom 200 % with different profile.

If the fast speed of coding is a priority I recommend choosing Superfast profile. The profile Very Fast and Faster provide very similar image quality like a Superfast but speed of coding is slower. Ultrafast profile is special case. This profile has extremely poor image quality, but needs the shortest time to encode. This profile is recommended only for some specialized case like a real time encoding on some low powerful devices where hardware encoding is not available. For example mobile phones which have not HEVC hardware encoder integrated directly to chip.

References

[1] Xiph.org::Derf\'s Test Media Collection. Xiph.org Video Test Media. [online]. [cit. 2015-05-04]. Available from: https://media.xiph.org/video/derf

[2] O. Zach, M. Slanina, "A comparison of H.265/HEVC implementations" ELMAR (ELMAR), 2014 56th International Symposium , vol., no., pp.1,4, 10-12 Sept. 2014 doi: 10.1109/ELMAR.2014.6923337

[3] x265 HEVC High Efficiency Video Coding H.265 Encoder. x265. [online]. [cit. 2014-12-07]. Available from: http://x265.org/index.html

[4] H.R. Sheikh and A.C. Bovik, "Image information and visual quality", IEEE Trans. on Image Processing, vol. 15, no. 2, pp. 430-444, February 2006.

[5] Software - VQMT: Video Quality Measurement Tool (École polytechnique fédérale de Lausanne). [online].[cit. 2015-05-04]. Available from: http://mmspg.epfl.ch/vqmt

[6] Software – ITS. Video Quality Metric (VQM) Software. [online]. [cit. 2015-05-04]. Available from: http://www.its.bldrdoc.gov/resources/video-quality-research/software.a

CST STUDIO SUITE 2015

From Components to Systems. Simulate, Optimize, Synthesize.

From the first steps to the finishing touches, CST STUDIO SUITE® is there for you in your design process. Its tools allow engineers to develop and simulate devices large and small across the frequency spectrum. The powerful solvers in CST STUDIO SUITE 2015 are supplemented by a range of new synthesis and optimization tools, which are integrated directly into the simulation workflow. These can suggest potential starting points for component design, allow these components to be combined into systems, and finally analyze and fine-tune the complete systems.
Even the most complex systems are built up from simple elements. Integrate synthesis into your design process and develop your ideas.

Choose CST STUDIO SUITE – Complete Technology for EM Simulation.

Filtering Antennas: Introduction of Different Concepts

Martin Kufa
Dept. of Radio Electronics
Brno University of Technology
Brno, Czech Republic
martin.kufa@phd.feec.vutbr.cz

Abstract—**The paper is focused on a comparison of a different approaches to design of a filtering antennas or a filtennas. In this paper, a combination of a band-pass filter integrated into SIW technology with an antenna or an antenna array is discussed. Integration of a radiating element as a last resonator of a band-pass filter is presented as well. An optimization of an antenna array without any filtering parts is mentioned as a last approach to design of filtering antennas. On the basis of presented approaches, a new way and new objectives to design of a filtering antenna arrays is defined.**

Index Terms—**Filtering antenna; filtenna; pass-band filter; equivalent circuit.**

I. Introduction

Nowadays, wireless devices play an increasingly important role. We use such devices in daily life almost continuously. In future, more and more new devices will use wireless connections. With the proliferation of new devices, we start to face two problems. First, existing frequency bands are overcrowded, and therefore, higher and higher frequency bands have to be used (that way, the growing demands on the transmission speed and the amount of data transferred can be met also). Second, flexibility and mobility of wireless devices have to be improved.

In order to minimize dimensions of mobile devices, we can implement fractal theory to the design of an antenna. The fractal design can also reduce the size of filters. Further reduction of the filter size can be achieved by using the concept of defected ground structures. Finally, an antenna and a filter can be integrated into a single, compact structure of minimal dimensions.

The combination of an antenna and a filter into a single, compact structure is called filtering antenna or filtenna. A simultaneous frequency and space filtering is the main task of the filtenna. In an ideal case, a band-pass filter does not need to be used on the receiving side because the filtering is done by the filtenna.

This paper reviews recent developments in the field of filtering antennas (filtennas). We will concentrate on three main approaches to the design of filtennas:

- A separate planar filter and a separate planar antenna are integrated. As a planar filter, a microstrip band-pass filter or a SIW (substrate integrated waveguide) band-pass filter are used. As a planar antenna, a patch antenna or a patch array are exploited.
- In the structure of a planar filter, the last resonator of a filter is replaced by a patch antenna or a monopole. Such a structure can be designed by a filter synthesis approach.
- The filtering antenna (the filtenna) is created only by radiating elements without any filtering parts. The frequency response of the realized gain is synthesized by optimization of all radiating elements, distances of the neighboring elements and amplitude and phase excitations of each individual element.

The following subchapters describe the methods in detail.

II. Combination of Band-Pass Filter and Antenna

The first part of the filtering antenna is created by a band-pass filter and the second one is formed by any radiating element. In this case, a band-pass filter and an antenna or antenna array are designed absolutely separately on one common substrate. Several papers using a combination of a band-pass filter integrated into SIW technology will be introduced in a following sub-section.

In [1], authors have described a third-order SIW inductive window filter which is coupled with a slot antenna or two-slot antenna array. The structure is called a SIW one-slot filtenna or a SIW two-slot filtenna, respectively. In [1], the slot antenna array without the SIW filter is shown it does not behave like a filtenna. A SIW two-slot filtenna is depicted in Fig. 1.

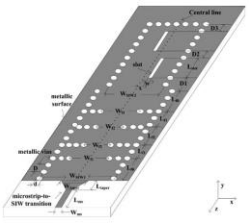


Fig. 1 The structure of the SIW two-slot filtenna [1].

The SIW filter can be combined with a Yagi antenna. In [2], a five-cavity SIW filter was connected with a six-element printed Yagi antenna consisting of a reflector, a dipole and

four directors. Layout of the described filtenna is shown in Fig. 2.

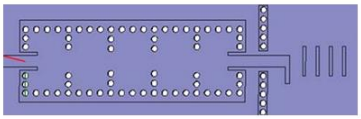


Fig. 2 The structure of the SIW filter–Yagi filtenna [2].

A structure consisting of a third-order cavity window band-pass filter integrated into SIW and a planar coaxial collinear (COCO) radiation element was described in [3]. In the first instance, authors presented a conventional planar COCO antenna which is created by serially fed patch antennas on both sides of a substrate. In the second instance, the third-order SIW band-pass filter was added before the COCO antenna. In the paper, the authors did not address the question of a frequency response of gain and space filtering. A top and a bottom layout of the SIW filter–COCO filtenna is depicted in Fig. 3.

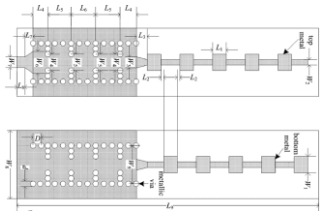


Fig. 3 The layout of the top side (upper) and the layout of the bottom side (below) of the SIW filter–COCO filtenna [3].

Tunable filtennas using a band-pass filter with a varactor were discussed in [4], [5] and [6]. In these papers, a microstrip band-pass filter with a varactor is connected with a wideband dual-side Vivaldi antenna. Thanks to the varactor, the filtenna can be tuned at required frequency in the frequency range from 6.1 GHz to 6.5 GHz. But authors did not discuss a spatial filtering in the frequency range. An example of the Vivaldi filtenna with the varactor is shown in Fig. 4.

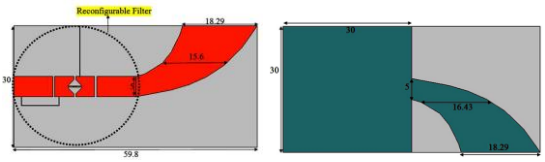


Fig. 4 A top layer (left side) and the bottom layer (right side) of the Vivaldi filtenna with varactor [6].

Slightly different approach to design the filtering antenna was presented in [7]. In this case, authors combine two third-order stopband filters with binomical balun and single dipole antenna. The filters were created by spiral resonators and were integrated into microstrip feeder (left part in Fig. 5). Due to this combination, the frequency response of the realized gain was suppressed out of the working band. The structure is shown in Fig. 5.

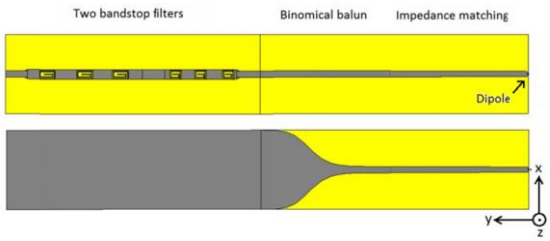


Fig. 5 The combination of two bandstop filters with binomical balun and single dipole antenna [7].

III. Integration of Radiating Element into Filter

Several papers describe a design of filtennas using a filter synthesis approach. Filtennas consist of a band-pass filter with an integrated radiating element as a last resonator of the filter. In this case, the filtenna is designed as complete device, but filtering and radiating parts can be clearly identified.

In [8], a filter consisting of three microstrip square open-loops was designed first. Second, the last open-loop resonator and an output port were replaced by a coupled line and a *Γ*-shaped antenna [8], respectively by center fed circular patch antenna and a coupled annular ring [9]. The coupled line was used as an admittance inverter with the length close to quarter of wavelength. The three square open-loops filter and the corresponding two square open-loops filtenna are shown in Fig. 6.

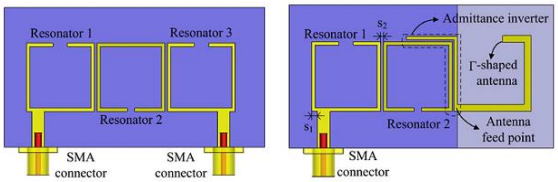


Fig. 6 The structure of the three square open-loops filter (on the left side) and the structure of the two square open-loops filtenna (on the right side) [8].

In [10], authors presented a microstrip patch antenna integrated into a band-pass filter. The band-pass filter was created by a three-pole hairpin filter and connected to a patch antenna representing the fourth pole. The patch played the role of the last resonator of the filter. In [10], an equivalent circuit including a patch and filter approach for design of the filtenna was designed. That way, dimensions of the whole structure were partially reduced. A layout of this filtenna is shown in Fig. 7 and the equivalent circuit of the filtenna is given in Fig. 8.

Fig. 7 The photo of the filtenna with the hairpin filter [10].

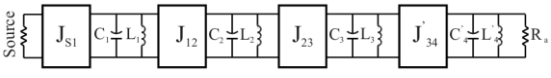


Fig. 8 The equivalent circuit including the patch as the last resonator of the filter [10].

In [11], a filtering microstrip *U*-shaped antenna with a *T*-shaped resonator in a feeder was published. A gain response of the combination of a *U*-shaped antenna and a *T*-shaped resonator behaves like a second-order quasi-elliptic band-pass filter with two zeroes at the band edges. The *U*-shaped antenna works as a radiation element and as well as the second resonator in a filter structure. A gap between the *T*-shaped resonator and *U*-shaped antenna forms an admittance inverter (*J*-inverter). The design of the filtering antenna was based on a circuit approach and a synthesis of a band-pass filter. The equivalent circuit and the layout of the *U*-shaped filtenna are shown in Fig. 9.

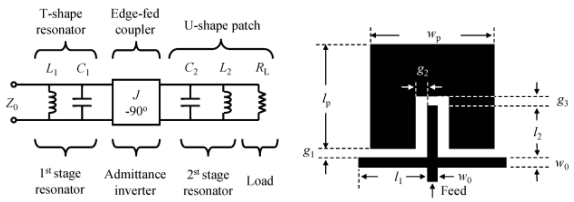


Fig. 9 The equivalent circuit (left side) and the layout (right side) of the filtering *U*-shaped antenna [11].

The design of a filtering microstrip antenna array based on [11] was described in [12] and [13]. A filtering array is created by four *U*-shaped patch antennas and a very sophisticated feeder which together behaves like a third-order band-pass filter. The first pole is formed by an *E*-shaped power divider which is coupled with an external 50 Ω feeder by the first gap (first *J*-invertor). The second pole is created by a half-wavelength balun which is attached by the second gap with the *E*-shaped divider (second *J*-inverter). And the last pole is generated by *U*-shaped patch antennas coupled with the balun by a gap (last *J*-invertor). Design of the filtering *U*-shaped antenna array is based on a circuit approach and synthesis of a band-pass filter. An equivalent circuit of the filtering antenna array and a layout are depicted in Fig. 10.

In [12] we can found a comparison of the filtering *U*-shaped antenna array and a conventional antenna array which is connected with a hairpin band-pass filter.

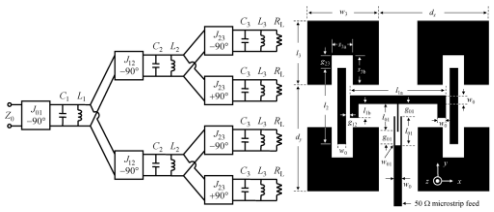


Fig. 10 The equivalent circuit (left side) and the layout (right side) of the filtering *U*-shaped antenna array [12].

A printed meander-line antenna with quarter-wavelength resonator filter which operate together as the second-order filtenna was published in [14]. The meander-line antenna plays the role of radiating element but also the last segment of the band-pass filter. The first segment of the band-pass filter is designed as a shunt resonator. Thanks to the shunt resonator, a frequency response of antenna gain has two extra transmission zeroes. A structure of the meander-line filtenna is depicted in Fig. 11.

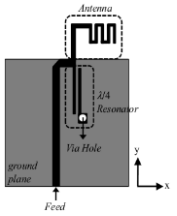


Fig. 11 The structure of the meander-line filtenna [14].

The described issue was followed in [15] and [16]. The design procedure was identical with [14] but the meander-line antenna was replaced by the *Γ*-shaped antenna and the quarter-wavelength resonator was replaced by a defected ground plane (DGS) resonator. The DGS resonator created the first segment of the second-order filtering antenna and the *Γ*-shaped antenna formed the last segment of the filter (see Fig. 12).

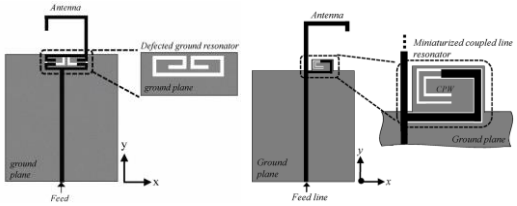


Fig. 12 The layout of the *Γ*-shaped filtenna with DGS designed in [15] (left side) and [16] (right side).

A similar design of a filtering antenna was presented in [17]. The filtering antenna was created by an inverted *L*-shaped antenna and a quarter-wavelength resonator. The structure was placed on one side of a substrate and the resonator was separated from the ground plane by a slot as shown in Fig. 13.

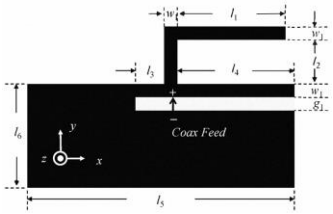


Fig. 13 The structure of the inverted *L*-shaped filtering antenna [17].

In [18], authors designed a filtering antenna which was composed by a parallel coupled microstrip line band-pass filter and an inverted *L*-shaped antenna. The inverted *L*-shaped antenna was designed as the last resonator of the parallel coupled microstrip line band-pass filter. Authors compared results of the filtering antenna and the combination of a parallel coupled microstrip line filter which is conventionally connected with the inverted *L*-shaped antenna. A structure of the *N*-order filtering antenna is shown in Fig. 14.

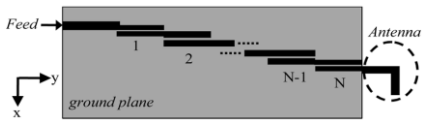


Fig. 14 The structure of the *N*-order filtering antenna [18].

A low-pass filter with reduced fractal defected ground structure (DGS) [19]–[21] can be used as a basic element to design of a filtenna [22]. The low-pass filter consisted of the 11th order low-pass filter with stubs separated by transmission lines on the top side of the substrate and six units of reduced fractal Minkowski DGS connected by narrow slots on the bottom side of the substrate. This filter exhibits two passband ranges and owing to the addition of three capacitive elements at the end of the first, third and fifth stubs and removed the second output port, the filter behaves like the filtenna at the second operating band (Fig. 15) [22].

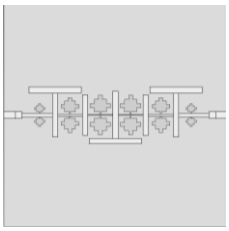


Fig. 15 Filtenna based on the low-pass filter with reduced fractal defected ground structure (DGS) [22].

IV. Filtering Antennas with Synthesized Realized Gain

The last main approach to the design of filtennas is based on the antenna array, where the frequency response of the realized gain is synthesized by optimization of the dimensions, amplitude and phase excitations and distances between neighboring radiating elements. Due to described approach, authors could shape the frequency response of the realized gain and obtain the required response.

In [7], authors described the series fed patch antenna array consisting of five radiating elements (Fig. 16). In the structure, each patch antenna (length and width of the patches) and each transmission line (length and width of the feeders) have different dimensions, which were the subjects of the optimization process. The optimized whole structure behaves like filtering antenna at the band 5 GHz.

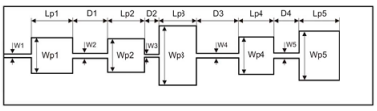


Fig. 16 Series fed five patch antenna array [7].

Another way to synthesis of the realized gain of the antenna array was presented in [7] and [24] as well. In this case, authors assumed an idealized antenna model without couplings between neighboring radiating elements. The idealized antenna array was optimized by particle swarm optimization and the objects of the optimization were: distances among radiating elements, amplitudes of the all elements and phases of the excitations sources. The idealized antenna array is illustrated in Fig. 17.

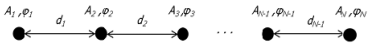


Fig. 17 Idealized antenna array [7].

The synthesis of a dipole antenna array was described in [23] and [24]. In this case, authors used a multi-objectives self-organizing algorithm. The synthesis of the dipole antenna array was divided into two main parts. Firstly, authors optimized the antenna array without any feeder among neighboring dipoles (Fig. 18 left). The objects of the optimization were amplitudes and phases on inputs of the each dipole and its length. Distance between array and a ground plane was set to fixed length equal to quarter of a wavelength.

Secondly, the feeder among the dipoles and ground plane were considered to the model (Fig. 18 right). In this configuration, goals of the optimization were achieved owing to changes of the amplitudes and phases of the excitation current, distances among neighboring dipoles and its lengths. Distance between antenna array and the ground plane was same as in the previous model. Due to described models, authors could shape gain and level of side lobes (in the first case) and control the responses of the gain and the reflection coefficient (in the second case).

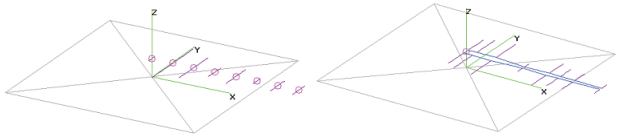


Fig. 18 Dipole antenna array without feeders (left) and with feeder (right) [23].

V. Conclusions

This paper clearly shows that approaches to the design of filtering antennas can be divided into three main branches:

- The first class of approaches connects a band-pass filter and an antenna into a compact device on a common substrate. In some cases, the authors do not address the question of space filtering, and the filtenna is a device connecting the filter and the antenna on common substrate only.
- In the second class of approaches, the filtering antenna is created by a band-pass filter with a radiating element on the position of the last resonator in the filter topology. These filtering antennas are designed using a band-pass filter synthesis approach. But even in this case, filtering and radiating parts of the structure can be identified.
- The third class of approaches, the filtering antenna array is consisted only from radiating elements (patches or dipoles) and frequency filtering is obtained by optimization of the all parameters in the whole antenna array. Described approach does not include the filtering theory into the design of the filtering antennas or the filtennas.

Up to now, authors have not used the band-pass synthesis for the design of the filtering antenna without filter parts. And in this case of approach, the synthesis of the shape of the frequency responses of the realized gain and reflection coefficient and control of the direction of the main lobe can be a new topic for research too.

REFERENCES

[1] Ch. Yu, and W. Hong. "37-38 GHz substrate integrated filtenna for wireless communication application." Microwave and Optical Technology Letters. 2012, vol. 54, issue 2, pp. 346-351.

[2] S. Yu, W. Hong, Ch. Yu, H. Tang, and J. Chen. "Integrated millimeter wave filtenna for Q-LINKPAN application." 2012 6th European Conference on Antennas and Propagation (EUCAP).

[3] Ch. Yu, W. Hong, Z. Kuai, and H. Wang. "Ku-band linearly polarized omnidirectional planar filtenna." IEEE Antennas and Wireless Propagation Letters. vol. 11.

[4] Y. Tawk, J. Costantine, and C. G. Christodoulou. "A varactor-based reconfigurable filtenna." IEEE Antennas and Wireless Propagation Letters. vol. 11.

[5] Y. Tawk, M. E. Zamudio, J. Costantine, and C. G. Christodoulou. "A cognitive radio reconfigurable "filtenna"." 2012 6th European Conference on Antennas and Propagation (EUCAP).

[6] M. Zamudio, Y. Tawk, J. Costantine, J. Kim, and C. G. Christodoulou. "Integrated cognitive radio antenna using reconfigurable band pass filters." Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP), vol., no., pp.2108-2112, 11-15 April 2011.

[7] D. Wolansky, P. Vsetula, Z. Raida, and P. S. Hall. "Antennas with synthesized frequency dependency of gain." 2012 6th European Conference on Antennas and Propagation (EUCAP). 2012.

[8] W. J. Wu, Y. Z. Yin, S. L. Zuo, Z. Y. Zhang, and J. J. Xie. "A new compact filter-antenna for modern wireless communication systems." IEEE Antennas and Wireless Propagation Letters. vol. 10.

[9] W. J. Wu, J. Wang, R. Fan, and Q. Zhang. "A broadband low profile microstrip filter-antenna with an omni-directional pattern." Antennas and Propagation (APCAP), 2014 3rd Asia-Pacific Conference on , vol., no., pp. 580–582, 26–29 July 2014

[10] J. Verdu, J. Perruisseau-Carrier, C. Collado, J. Mateu, and A. Hueltes. "Microstrip patch antenna integration on a bandpass filter topology." In proc. 12th Mediterranean Microwave Symposium (MMS2012), no. EPFL-CONF-179874. 2012.

[11] Ch.-K. Lin, and S.-J. Chung. "A compact filtering microstrip antenna with quasi-elliptic broadside antenna gain response." IEEE Antennas and Wireless Propagation Letters. vol. 10.

[12] Ch.-K. Lin, and S.-J. Chung. "A filtering microstrip antenna array." IEEE Transactions on Microwave Theory and Techniques. 2006, vol. 59, issue 11, s. 487-515.

[13] S.-J. Chung, and H.-N. Wang. "Compact multi-function antennas designed using filter synthesis technique." 2012 42nd European Microwave Conference (EuMC), vol., no., pp.1331-1334, Oct. 29 2012-Nov. 1 2012

[14] Ch.-T. Chuang, and S.-J. Chung. "New printed filtering antenna with selectivity enhancement." European Microwave Conference (EuMC 2009), vol., no., pp.747-750, Sept. 29 2009-Oct. 1 2009.

[15] Ch.-T. Chuang, and S.-J. Chung. "A new compact filtering antenna using defected ground resonator." 2010 Asia-Pacific Microwave Conference Proceedings (APMC), 2010 vol., no., pp.1003-1006, 7-10 Dec. 2010.

[16] Ch.-T. Chuang, and S.-J. Chung. "A compact printed filtering antenna using a ground-intruded coupled line resonator." IEEE Transactions on Antennas and Propagation. vol. 59, issue 10, pp. 3630-3637.

[17] Ch.-K. Lin, and S.-J. Chung. "A compact simple structured filtering antenna utilizing filter synthesis technique." 2010 Asia-Pacific Microwave Conference Proceedings (APMC), vol., no., pp.1573,1576, 7-10 Dec. 2010

[18] Ch.-T. Chuang, and S.-J. Chung. "Synthesis and design of a new printed filtering antenna." IEEE Transactions on Antennas and Propagation, vol.59, no.3, pp.1036-1042, March 2011.

[19] M. Kufa, and Z. Raida. "Lowpass filter with reduced fractal defected ground structure." Electronics Letters. 2013, vol. 49, no. 3, pp. 199-201.

[20] M. Kufa, and Z. Raida. "Comparison of planar fractal filters on defected ground substrate." Radioengineering, vol. 21, no. 4, pp. 1019-1024, December 2012.

[21] M. Kufa. "Planar fractal filters on defected ground substrate." Diploma thesis. Brno University of technology, 2012.

[22] M. Kufa, and Z. Raida. "Design of filtenna with fractal defected ground structure." European Conference on Antennas and Propagation (EuCAP), 2013 7th , vol., no., pp. 1278-1280, 8-12 April 2013.

[23] P. Vsetula, and Z. Raida. "Dipole antenna array with synthesized frequency dependency of gain and reflection coefficient." 2013 International Conference on Electromagnetics in Advanced Applications (ICEAA). 2013.

[24] P. Vsetula. "Antenna arrays with synthesized frequency response of gain." Doctoral thesis. Brno University of Technology, 2014.

Study of Radiation Properties of SIW Horn Antenna at 60 GHz

Jiří Lambor
Department of Radio Electronics
Faculty of Electrical Engineering and Communication, Brno University of Technology
Brno, Czech Republic
xlambo01@stud.feec.vutbr.cz

***Abstract*—This paper deals with methods for improving radiation properties of an SIW horn antenna at 60 GHz. Effect of metal block and dielectric extension in front of aperture are investigated. Final antenna has significantly improved radiation patterns in E-plane and H-plane. Matched bandwidth is 15% and the antenna exhibits maximal directivity 13.7 dBi at 60 GHz.**

Keywords—SIW horn antenna, 60 GHz, dielectric extension, metal blocks

I. INTRODUCTION

Wireless communication in 60 GHz band is suitable for short range communication, e.g. for Personal Area Network (PAN), due to the fact that millimeter waves demonstrate higher free space attenuation with respect to microwaves. This property allows operating with higher level of security and reduced interference with other wireless systems.

A conventional H-plane SIW horn antenna suffers from wide beam in the E-plane because the extension of an aperture is only in H-plane. Consequently, such an antenna has small gain and is not matched very well. To improve radiation properties of that antenna, there are few methods which can be used: (1) exploitation of a dielectric extension [1], (2) exploitation of metal blocks [2], and (3) usage of an air hole performed dielectric extension [3]. In this paper, the first two approaches are investigated.

II. DESINGN OF ANTENNA AND SIMULATION RESULTS

At first, a simple H-plane SIW horn antenna is used without any improving technique. The antenna is designed for the Arlon CuClad 217 substrate with the height of 1.52 mm. The design is done in CST Microwave Studio with the help of Matlab script [4] which follows conventional design approaches for common horn antenna. Drawing of the designed antenna is depicted in Fig. 1 and values of the parameters are in Tab. 1. The antenna is defined by the width W_H and length L_H of the aperture. The radiation patterns in E- and H-plane of this reference antenna are depicted in Fig. 2. The maximal directivity is 6.82 dBi and we can see wide main beam in E-plane.

TABLE I. REFERENCE ANTENNA DESIGN PARAMETERS

W [mm]	L_r [mm]	W_H [mm]	L_H [mm]	W_{feed} [mm]
8	3.3	5.1	1.3	2.5

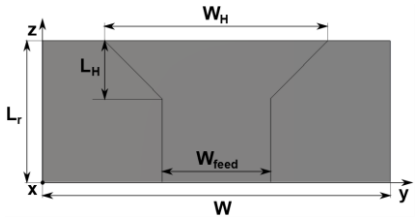


Fig. 1. Drawing of reference antenna

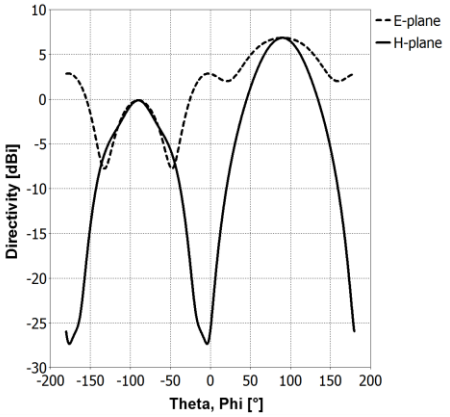


Fig. 2. Radiation patterns of reference antenna in E- and H-plane at 60 GHz

Further, from 1 to 5 metal blocks are placed in front of the aperture of the horn antenna which is showed in Fig. 3. For desired number of metal blocks, we look for maximal directivity and good matching by changing the gap of the width S_X between two nearby blocks and the length of the blocks which is denoted by L_{SX}. Resultant design parameters are in Tab. 2. The effect of the dielectric extension without metal blocks is also investigated. The extension is gradually added with length equal to metallic block L_{SX} plus gap between blocks S_X as parameter L_{SEQ}. Graphical dependence of the antenna directivity on the number of the metal blocks and length of the extension are depicted in Fig. 4.

TABLE II. PARAMETERS OF METAL BLOCKS AND GAPS

L_{S1} [mm]	L_{S2} [mm]	L_{S3} [mm]	L_{S4} [mm]	L_{S5} [mm]
0.9	1.1	1.3	1.5	1.5
S_1 [mm]	S_2 [mm]	S_3 [mm]	S_4 [mm]	S_5 [mm]
0.4	0.6	0.4	0.15	0.15

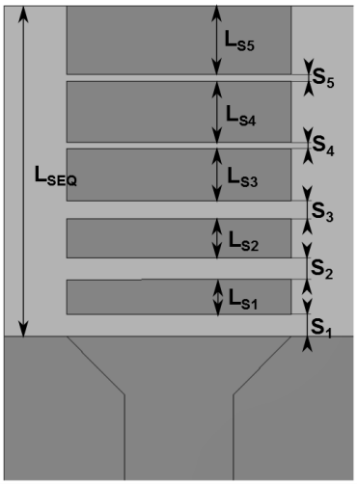


Fig. 3. Drawing of the antenna with metal blocks and dielectric extension

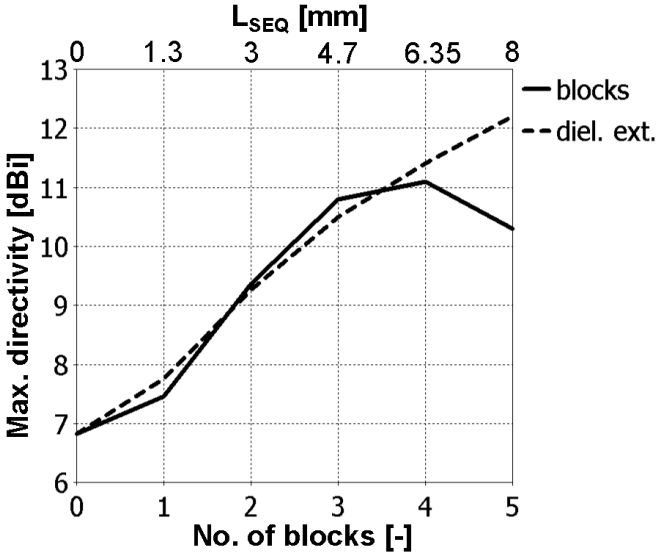


Fig. 4. Dependece of directivity on dielectic extension length (non-linear scale) and no. of blocks

As we can see in Fig. 4, the effect of the length of the dielectric extension has nearly linear dependence on directivity. Metallic blocks have non-linear dependence. Optimal value in this case is three. Above three blocks influence of the metal blocks is negligible on the value of the directivity and effect of the dielectric extension become more significant.

In next step, the influence of the dielectric extension on the directivity of the horn antenna with three metal blocks is investigated. The extension L_{ld} (Fig. 6.) is changed from 0 mm to 12 mm. The results of this investigation are depicted in Fig. 5. Obviously, usable range of this extension is from 1 to 9 mm, where the directivity increases relatively rapidly.

The value of 5 mm for L_{ld} parameter is chosen since there is maximal suppression of side lobes and also for reasonable overall dimensions of the antenna.

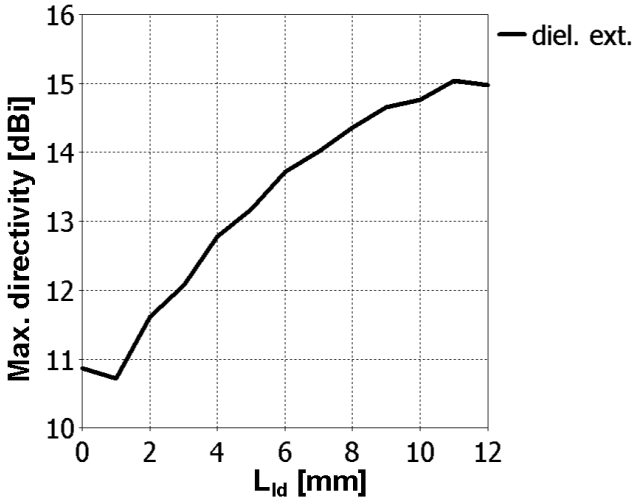


Fig. 5. Dependece of directivity on dielectic extension length with three blocks

Design parameters and drawing of the final antenna are in Tab. 3 and Fig. 6. Impedance matching of the antenna is tuned by L_{S3} and S_3 to have reflection coefficient below -10 dB in band near to 60 GHz.

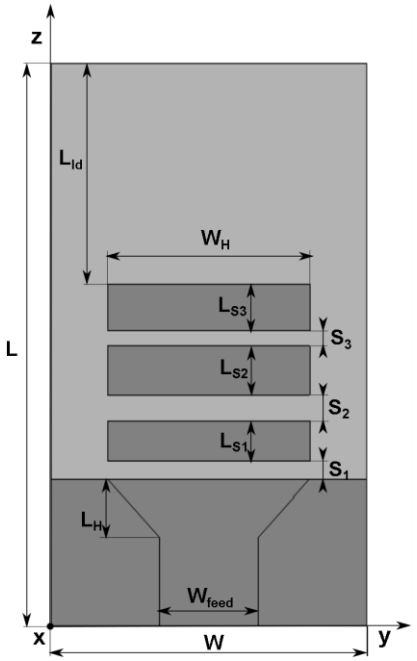


Fig. 6. Drawing of the final antenna

TABLE III. FINAL DESIGN PARAMETERS

W [mm]	L [mm]	W_H [mm]	L_H [mm]	W_{feed} [mm]	L_{ld} [mm]
8.0	12.8	5.1	1.3	2.5	5.0
L_{S1} [mm]	S_1 [mm]	L_{S2} [mm]	S_2 [mm]	L_{S3} [mm]	S_3 [mm]
0.9	0.4	1.1	0.6	1.05	0.35

The radiation patterns in E- and H-plane of the final antenna are depicted and compared with the reference antenna in Figs. 7 and 8. Comparison of the reflection coefficient is in Fig. 9.

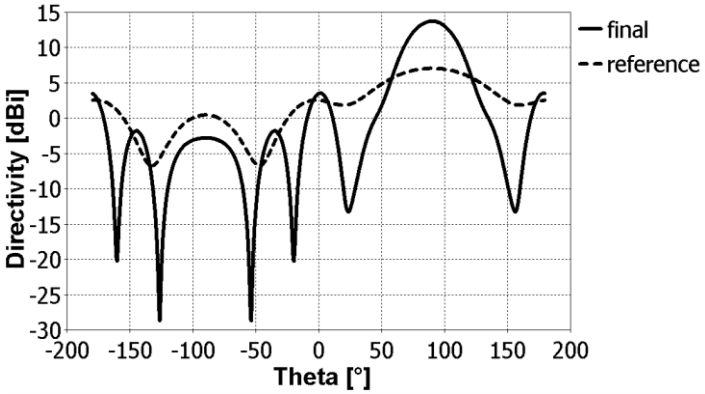


Fig. 7. Radiation pattern of the reference and final antenna in E-plane at 60 GHz

The maximal directivity of the final antenna is 13.7 dBi. The radiating pattern in H-plane is significantly improved in comparison with the reference antenna. The pattern in E-plane has also narrower main beam.

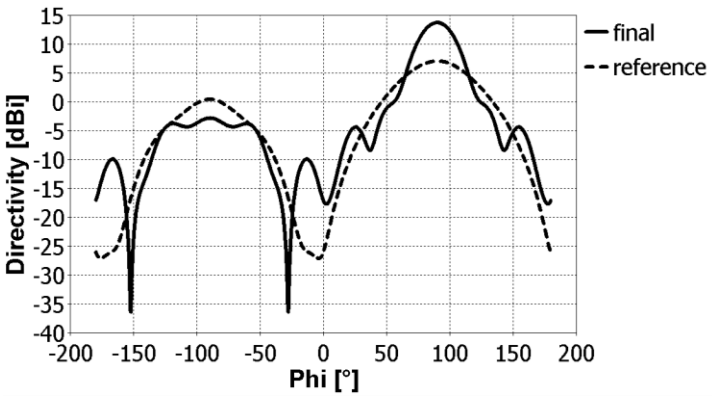


Fig. 8. Radiation pattern of the reference and final antenna in H-plane at 60 GHz

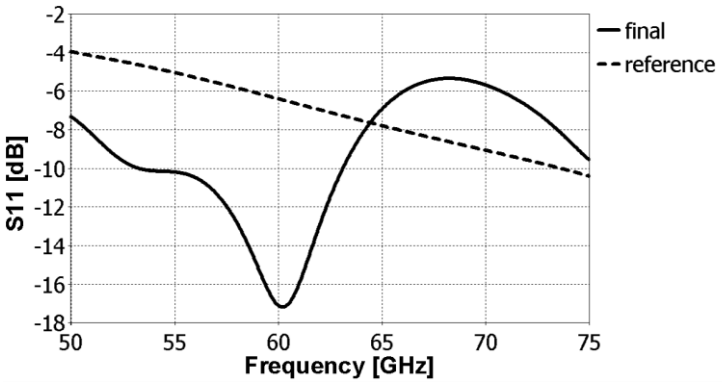


Fig. 9. Comparison of the reflection coefficient of the reference and final antenna

Reflection coefficient is below -10 dB in band from 55 GHz to 64 GHz. It is significant improvement when compared with reference antenna's reflection coefficient.

III. CONCLUSION

The effect of the metal blocks and the dielectric extension on the maximal directivity of the SIW horn antenna has been investigated. The optimal value of the blocks is three blocks in this case. The length of the dielectric extension can be used to further increase of the directivity, but we must consider enlarged overall dimensions of the antenna and level of the side lobes.

Final antenna achieved maximal directivity of 13.7 dBi with overall dimensions 8x12.8 mm, so the antenna is still relatively small. Compared to the reference antenna, the increase of the directivity is approximately 7 dB. The radiation patterns in E-plane and H-plane have been significantly improved. Matched bandwidth is 15 %.

REFERENCES

[1] H. Wang, D. G. Fang, B. Zhang, and W. Q. Che, "Dielectric loaded substrate integrated waveguide (SIW) H-plane horn antennas," IEEE Trans. Antennas Propag., vol. 58, no. 3, pp. 640–647, Mar. 2010.

[2] M. Esquius-Morote, B. Fuchs, J.-F. Zurcher, and J. R. Mosig, "A Printed Transition for Matching Improvement of SIW Horn Antennas," Submitted to IEEE Trans. Antennas Propag., 2012.

[3] C. Yang., Q. Zu-Ping, Z. Ying-Song, J. Jun, and C. Wen-Quan, "Bandwidth Enhancement of SIW Horn Antenna Loaded With Air-Via Perforated Dielectric Slab". In: IEEE Antennas and Wireless Propagation Letters. IEEE, 2014, p. 571 - 574. ISSN 1536-1225.

[4] P. Procházka, "Planární Parabolická Reflektorová Anténa", Master's thesis, Brno University of Technology, Brno, Czech Republic, 2015.

Implementation of Cyclostationary Based Detection

Ondřej Maňas, Roman Maršálek
Dept. of Radio Electronics
University of Technology, Brno, Czech Republic
Email: xmanas06@stud.feec.vutbr.cz

Ahmed Elsokary, Václav Valenta
Institute of Electron Devices and Circuits
Ulm University, Ulm, Germany
Email: Ahmed.Elsokary@uni-ulm.de

***Abstract*—Cognitive radios offer re-utilization of unused licensed channels by establishment of secondary networks based on dynamic spectral sharing. A key aspect of CR networks is proper exploring of unused frequencies. In this paper, a well known sensing method employing repetitive properties of OFDM based signal was selected for implementation. The method is fully implemented into special wideband sensing platform and evaluated under signals of real DVB-T services. Implementation steps and baseband preprocessing are described. Moreover, a new approach of implementation of the test metric calculation is introduced.**

I. Introduction

The spectral scarcity is recently one of the most challenging topics in communication. The broadcast environment is treated as a natural resource, usually managed by national organizations. Utilization of particular frequencies is based on licensing, which provides rights for usage of those frequencies only by one privileged user known as Primary User (PU). In terms of static frequency allocation, PU does not use the whole licensed bandwidth all the time and thus some part of PU's channels become underutilized [1]. As a result, those channels provide opportunities for re-utilization by unlicensed Secondary Users (SU). However, this can be done only in terms of ensured coexistence. In 1998 J. Mitola introduced a concept of Cognitive radio [2]. Observing white spaces, CR enables deployment of secondary networks based on dynamic sharing of frequency bands between PUs and SUs.

Primary deployed services operating on perspective frequencies mostly use OFDM technique e.g. DVB, DAB, LTE, UMTS, WiMAX, etc. Moreover, utilization of OFDM-based secondary users is even standardised in 802.11af as WLAN family standard. These essential facts are a motivation for research that involves sensing techniques dedicated for OFDM-based systems. A detection method proposed in [4] and [5] employs cyclostationary properties of OFDM streaming added by repetitive insertion of cyclic prefix (CP). This method seems to be promising in the category of narrow band detectors dedicated for fine sensing. The method can be also potentially used in white spaces monitoring performed on digital television broadcast channels, which are recently proved to be among the most opportunistic environments for secondary networks set-up.

In this paper, hardware implementation of a detector based on CP repetition will be described. A fully implemented core dedicated for a multi-cycle based test metric determination will be introduced. Moreover, a baseband preprocessing with fractional re-sampling will be introduced. The final design will be evaluated on real DVB-T 8k signals received in Ulm, Germany.

For the evaluation, a sensing platform developed directly for cognitive networks research and deployment will be used [6]. An RF front-end is based on three custom RFIC chips (Radio Frequency Integrated Circuit) realized in an 0.25 μm SiGe BiCMOS technology. The front-end uses a fully differential up/down-conversion heterodyne architecture suitable for wide-band receivers. As a result, the platform can achieve frequency range up to 4.5 GHz with a conversion gain more than 17.2 dB. A quadrature output with bandwidth 190 MHz is converted by dual-channel 14-bit 250 MS/s ADCs and processed by Xilinx Virtex 6 evaluation board.

A detector design description is spread into three sections. Sections II, III, IV include theoretical background as well as implementation description. Section V is dedicated to simulation and measurements. Finally, a conclusion is given in Section VI.

II. Baseband preprocessing

The selected detector operates under DVB-T baseband sampling rate F_{dvb} $\frac{64}{7}$ MHz (9.1428 MHz). On the other hand, the quadrature ADC output is clocked by sampling rate F_{adc} on frequency 245.76 MHz. Therefore, a preprocessing block has to be involved. For this purpose, a decimation filter has been realized. It consists of three stages. First two stages $P(z)$ and $Q(z)$ realize anti-aliasing filtering and down-sampling by 4, respectively. These two stages are followed by a last stage $G(z)$ realizing final filtering. Filters are designed in that way to follow standardised DVB-T transmitter spectral mask [3]. After decimation, the resulted sampling rate is 15.36 MHz. The decimation block consists of single-rate FIR (Finite Impulse Response) filters. These filters are generated by a development software ISE 14.7 supported by Xilinx.

To meet F_{dvb} required for proper detection, signal needs to be re-sampled by fraction factor of 1.68. For this purpose, Lagrange extrapolation implemented by Farrow structure has been employed. Farrow filter structure is realized by two FIR filters in parallel with coefficients based on 6th order Chebyshev low-pass filter. A sequence of fractional delay is repeated every 33 output samples. It means that no fractional delay estimation for every sample is needed. The sequence is stored in a ROM.

The baseband preprocessing system clock is F_{adc}. Sampling frequencies are emulated only by control signal activated on new valid data. However, output data system clock is $F_{adc}/2$. No clock crossing technique is employed, since both

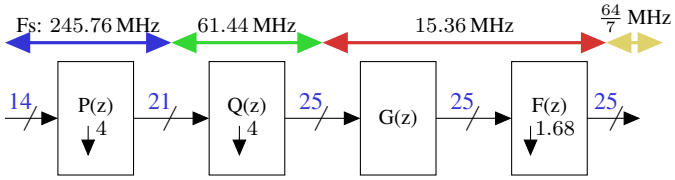


Fig. 1: Block diagram of the baseband preprocessing and the sampling frequency at different stages

clocks are synchronized in phase. Block diagram of the baseband processing with sampling rate plan and used arithmetic precision is depicted in Fig. 1.

III. Cyclostationary based detection

The detection of OFDM-based systems employs the presence of CP (cyclic prefix) inserted in every symbol. This periodic feature can be extracted by an autocorrelation of the signal with its delayed version. A delay τ is defined by symbol length without CP and transmitter sampling rate. In case of user presence, the autocorrelation contains a square wave with pulse-width equivalent to CP length (data are not correlated). Fourier analysis of that square wave figures out set of non-zero sine wave components at multiples of the lowest frequency. That frequency is in our case called a cyclic frequency α. The method of feature extraction described here, is known as Cyclostationary Autocorrelation Function (CAF) eq. (1) [4].

$$\begin{aligned}\hat{R}_x(k) &= \frac{1}{N}\sum_{n=0}^{N-1} x(n)x^*(n-\tau)\mathrm{e}^{-j2\pi kn/N} \\ &= X(k) + jY(k),\end{aligned} \tag{1}$$

where $x(n)$ is sensed signal, variable k is a set of frequency components for which is $\hat{R}_x(k)$ estimated - $\alpha \in k$.

A disadvantage of CP based detector is that for proper OFDM detection, a set of information is needed. Those are Carrier frequency, DVB-T mode (2k or 8k - τ) and CP length ($\frac{1}{\alpha}$).

A. Implementation

Basically, CAF can be implemented by a memory block, complex multiplier and FFT (Fast Fourier Transform), as depicted in Fig. 3. From right side, the delay τ is realized by one 36 Kb BRAM block. To make delay one OFDM symbol, 8192 memory rows are needed. Each memory row is first read out and and replaced at the next clock cycle by an actual sample. Between the complex multiplier and the FFT, the signal $x(n)x^*(n-\tau)$ is decimated by CIC filter with decimation factor $M = 32$. This step improves FFT resolution at lower frequencies where first α is situated. A number of k now depends on FFT width which is 2048. The FFT block is moved to another faster clock domain. A signal data are valid at the emulated sampling frequency $\frac{64}{7}$ MHz at system clock 122.88 MHz. The data are buffered in a FIFO memory and after that processed by the FFT. The FFT is clocked 245.76 MHz as well as the incoming data stream from the FIFO memory. In this configuration, Fourier transform requires

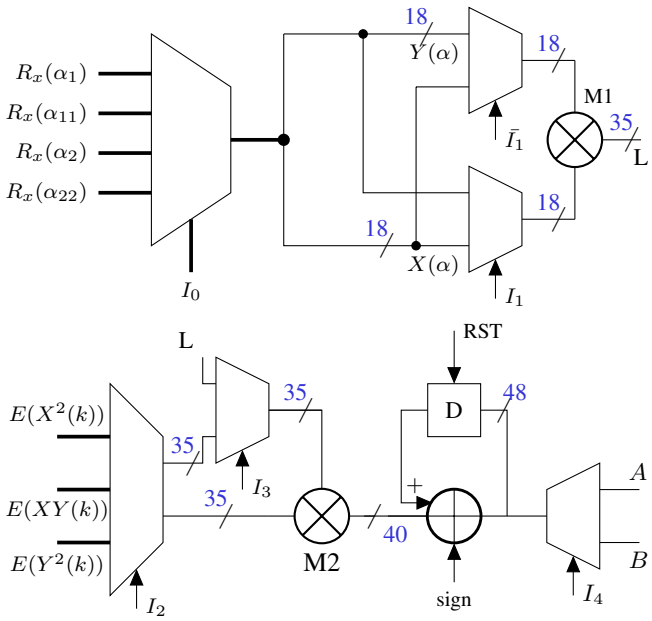


Fig. 2: Test metric core with used aritmetic precision

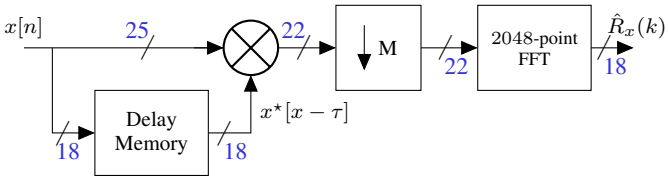


Fig. 3: CAF implementation according to [4] with chosen arithmetic precision

less time than in the case of the FFT clocked by the original system clock 122.88 MHz. This improvement also decreases the final sensing time. In the clock domain of 245.76 MHz, a test metric determination core described in next section is implemented as well. A DSP blocks as FFT, CIC and memories are generated by the development software. The designed fixed-point arithmetic precision are selected to fit into 25 x 18 hardware multiplier structure.

IV. Test metric

In [4] and [5], a test metric T_t is determined from 2 x 2 covariance matrix which follows

$$\begin{bmatrix} E\left[X^2(k)\right] & E\left[X(k)Y(k)\right] \\ E\left[X(k)Y(k)\right] & E\left[Y^2(k)\right] \end{bmatrix}, \tag{2}$$

where operation X^2, Y^2 and XY represent multiplication. The operator E can be rewritten as

$$E(x) = \frac{1}{N}\sum_{n=0}^{N-1} x(n). \tag{3}$$

Simplified final equation for T_t is described in eq. (4). An improved performance of the detector can be achieved by

$$T_t(\alpha) = \frac{X^2(\alpha)\hat{E}[Y^2(k)] + [Y^2(\alpha)\hat{E}[X^2(k)] - 2X(\alpha)Y(\alpha)\hat{E}[X(k)Y(k)]}{\hat{E}[X^2(k)]\hat{E}[Y^2(k)] - \left(\hat{E}[X(k)Y(k)]\right)^2} \tag{4}$$

involving also higher multiples of α. This method is known as multi-cycle detection [7]. For this purpose, we rename the lower cyclic frequency α to α_1 and first higher multiples of α to α_{11}. In the same manner we define negative cyclic frequencies as α_2 and α_{22}. A multi-rate test metric T_T is estimated according to eq. (5).

$$T_T = \frac{T_t(\alpha_1) + T_t(\alpha_2) + T_t(\alpha_{11}) + T_t(\alpha_{22})}{4} \tag{5}$$

By substitution of eq. (4) in the eq. (5), we obtain simplified eq. (6), where A is resulted numerator summation for each α_n. B is equal to denominator of eq. (4).

$$T_T = \frac{A(\alpha_1, \alpha_{11}, \alpha_2, \alpha_{22})}{4B} \tag{6}$$

The estimation of T_T contains 40 multiplying operations, 6 add/sub operations and one division. Multiplication by 2 are not counted. However, the estimation of the test metric T_T can be fitted into an ALU (arithmetic logic unit) containing two multipliers and one add/sub accumulator. All the ALU's inputs and output are controlled by a sequence of instructions stored in a ROM. Implementation of the division is described in the next paragraph. The test metric core is fitted to faster clock domain at frequency 245.76 MHz. A simplified block diagram of the proposed ALU is depicted in Fig. 2. Control signals I_n are part of the instruction word. Value truncations, scaling operations and pipe-lining registers are not included in the diagram.

A final hypothesis (signal present) is determined based on threshold T_{thr} and comparison as in eq. (7).

$$T_T < T_{thr} \tag{7}$$

By substitution of eq. (6) in the eq. (7), we obtain eq. (9). The division in the test metric is replaced by a multiplication of the threshold and the denominator B. This reduces the complexity by eliminating a large divider.

$$A(\alpha_1, \alpha_{11}, \alpha_2, \alpha_{22}) < 4T_{thr}B \tag{8}$$

V. Simulation and Measurements Results

A. Software simulation

For simulation purposes, a simple model of DVB-T transmitter (only data and CP) operating in AWGN channel was designed. The parameters are set as in the realistic signals (Mode = 8k, CP = 1/4 [8]). Further, the input signal is quantized to ADC 14–bit precision. To evaluate the detection performance, a simulation is performed on the software model of the detector. The expected performance is shown in Fig. 4. For a signal with -14 dB SNR, we notice a degradation in detection performance due to the use of fixed point arithmetic. The fixed point precision is included in Fig. 1, Fig. 3 and Fig. 2 as blue marks. The scale factors are determined from previous simulations to optimize detector sensitivity. The second model is implemented in FPGA. A receiver operation characteristic (ROC) curve of both detectors are presented in Fig. 4. The FPGA resource usage of the implemented detector is summarized in TABLE I.

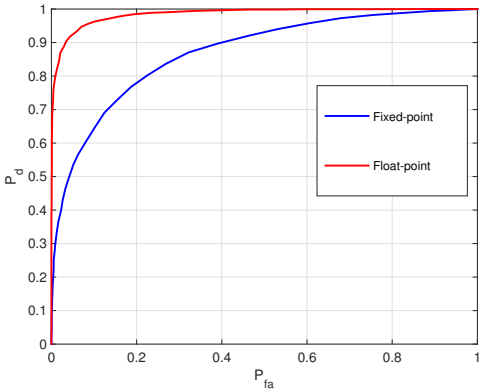


Fig. 4: ROC based on software models – SNR = -14.7 dB

TABLE I: Resources usage

Entity	Count	Percentage
Logic Slices	4397	1
DSP slices	46	6
Block RAM	15	3

B. Measurements

The functionality of the implemented detector was tested with real DVB-T channels. The list of all deployed DVB-T channels in Germany and their parameters are publicly accessible in [8]. For the first test, a DVB-T channel at the carrier frequency 482 MHz, with symbol length 8k and cyclic prefix 1/4 was selected. A channel power spectral density measured by spectrum analyzer is depicted in Fig. 5.

By settings the FFT-width and decimation factor DF as mentioned in section III, a frame of 8 OFDM symbols is needed. This sets the acquisition time to 8.1 ms, at a sampling frequency of $64/7$ MHz. The processing time is 100 μs. This includes the FFT operation and test metric estimation. Therefore, the final sensing time is 8.2 ms.

In the first step, CP features were successfully extracted. The spectrum of cyclic frequencies is presented in Fig. 6. Cyclic features related to CP repetition have been measured. The lowest cyclic frequency is at 892.85 Hz.

To allow an efficient threshold selection, the PDF of the test metric was evaluated for an H_0 and H_1 scenarios. For

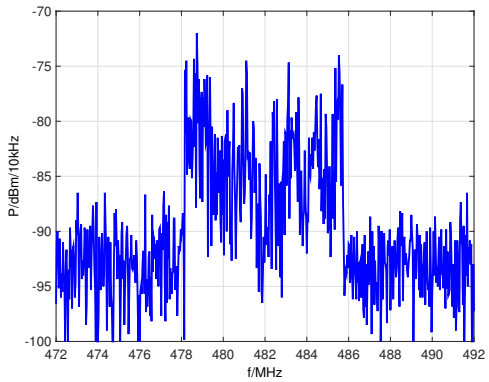


Fig. 5: Spectrum of the input signal from antenna measured by spectrum analyzer at 10kHz resolution bandwidth

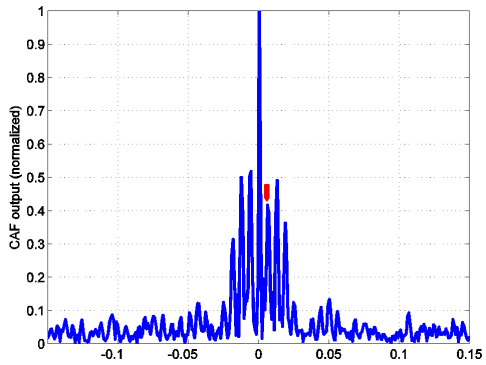


Fig. 6: Cyclic spectrum of received DVB-T signal - The lowest cyclic frequency is marked by a red arrow.

H_1, the receiver was tuned to an occupied DVB-T channel at 482 MHz. For H_0, the receiver was tuned to an unused channel. The measurements were done 4096 times and the resulting PDF of the resulting test metric is plotted in Fig. 7. From the PDF for H_0, we can observe typical probability distribution chi-square χ^2_2. By Neyman-Pearson test [5]

$$t = F^{-1}_{H0}(1 - P_{fa}) \tag{9}$$

for P_{fa} equal to 1 %. We have estimated threshold t from cumulative distribution function F_{H0}, which achieved P_d equal to 97.7 %.

The test is done on a real TV signal for The SNR is not precisely determined. The evaluation of the detection performance under various SNR conditions remains as a future task.

VI. Conclusion

A cyclostationary detector based on CP for DVB-T signals has been implemented. The detector achieves a maximum clock frequency of 245.76 MHz and a detection time of 8.2 ms. A new approach to efficiently implement the test metric evaluation on FPGA has been shown.

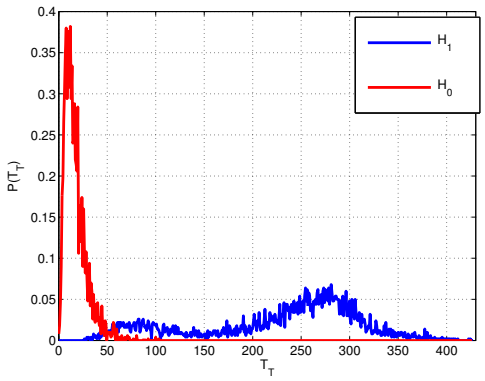


Fig. 7: PDF of the test metric T_T of an occupied (H_1) and non-occupied (H_0) DVB-T channel

Furthermore, the performance of the proposed detector implemented in floating point and fixed point arithmetic has been compared. The functionality of designed detector has been evaluated on DVB-T channels in real RF environment.

Acknowledgment

This work is partially funded by DFG, the German Research Foundation.

References

[1] V. Valenta, R. Maršálek, G. Baudoin, M. Villegas, M. Suarez and F. Robert, "Survey on spectrum utilization in Europe: Measurements, analyses and observations," in *Proc. of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks & Communications (CROWNCOM)* June 2010, pp. 1–5.

[2] J. Mitola, and G.Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications* Aug 1999, pp. 13–18.

[3] European Telecommunications Standards Institute, "Digital Video Broadcasting (DVB): Framing structure, channel coding and modulation for digital terrestrial television," *European Telecommunications Standards Institute*, ETSI-EN-300, 2011.

[4] V. Turunen, M. Kosunen, A. Huttunen, S. Kallioinen, P. Ikonen, A. Parssinen, J. Ryynanen,"Implementation of Cyclostationary Feature Detector for Cognitive Radios." *4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications, 2009. CROWNCOM '09.* June 2009, pp. 1–4.

[5] M. Kosunen, V. Turunen, K. Kokkinen, J. Ryynanen,"Survey and Analysis of Cyclostationary Signal Detector Implementations on FPGA," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* December 2013, pp. 541–551.

[6] P. Lohmiller, A. Elsokary, H. Schumacher, S. Chartier, S., "Towards a broadband front-end for cooperative spectrum sensing networks," *2014 44th European Microwave Conference (EuMC)* October 2014, pp. 147–150.

[7] J. Lunden, V. Koivunen, A. Huttunen, H.V. Poor, "Spectrum Sensing in Cognitive Radios Based on Multiple Cyclic Frequencies," *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications* August 2007, pp. 37–43.

[8] B. Heimermann, P. Jochum, O. Kluth, T. Khler, T. Lang, T. Mann-Raudies, J. Mezger, H. Nehse, V. Olischlger, M. Painter, S. Rettner, N. Schulze, "DVB-T – Sender in den einzelnen Bundeslndern, geordnet nach Programm–Ensemble"*irt.de* June 2015

Velocity Impact on Multi-hop Communication in MANET Environment

Martin Matis
Department of Electronics and Multimedia Communications
Technical University
Košice, Slovakia
martin.matis@tuke.sk

Ľubomír Doboš
Department of Electronics and Multimedia Communications
Technical University
Košice, Slovakia
lubomir.dobos@tuke.sk

Ján Papaj
Department of Electronics and Multimedia Communications
Technical University
Košice, Slovakia
jan.papaj@tuke.sk

***Abstract*—In this paper is described one of possible solutions for communication in areas without infrastructure, or in areas, where the infrastructure is functioning, but other type of communication between mobile nodes is preferred. Further there is analysed velocity impact on multi-hop communication. This type of communication is provided by multi-hop communication based on MANET (Mobile Ad-hoc Network) routing protocols with utilization of devices of daily use (smartphones, tablets and etc.). MANET routing protocol DSR (Dynamic Source Routing) was analysed in this paper and simulations were created based on this protocol. After analyse of simulations it is described, which type of environment and which level of velocity is better to use for DSR routing protocol.**

Keywords—MANET; DSR protocol; multi-hop communication; networks without infrastructure; inter-device communication

I. Introduction

Communication and transfer of information are one of the most important things between people. It is not a problem in our society to use global infrastructure network or radio links network, which is provided with some organizations or administrators. For this transfer of information is mostly used equipment of daily use, such as smartphones, laptops or tablets.

This system of communication is great till infrastructure is correct and functioning. In the situation like natural disaster it is impossible to use available devices of daily use. On the other hand, even if the infrastructure is functioning, the other type of transfer of data without infrastructure can be used, because it can be better or preferable. Our effort is utilizing of devices of daily use for both cases in the small areas in terms of hundreds of meters. Multi-hop communication can solve a problem in areas without infrastructure or in areas, where communication without infrastructure can be useful. Many methods can be used for communication and transfer of data (messages, pictures, videos, voices, and act.) without infrastructure, but we will focus on the only one method, which is called MANET.

We want to confirm usage of the one MANET routing protocol in cases, when no infrastructure exists and nodes or devices are mobile with different level of velocity and random movement. We want to compare a success of delivery for MANET routing protocol DSR (Dynamic Source Routing) for different velocities and different cases of occurrences source (S) and destination (D) nodes.

II. Communication without Infrastructure

Communication between people is matter of course in our society during everyday life. There are many supported devices (laptops, smartphones, tablets, and act.) which can transfer data by supported applications to everyone. The problem begins, when:

- The infrastructure is unusable for various reasons. For example the provider stops to support the services, the infrastructure is destroyed by natural disaster or a group of people is in an area with no cell coverage.
- The infrastructure is usable, but we want to transfer data in small area without this infrastructure, only from D2D (device-to-device). For example communication and transfer of data can be in progress between colleagues on the same floor of the building, or in the same building, on the conferences or in other places where high density of nodes is relative.

There are many possibilities how to communicate in those areas. A goal is to use devices of daily use with wireless interfaces, such as WiFi, Bluetooth, NFC, and etc. These devices will be able to create other type of network based on multi-hop communication. Therefore this communication can and knows to create a path from S to D using only other nodes in the neighbourhood based on many criterions which are specified by the routing protocols [1][2][3].

III. Multi-Hop Communication

A. Movement and Velocity

Mobile nodes have some patterns of movement, which can be for example random, forward learned or influenced by social behaviour.

Movement and velocity can form a network, which will be dense or sparse. We can know which type of routing will be the optimal solution for successful transmission based on this movement and velocity information [4][5].

B. Mobile Ad-Hoc Networks (MANET)

MANET is a collection of mobile nodes, which forms a temporary and dynamic network without aid of centralized

administration or standard support devices regularly available, such as conventional networks. These nodes generally have a limited transmission range and so each node seeks the assistance of its neighbouring nodes in forwarding packets and hence the nodes in MANET can act as both routers and hosts. Thus the node may forward packets between other nodes as well as run user applications. These types of networks are suitable for situations in which either no fixed infrastructure exists or deploying network is not possible. Mobile nodes will often be battery powered, which limits the capacity of CPU, memory and bandwidth. This will require network functions that are resource effective. MANETs have found many applications in various fields like military, emergency, conferencing and sensor networks. Each of these application areas has their specific requirements for routing protocols.

The unique feature of these protocols is their ability to trace routes in spite of a dynamic topology. In the simplest scenarios, nodes may be able to communicate directly with each other. However, MANET has to also support communication between nodes that are only indirectly connected by a series of wireless hops (multi-hops) through other nodes [6].

In general, MANET is a network where every node is potentially a router and every node is potentially mobile. MANET is not a traditional wired network because of presence of wireless communication and mobility, which requires that the routing protocols used in MANET should be based on new and different principles [1][7][8].

MANET is very practical solution for an area without infrastructure with many advantages like: a) no infrastructure and lower cost, b) mobility, c) decentralized and robust, d) easy to build paths and spontaneous infrastructure; but on the other side many disadvantages like: a) higher error rate, b) lower data rate, c) dynamic topology and scalability, d) security, e) high velocity [8].

C. Routing Solution for MANET

The main idea for routing protocols in MANET is to correctly and effectively estimate the path between a pair of nodes in the network. The routing protocols are created in view of the QoS parameters [1].

MANET network is dynamically changing network which always creates new connections between nodes and loses the old connections based on routing protocols. Those protocols can be categorized into tree basic groups [7]:

1) Reactive (Source-initiated): It represents a class of routing protocols where the route is created only when the source requests the route to the destination. The route is created through a procedure of route discovery which involves flooding the network with route request packets (RREQ) from start node to the destination, till it will be found. Once the route is formed or multiple routes are obtained to the destination, the route discovery process comes to the end. A route maintenance procedure maintains the continuity of the route in the time by the source. The main representatives of reactive routing are: DSR (Dynamic Source Routing), AODV (Ad-Hoc On-Demand Distance Vector), TORA (Temporally Ordered Routing Algorithm) and etc.;

2) Proactive (Table-driven): These protocols always maintain up-to-date information about routes from each node to every other node in the network. Routing information is stored in the routing table of each node and route updates are propagated throughout the network to keep the routing information as actual as possible. Most of the routing protocols are not suitable for highly dynamic networks due to extra control overhead, which is generated to keep the routing tables consistent and fresh for each node in the network. The main representatives of reactive routing are: DSDV (Destination-Sequenced Distance-Vector), OLSR (Optimized Link State Routing), WRP (Wireless Routing Protocol) and etc.;

Hybrid: This routing scheme combine elements of reactive and proactive protocols. The main representatives of reactive routing are: ZRP (Zone Routing Protocol).

Another division of routing protocols can be: singlepaths or multipaths. For MANET routing protocols were created much more detailed divisions and many categories based on their property and utilization like: Location-aware, Hierarchical, Multicast, Geographical Multicast or Power-aware.

D. DSR Routing Protocol

DSR routing protocol (RFC 4728)[7] is one of based of MANET routing protocol. It is simple and efficient, on-demand routing protocol designed for usage in multi-hop wireless ad-hoc networks. DSR allows the network to be completely self-organisable and self-configurable. The protocol is composed from two main mechanisms of "Route Discovery" and "Route Maintenance", which allow discovering path from S to D and retaining connection between them during overall communication [7][8].

Two main mechanisms, which are used in the DSR routing protocol are as follows:

1) Route discovery: S sends route request packet (RREQ) to all neighbour nodes. These nodes send RREQ to their neighbours and etc., until the D is found. RREQ packet stores and carries the nodes' ID through which the RREQ was sent. If the finding was successful, the destination node sends route replay packet (RREP) back to S using the stored path, respectively paths. When the RREP arrives to S, than the path is established and S can use this path for sending the data packets;

2) Route maintenance: It is responsible for finding alternative path in such a situation when an actual used path is interrupted during sending of packets. This process utilizes backup path or paths, which was found during Route discovery. If some of nodes detects failed link, than it will send Rout Error message (RERR) to S, which tries to use alternative path selected from route cash if it exists. If alternative path does not exist standard maintenance process will be stopped and sending of packets is cancelled, so the transfer is unsuccessful.

Algorithm 1: Standard DSR routing protocol:

1. **Assumption**: Every node is initially located in specified area. Source (S) node wants to transfer a message to Destination (D) node based on DSR routing;
2. **Initialization**: Start positions are generated or are loaded for every node. Source and destination node are set. A number of packets (NoP) are set for a message;
3. S node sends RREQ due to finding paths to D;
4. **For** *S* wants to send the message which composes from packets (*Packets*=1→*NoP*)
5. **If** *Source* obtains number of RREP packets ≤ 1, **do**
   -*S* waits and tries to send RREQ again. Only two attempts are permitted for one message. Then, the transfer will be unsuccessful;
6. **Elseif** *Source* obtains number of RREP packets > 1, **do**
   -*S* starts to send a packet *P*;
   End
7. **If** *S* obtains RERR packet, **do**
   -The connection is lost and the path is destroyed in the next time slot;
   -All backup paths are checked, due to loosing connection;
   -*S* tries to use backup path/s for next packet, because number of RREP packets > 1;
8. **If** backup path is valid, **do**
   -*S* uses this valid path or other valid path for the packet;
9. **If** The backup path/s, which is/are used, is/are valid during the transfer of message, **do**
   -*S* uses only this/those path/s and the message will be transferred by this/those path/s;
10. **Elseif** The backup path/s, which is/are used, is/are not valid, **do**
    -*S* tries to find new path one more time.
    End
11. **Elseif** No backup paths are valid, **do**
    -*S* tries to find a new path one more time.
    End
12. **Elseif** The path, which is used, is valid during the transfer of message, **do**
    -S uses only this path and the message will be transferred by this path.
    End
    End (For)

IV. Simulation and Results

Our confirmation of usage of MANET routing protocol for the network without infrastructure was simulated in Matlab environment, where we set values of parameters (TABLE I).

Our simulation was orientated on DSR routing protocol and its success of delivery of messages depends on velocity for these situations (Fig. 1):

- S and D were in the same subnetwork (island), where end-to-end paths could be used from start of simulation and
- S and D were in different subnetworks (islands), where maintenance process had to be used.

TABLE I Set of Parameters for DSR Simulation

Parameter	*Value*	*Description*
Number of nodes	100	Nodes, which were located and used for simulation.
Radio range [m]	130	It determines other devices in the neighbourhood of nodes.
Area [m]	1000x1000	It is a start space, where the mobile devices can occur.
Number of repetition	50	It is for getting average values
Velocity [m/s]	1,4 2,8 5,6 11,2 22,4	-Slow walking. -Fast walking -Running/cycling -Driving in a city -Driving out of a city
Packets	5	A message composed of packets.

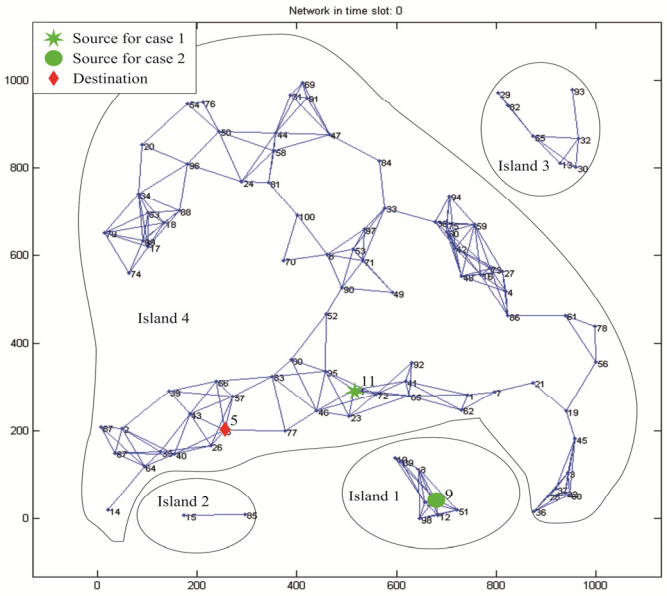


Fig. 1 Start positions of nodes with radio range for simulations, which display sources and destination nodes for both simulated situations

A. Success of Delivery

Success of delivery (Fig. 2a) had decreasing trend due to increasing velocity. Velocity had destroying impact on the success of delivery for DSR routing. Maintenance process crashed due to velocity and many founded paths were interrupted. Success of delivery was higher, when the communication started in one island, because maintenance process could use backup paths or find some new paths. The success of delivery (TABLE II) was much lower in the case, when the simulation started between nodes, which were in different islands, because first finding of paths was unsuccessful and only maintenance process in next time slot could find new paths to destination, when islands with S and D were connected.

B. Average Number of Transmitted Packets

Average number of transmitted packets (Fig. 2b) had decreasing trend due to velocity. When S and D were in the same network, the number of transmitted packets was higher as in the situation when S and D were in the different islands at start of simulations (TABLE III).

TABLE II Success of Delivery for Two Described Situations

Between nodes	**Success of delivery [%]**				
	Velocity [m/s]				
	1,4	***2,8***	***5,6***	***11,2***	***22,4***
S-11,D-5	99,6	92,8	63,6	37,2	31,2
S-9,D-5	21,2	14	5,6	5,2	2

TABLE III Average Number of Transmitted Packets

Between nodes	**Average number of transmitted packets**				
	Velocity [m/s]				
	1,4	***2,8***	***5,6***	***11,2***	***22,4***
S-11,D-5	4,98	4,64	3,18	1,86	1,56
S-9,D-5	1,06	0,7	0,28	0,26	0,1

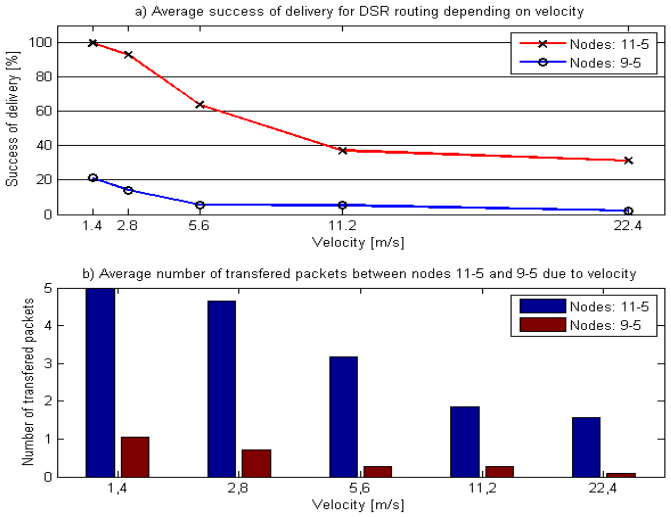


Fig. 2 a) Average success of delivery for simulations between nodes 11-5 and 9-5, which represent two simulated situations; b) Average number of transmitted packets during simulations

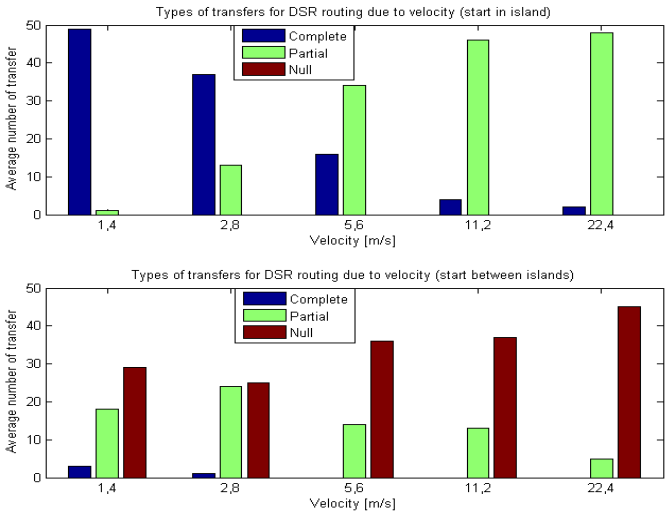


Fig. 3 Average numbers of types of transfers for situation, when S and D a) were in the same subnetwork and b) when were not

C. *Average Number of Types of Transfers*

These results display, how and in which quantities the types of transfers occur (TABLE IV). All transfers were divided into three types of transfers:

1) Complete transfer
2) Partial transfer
3) Null transfer

Number of Complete transfer was the highest when the velocity was the lowest and S and D were in the same subnetwork at the start of simulation (Fig. 3a). The number of Complete transfer was decreasing due to velocity but on the other hand, the number of Partial transfer (only part of the all messages, some packets was transmitted) was increasing due to velocity. When S and D were in different subnetworks (Fig. 3b), then the number of Complete transfer was minimal. The number of Null transfer was increasing due to velocity.

TABLE IV AVERAGE NUMBER OF TYPES OF TRANSFERS

Between nodes		**Average number of types of transfers**				
		Velocity [m/s]				
		1,4	2,8	5,6	11,2	22,4
S-11 D-5	Complete	49	37	16	4	2
	Partial	1	13	34	46	48
	Null	0	0	0	0	0
S-9 D-5	Complete	3	1	0	0	0
	Partial	18	24	14	13	5
	Null	29	25	36	37	45

V. CONCLUSION

In this paper was presented one of possible solutions, based on which it is possible to communicate in areas without infrastructure. It is MANET reactive routing protocol DSR, which is usable for these types of networks, but not for all simulated situations. We analysed two main start situations, when the S and D were in the same subnetwork and when S and D were not in the same subnetwork, but in different islands. For all start situations the movement with different level of velocity was generated.

DSR routing protocol functioned with high success (more than 99%) in situation, when maximum level of velocity was set on 1,4 m/s and S and D were in the same subnetwork at the start of simulation. In all other cases, when the velocity was increasing, the success of delivery was decreasing. In the case, when S and D were in the different subnetworks, the success of delivery was lower as in other situation and decreasing with increasing velocity. The highest average number of transmitted packet was during transfer with velocity 1,4 m/s and for situation, when S and D were in the same subnetwork. The result of types of transfers displays very low quantities of Complete transfer (full message) during transfer, when S and D were in different subnetworks at the start of simulations. Quantities of Complete transfer were increasing, when S and D were in the same subnetwork at the start of simulation.

ACKNOWLEDGMENT

The research presented in this paper was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project VEGA 1/0075/15.

REFERENCES

[1] S. Basagni, et al., eds. Mobile Ad Hoc Networking: The Cutting Edge Directions. Vol. 35. John Wiley & Sons, 2013.

[2] D. Johnson, Y. Hu, D. Maltz. The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4. Vol. 260. RFC 4728, 2007.

[3] H. Nishiyama, N. Kato, "Relay-by-smartphone: realizing multihop device-to-device communications," Communications Magazine, IEEE , vol.52, no.4, pp.56,65, April 2014

[4] V. Athanasios, at al.: Routing in Opportunistic Networks.Springer, 2013.

[5] Ch. Mayer, O. P. Waldhorst. "Offloading infrastructure using Delay Tolerant Networks and assurance of delivery." Wireless Days (WD), 2011 IFIP. IEEE, 2011.

[6] Ch. Raffelsberger, H. Hellwagner. "Combined Mobile Ad-hoc and Delay/Disruption-tolerant Routing.", 13th International Conference, ADHOC-NOW 2014, Benidorm, Spain, June 2014 Proceedings: 1-14

[7] A. Boukerche, et al. "Routing protocols in ad hoc networks: A survey." Computer Networks 55.13 (2011): 3032-3080.

[8] T. P. Singh, D. Vikrant, S. Maheshwari. "ADHOC NETWORKS: AN ANALYTICAL OVERVIEW." January 2012.

Hi-Speed USB Communication with FPGA

Roman Mego
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
roman.mego@phd.feec.vutbr.cz

***Abstract*—This paper is dealing with the implemented interface in FPGA for the USB communication using FTDI chip. The main goal is to determine the maximum data rate in direction from FPGA to PC. The interface is in form of the peripheral for soft processor. The paper introduces all possible ways of using the FTDI chip. Three of them were physically tested with different conditions, which can influence the final performance.**

Keywords—USB; Hi-Speed; FTDI; FT232H; FPGA; Altera Cyclone IV; data transfer; communication channel

I. Introduction

The field-programmable gate arrays (FPGA) are the semiconductor devices based on the configurable logic blocks, which can be programmed for desired function in the application. The modern FPGAs can also contain memory blocks, phase-lock loops (PLL), hard-wired processors or dedicated digital signal processing blocks, analog to digital converters (ADC) etc. [1], [2]. This makes the FPGAs ideal for the special applications with the small volume of production, where the use of the application specific integrated circuits (ASIC) is not economical, or where the system upgrade is desired. Typical application of the FPGAs is in the communication systems, signal processing, data acquisition or control systems.

The FPGAs are able to process the data with high bandwidth using the high level of parallelism. In some applications, the data presentation or the post-processing in personal computer (PC) is required. For this reason, the fast communication channel is needed. One of the most accessible is the universal serial bus (USB), where the USB 3.0 can reach the data rate up to 5 Gbit/s [3]. This article will be dealing with the USB 2.0, where the maximum data rate is 480 Mbit/s [4].

This paper is divided as follows. Section 2 introduces the hardware used in the benchmark. Section 3 shows the possible modes of communication in the USB module. Section 4 deals with the achieved speed in the different modes.

II. System Hardware

The tested system is realized on the custom development board with the FPGA development board and the UM232H development module [5]. The whole system is shown on Fig. 1.

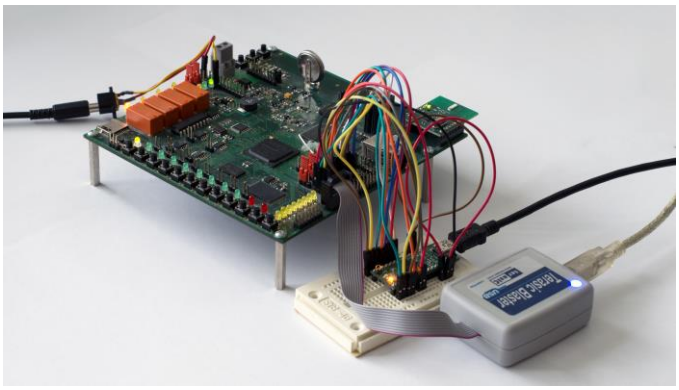


Fig. 1. Photography of the tested system.

A. FPGA Development Board

The development board is based on the Altera's Cyclone IV EP4CE40 [6]. This device is low-power FPGA suitable for low-cost applications. The overview of the basic parameters is in table I. The board also contains IS42S32400 SDRAM [7], which can be used as soft-processor memory, or for the other purpose. The next devices which can be found on the board, such as flash memory, real-time clock, Ethernet controller or SD card slot, are not used in the following application.

TABLE I. EP4CE40 Overview

Logic elements	39 600
Embedded memory (Kbit)	1 134
Embedded 18x18 multipliers	116
General-purpose PLLs	4
Global Clock Networks	20
User I/O Banks	8
Maximum user I/O	532

B. UM232H Development Module

The UM232H is USB to serial/FIFO development module utilizing the FT232H integrated circuit [8]. The FT232H is the single channel USB 2.0 bridge handling signals, which are converted to UART, parallel FIFO or synchronous serial protocols.

The advantage of the FT232H, or general of FTDI chips, is the availability of the drivers and libraries for the all most widely used operating systems like Windows, including CE and RT versions, Linux, Mac OS and Android.

III. Possible Modes of Communication

As it was previously mentioned, the FTDI chips support several ways of the communication. In general, it can be divided into two basic groups from the view, how the system sees the device:

- Virtual COM port (VCP),
- D2XX interface.

When the device is in the VCP mode, it appears as the legacy COM port. The maximum data rate is 3 Mbit/s. In the second mode, the device can be accessed through the proprietary libraries. In this mode, the chip can be configured as:

- Multi-Protocol Synchronous Serial Engine (MPSSE) [9],
- 245 FIFO [10],
- CPU FIFO,
- Fast Serial Interface [11] and
- FT1248 Dynamic Parallel/Serial Interface [12].

MSPSEE provides the flexible way to create the synchronous serial interface such as SPI, I2C or JTAG. The data format and the clock reaction can be configured in the application through the commands. In this mode, the data rate can be up to 30 Mbit/s.

The 245 FIFO mode provides compatibility with the FT245 synchronous or asynchronous interface. In the synchronous mode, the data is synchronous with the 60 MHz generated by FTDI. This clock corresponds to the maximum speed of USB 2.0. The data rate is up to 40 Mbyte/s, which is the maximum provided by the FTDI chips for USB 2.0. The drop is caused by the chip itself, because it needs some bandwidth to handle the communication. It has also one disadvantage. When it is used on the dual-channel device, the second channel cannot be used, because it uses all resources (memory and channel bandwidth) on the chip. It can be bypassed by using the asynchronous mode, but the maximum data rate is 8 Mbyte/s. The CPU FIFO is similar to the 245 asynchronous FIFO, but the difference is in the control signals.

The original purpose of the Fast Serial Interface is for data transfer through the optical isolation. It is kind of the synchronous serial interface similar to the SPI. The maximum clock frequency is 50 MHz, but the transmission of one byte includes the start and source bits, where the source bit identifies the channel. Maximum data rate is theoretically 5 Mbyte/s.

TABLE II. Pin usage on FTDI in Different Modes

Mode	Input	Output	In/Out
MPSSE	1	3	0
Synchronous 245 FIFO	3	3	8
Asynchronous 245 FIFO	2	2	8
CPU FIFO	4	0	8
Fast Serial Interface	2	2	0
FT1248	2	1	8

The FT1248 is another kind of the synchronous interface with 8 bi-directional data pins, which can be dynamically configured as 1, 2, 4 or 8-bits wide. The maximum clock frequency is 30 MHz.

Table II shows the overview of used pins on FT232H in the different modes. It can be seen, that only MPSSE and Fast Serial Interface are suitable for the simple galvanic isolation.

IV. Results

From the previously mentioned modes, only three were implemented in the FPGA for testing. These modes are:

- Fast Serial Interface,
- Asynchronous 245 FIFO Interface and
- Synchronous 245 FIFO Interface.

First two interfaces were implemented as the peripheral for the Nios II processor [13]. The third one was downloaded from the OpenCores [14]. All of these peripherals contains transmit and receive FIFOs, and have practically the same structure (Fig. 2).

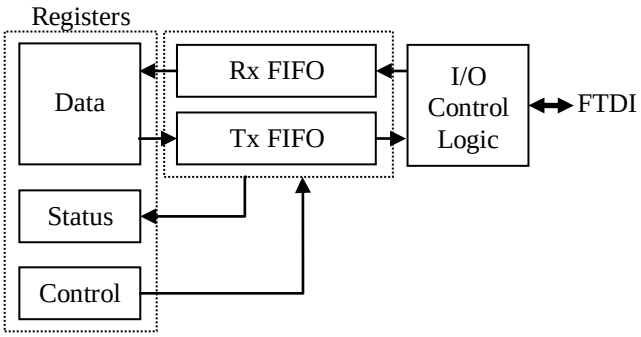


Fig. 2. Structure of the interface peripheral.

A. Fast Serial Interface

The first tested peripheral is the Fast Serial Interface. The expected 10 clock cycles for one byte has been increased to 13 due to delay of the detection of the start bit and the clear to send (CTS) signal. For this reason, the bit rate of the data transmission fell from the theoretical 5 Mbyte/s to the 3.8 Mbyte/s.

This speed can be achieved only when the size of the transmitted data is less than 512 bytes, which is also size of the internal buffer of the FT232H. When the FPGA tries to transmit bigger amount of the data, the FTDI chip starts data flow control using the CTS signal (Fig. 3). After that, the data rate fall to the 1.1 Mbyte/s.

Fig. 3. Captured communication from FPGA to FTDI.

B. Asynchronous 245 FIFO Interface

In the asynchronous FIFO mode, the bitrate is constant with the variable data length. There is another limitation. The bitrate in the direction from the FPGA to PC can be limited by the speed of the Tx FIFO feeding. The test was performed with the different system clock and also with the different settings of the compiler optimization.

From the view of the compiler optimization, there is only difference when the optimization is enabled or disabled. The level of the optimization does not have the impact on the data rate.

From the view of the system clock, it is obvious, that the processor will not be able to fill the FIFO in the real time. In the table III is shown, that the processor with the 48 MHz system clock fills the FIFO slower than the output logic is able to read it.

With the 100 MHz clock, the channel on the FIFO's output is practically fully utilized.

TABLE III. COMMUNICATION SPEED OF THE ASYNCHRONOUS 245 FIFO MODE, DEPENDING ON THE CLOCK SPEED AND COMPILLER OPTIMIZATION

System Clock	Achieved Data Rate			
	-o0	*-o1*	*-o2*	*-o3*
48 MHz	1.84 MB/s	6.83 MB/s	6.83 MB/s	6.83 MB/s
100 MHz	3.84 MB/s	7.88 MB/s	7.88 MB/s	7.88 MB/s

C. Synchronous 245 FIFO Interface

The conditions during the tests performed with the synchronous were the same as with the asynchronous interface. Now, the bit rate is limited only with the ability of the processor to fill the buffer. There is also no space for increasing of the system clock, because the maximum frequency of the Nios II processor is about 100 MHz, depending on the speed grade of the FPGA and number of connected peripherals. The achieved data rates are shown in the table IV.

TABLE IV. COMMUNICATION SPEED OF THE ASYNCHRONOUS 245 FIFO MODE, DEPENDING ON THE CLOCK SPEED AND COMPILLER OPTIMIZATION

System Clock	Achieved Data Rate			
	-o0	*-o1*	*-o2*	*-o3*
48 MHz	1.91 MB/s	9.56 MB/s	9.56 MB/s	9.56 MB/s
100 MHz	3.98 MB/s	19.39 MB/s	19.39 MB/s	19.39 MB/s

V. CONCLUSION

This paper showed one of the possible ways to implement high-speed communication channel from the FPGA to the PC. The use of the FTDI chips is one of the best solutions, because there is no need to write drivers for the host operating system and implementing the USB stack for the device, without the significant speed loss.

The standard serial interfaces such as UART and SPI can be used for lower bit rates. For higher bit rates, the one of the proprietary interfaces should be used.

The first tested interface was Fast Serial Interface suitable for optical isolation. This mode can be theoretically used for transfers up to 5 Mbyte/s, but it is practically able for transfers around 1 Mbyte/s with bigger block of data.

The second interface was in the asynchronous 248 FIFO mode. This mode can be used to transfers up to 8 Mbyte/s, what was proven by the measurement. There is also space for other operations on the side of the Nios II processor.

The last one was the synchronous 248 FIFO interface. Theoretical data rate is up to 40 Mbyte/s, which is close to the maximum speed of the USB 2.0. When this interface was driven by the software, the achieved data rate was nearly 20 Mbyte/s. This speed could be probably higher with the direct memory access (DMA) block in the system.

REFERENCES

[1] Altera Corporation. (2015, June 12). Cyclone V Device Overview [Online]. Available: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/hb/cyclone-v/cv_51001.pdf

[2] Altera Corporation. (2015, May 4). MAX 10 FPGA Device Overview [Online]. Available: https://www.altera.com/en_US/pdfs/literature/hb/max-10/m10_overview.pdf

[3] USB Implementers Forum, Inc. (2013, July 1) Universal Serial Bus 3.0 Specification [Online]. Available: http://www.usb.org/developers/docs/documents_archive/usb_30_spec_070113.zip

[4] USB Implementers Forum, Inc. (2015, July 2). Universal Serial Bus Revision 2.0 specification [Online]. Available: http://www.usb.org/developers/ docs/usb20_docs/usb_20_0702115.zip

[5] Future Technology Devices International, Ltd. (2012, October 24). UM232H Single Channel USB Hi-Speed FT232H Development Module Datasheet [Online]. Available: http://www.ftdichip.com/Support/Documents/DataSheets/Modules/DS_UM232H.pdf

[6] Altera Corporation. (2014, April 14). Cyclone IV Device Familly Overview [Online]. Available: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/hb/cyclone-iv/cyiv-51001.pdf

[7] Integrated Silicon Solution, Inc. (2010, March 4). IS42S32400 4M x 32 128Mb Synchronous DRAM [Online]. Available: http://www.issi.com/WW/pdf/42S32400.pdf

[8] Future Technology Devices International, Ltd. (2015, March 23). Single Channel Hi-Speed USB to Multipurpose UART/FIFO IC [Online]. Available: http://www.ftdichip.com/Support/Documents/DataSheets/ICs/DS_FT232H.pdf

[9] Future Technology Devices International, Ltd. (2010, March 12). FTDI MPSSE Basics [Online]. Available: http://www.ftdichip.com/Support/Documents/AppNotes/AN_135_MPSSE_Basics.pdf

[10] Future Technology Devices International, Ltd. (2010, March 5). FT2232H Used In An FT245 Style Synchronous FIFO Mode [Online]. Available: http://www.ftdichip.com/Support/Documents/AppNotes/AN_130_FT2232H_Used_In_FT245%20Synchronous%20FIFO%20Mode.pdf

[11] Future Technology Devices International, Ltd. (2009, October 21). FT2232D/H Fast Opto-Isolated Serial Interface Mode [Online]. Available: http://www.ftdichip.com/Support/Documents/AppNotes/AN_131_FT2232D_H_Fast%20Opto-Isolated%20Serial%20Interface%20mode.pdf

[12] Future Technology Devices International, Ltd. (2011, February 16). FT1248 Dynamic Parallel/Serial Interface Basics [Online]. Available: http://www.ftdichip.com/Support/Documents/AppNotes/AN_167_FT1248_Parallel_Serial_Interface_Basics.pdf

[13] Altera Corporation. (2015, April 2). Nios II Classic Processor Reference Guide [Online]. Available: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/hb/nios2/n2cpu_nii5v1.pdf

[14] OpenCores. (2014, March 8). FT2232H USB Avalon Core [Online]. Available: http://opencores.org/project,ft2232hcore

Higher-Order Modes in Dielectric Resonator Antennas

Michal Mrnka
Department of Radio Electronics
Brno University of Technology
Brno, Czech Republic
xmrnka01@stud.feec.vutbr.cz

***Abstract*—The review article summarizes the current endeavors in the field of higher-order mode utilization for gain enhancement of dielectric resonator antennas (DRA). Brief introduction into the field of DRAs is given as well as the explanation of the electromagnetic mode nomenclature for such resonators. Various application areas with frequencies ranging up to the millimeter and sub-millimeter bands are mentioned.**

Keywords—dielectric resonator antennas; higher order mode; directive antennas

I. INTRODUCTION

Until 1980s the dielectric resonators with high quality factors Q had been used as circuit elements in microwave filters and oscillators designs only [1-3] but the utilization of the electromagnetic (EM) fields escaping the resonator was firstly proposed by Long. et. at in 1982 [4]. Ever since then, the interest of researchers in this type of the new antenna element has grown rapidly and the number of publications dedicated to the DRA field still grows annually. Very good reference materials can be found in [5-8].

The operation of the DRA is based on the excitation of the proper electromagnetic (EM) mode inside the dielectric resonator. When the relative permittivity of the resonator is reduced as compared to the resonators in microwave circuits, it is easier for EM radiation to escape the resonator; thus the relative permittivity lies between 5 and 30 mostly, if the resonator is to be used as an efficient radiator. Three main shapes of the resonators has been studied analytically and experimentally (i.e. cylindrical [4] (see Fig. 1), rectangular [9] and hemispherical [10-11]), but other shapes based on the notches[12], combination of the shapes, stacking up the resonators [13-14] etc. have been proposed in order to improve the performance of the simple shaped DRAs. Moreover, introducing inhomogeneity [15] and anisotropy [16] has also revealed certain improvement in the DRA performance mostly in the terms of the input impedance and the directivity. Although, the DRA can be excited by a plethora of excitation schemes [5] the most frequently used solution is the excitation by an aperture coupled microstrip line [17]. This solution provides symmetrical excitation, reasonable bandwidth and simple fabrication. In certain applications DRA feeding by an aperture in a Surface Integrated Waveguide (SIW) might be more advantageous [18].

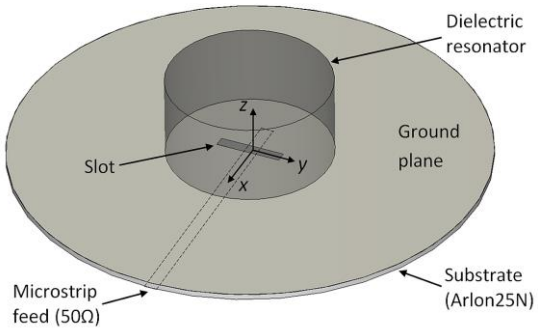


Fig. 1. Geometry of an aperture-fed cylindrical DRA

II. ELECTROMAGNETIC MODES IN DIELECTRIC RESONATORS

A dielectric resonator can resonate with infinite number of modes, but for a specific combination of material properties, dimensions and operating frequency only a finite number of potentially existing modes is allowed to be excited.

A simple cylindrical DRA is in the majority of cases excited either with the transverse magnetic $TM_{01\delta}$ (also designated as TM_{011}) mode providing vertical electric dipole like radiation pattern or with the hybrid electromagnetic $HEM_{11\delta}$ (also known as HEM_{111}) mode which provides radiation pattern similar to the horizontally placed magnetic dipole [5],[8]. Equivalent modes with analogous radiation patterns can be excited in the rectangular DRA as shown in [5] and [20].

Nevertheless, in the last decade, many higher-order modes have been identified and successfully excited. In particular, Avadanei et.al. in [21] and [22] showed which modes can be excited in a cylindrically shaped DRA of high dielectric constant ε_r=82.7 when the resonator was fed by a centrally placed and off center placed slot in the ground plane. The resonant frequencies, radiation patterns and efficiencies of these modes were discussed and compared. Provided the modes are spaced closely in the frequency and show similar radiation patters and if lower dielectric material is used for the antenna construction, the multiband operation can be utilized for increasing the antenna impedance bandwidth. In this case, the higher-order mode/modes is/are excited in addition to the low-order mode. For example a DRA described in [23]

operated with two hybrid modes $TE_{\delta 11}$ (sometimes designated as TE_{111}) and $TE_{\delta 13}$ (TE_{113}) over an impedance bandwidth exceeding 40%.

Multiband operation can be achieved if the modes are further away in the frequency spectrum; by adjusting the geometry of the resonator and/or the excitation scheme multiple modes can be excited at different frequencies of interest. When an omnidirectional radiation pattern is desired at two various frequencies, the lower frequency can be covered with TM_{011} mode of a cylindrical resonator and the higher frequency with the TM_{012} mode of the same resonator. Such a design was presented by Yong et.al. in [24] and the antenna was able to cover the 3.5 GHz WiMAX and the 5.8 GHz WLAN bands with similar omnidirectional radiation patterns. Dual band DRA antennas based on the dual mode operation with broadside radiation patterns were analyzed by Fang in his doctoral thesis [25]. Fang focused mainly on rectangular DRAs operating with TE_{111} and TE_{113} modes and he proposed design formulas for calculation of the dual band DRA dimensions when the operating frequency, dielectric constant and dimension rations were given [26]. A cylindrical DRA with dual band operation and a broadside radiation pattern in both frequency bands was developed by Sun and Leung in [27]. They used a cylindrical resonator operating in the HEM_{111} and HEM_{113} modes to provide broadside radiation patterns at 1.8 GHz and 2.4 GHz.

Stand-alone utilization of the higher order modes in DRAs was proposed only recently with the main motivation being the increase of the DRA gain (see Chapter 1.3). On the contrary, the multiband/wideband structures utilized the higher order modes in addition to the low-order modes. Nevertheless, the antennas working exclusively with the higher order mode suffer mostly from relatively narrow impedance bandwidth. However, no research focusing on the increase of the bandwidth of these antennas has been documented so far. Another unexplored research directions are the generation of the circular polarization in this radiators by techniques known from conventional DRA design and the utilization of the high gain DRA elements in the array antennas. Studies of mutual coupling between antenna elements operating with higher order modes are lacking; such studies are important since they must precede the effective DRA antenna array design [8],[29].

III. Gain Enhancement of the Single Element DRA

A cylindrically shaped resonator, operating with the low-order hybrid electromagnetic mode $HEM_{11\delta}$ placed above a sufficiently large ground plane, is probably the most frequently used DRA configuration [4]. This mode generates a broadside radiation pattern with linear polarization and a gain of about 5 dBi.

Several approaches have been suggested to increase the gain of the DRAs. Arraying of single element DRAs [8] is probably the most versatile method in which the gain value can be directly controlled by the number of elements in the array. Nevertheless, increased size, complexity and costs of the resultant antenna are the main disadvantages.

Altering a single element DRA can be used in cases, where medium gains up to around 10 dBi are sufficient. In general, two tactics to increase the gain of the single element DRA exist. First, additional structures can be placed in the vicinity of the resonator operating in the low order mode thus forming a more geometrically complex structure. In this report, larger attention is paid to the second approach, where the gain boost is provided by the introduction of the higher order modes into the simple shape dielectric resonator.

The gain enhancement of a single resonator DRA is a relatively new approach, the first research papers dealing with this topic were published during the year 2004. The usefulness of a single element directive DRA can be illustrated fairly easily: if the gain of the single element DRA could be increased from the 5 dBi (gain of the fundamental mode) to about 11 dBi - that would mean that the same gain could be obtained with a single element DRA as by an array of 4 DRA elements operating in the fundamental mode.

A. Additional structures in the DRA surrounding

This approach utilizes the DRA operating in the low-order mode (mostly the $HEM_{11\delta}$ mode of the cylindrical resonator) and its radiation characteristics are changed by introducing additional structures/elements in its close vicinity.

During experimentations with the stacked DRA structures [13-14] certain possibility to alter the antenna gain was discovered. First systematic design of a directive DRA based on this technique was demonstrated in [30] where the DRA operating in the low-order mode was extended by additional resonators to form a structure resembling a Yagi-Uda antenna. The achieved gains were 7.8 dBi and 8.7 dBi corresponding to the two (i.e. one active element and one director) and three element structure (one active element and two directors), respectively. Nevertheless, the antennas were not fabricated and all the results corresponded to the simulations. Nasimuddin and Eselle used a shape modification of the DRA ground plane [31]; the ground plane was altered into the shape of a surface mounted short horn increasing the gain of the DRA to 9.8 dBi at 5.95 GHz with 10 dB return loss bandwidth of 3.2%. Mushroom-like electromagnetic bandgap (EBG) structure with a circular symmetry was used in [39] to enhance the gain of a cylindrical DRA to about 8.3 dBi. Particularly, the rectangular hybrid DRA antenna based on a superstrate described in [32] provided peak gain of 14.44 dBi and gain above 11 dBi in the complete ISM band at 60 GHz.

It should be mentioned that a higher complexity and an increased size are the common disadvantages of the directive DRA antennas based on this approach.

B. Higher order modes in single element DRA

An alternative to the above described approach is to excite a suitable higher order mode in the single element DRA. Some of the modes documented over the last years exhibit relatively high radiation efficiencies with the gain up to about 10 dBi. The obvious advantages are the more compact dimensions and considerably reduced complexity of the final antenna; on the other hand, narrower impedance bandwidth is the main shortcoming.

The second strategy utilizes higher-order radiating modes in a single dielectric resonator. This approach has already been adopted in both rectangular and cylindrical DRAs. Petosa and Thirakoune in [33] and [34] showed that a DRA based on higher-order $TE_{\delta 13}$ and $TE_{\delta 15}$ modes in a rectangular resonator can achieve gains of 8.2 dBi and10.2 dBi, respectively. The structure operating in $TE_{\delta 15}$ mode [35] required a maximum resonator dimension of about $1.1\lambda_0$ when built from dielectric material with relative permittivity ε_r = 10, where λ_0 is the free space wavelength. Guha et al. [36-38] managed to excite higher-order $HEM_{12\delta}$ mode in a cylindrical resonator by introducing an air-filled cavity in the ground plane below the resonator. This way, peak gain of about 10 dBi was achieved but only in a relatively narrow impedance bandwidth.

Despite the reduced impedance bandwidth, the DRA antennas based on the second approach seem to be suitable for array applications. However, no DRA array design based on these structures has been proposed in the open literature so far.

IV. Conclusion

Excitation of the higher-order modes inside conventional, well-known dielectric resonator antennas can provide improvement of the antenna performance in different aspects, e.g. broader bandwidth, increased gain etc. The main focus of this article was to highlight the gain enhancement of a single element DRA utilizing the HO modes and its potential advantages over other methods.

Acknowledgment

The presented research was supported by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483.

References

[1] Balanis, C.A., Advanced engineering electromagnetics 2e, Wiley, 2012.

[2] Kajfez, D.; Guillon, P., Dielectric Resonators, Norwood, MA, Artech House, Inc., 1986, 547 p.

[3] Fiedziuszko, S.J.; Holme, S., "Dielectric resonators raise your high-Q," Microwave Magazine, IEEE , vol.2, no.3, pp.50,60, Sep 2001

[4] Long, S.A.; McAllister, M.; Liang Shen, "The resonant cylindrical dielectric cavity antenna," Antennas and Propagation, IEEE Transactions on , vol.31, no.3, pp.406,412, May 1983

[5] Petosa, A., Dielectric Resonator Antennas Handbook, Artech House Antennas and Propagation Library, 2007.

[6] Volakis, J., L. Antenna Engineering Handbook. McGraw-Hill, 2007.

[7] Petosa, A.; Ittipiboon, A., "Dielectric Resonator Antennas: A Historical Review and the Current State of the Art," Antennas and Propagation Magazine, IEEE , vol.52, no.5, pp.91,116, Oct. 2010

[8] Luk, K. M., Leung, K. W. Dielectric Resonator Antennas. Research Studies Press Ltd. 2003.

[9] Kumar Mongia, R.; Ittipiboon, A., "Theoretical and experimental investigations on rectangular dielectric resonator antennas," Antennas and Propagation, IEEE Transactions on , vol.45, no.9, pp.1348,1356, Sep 1997

[10] McAllister, M.W.; Long, S.A., "Resonant hemispherical dielectric antenna," Electronics Letters , vol.20, no.16, pp.657,659, August 2 1984

[11] Kwok-Wa Leung; Luk, Kwai-Man; Lai, K.Y.A.; Deyun Lin, "Theory and experiment of an aperture-coupled hemispherical dielectric resonator antenna," Antennas and Propagation, IEEE Transactions on , vol.43, no.11, pp.1192,1198, Nov 1995

[12] Thamae, L.Z.; Zhipeng Wu, "Broadband Bowtie Dielectric Resonator Antenna," Antennas and Propagation, IEEE Transactions on , vol.58, no.11, pp.3707,3710, Nov. 2010

[13] Qinjiang Rao; Denidni, T.A.; Sebak, A.R., "Broadband compact stacked T-shaped DRA with equilateral-triangle cross sections," Microwave and Wireless Components Letters, IEEE , vol.16, no.1, pp.7,9, Jan. 2006

[14] Chair, R.; Kishk, A.A.; Lee, K.F.; Smith, C.E., "Wideband flipped staired pyramid dielectric resonator antennas," Electronics Letters , vol.40, no.10, pp.581,582, 13 May 2004

[15] Kishk, A.A.; Glisson, A.W., "Effect of air gap on cylindrical dielectric resonator antenna operating in TM_{01} mode," Electronics Letters , vol.30, no.2, pp.97,98, 20 Jan 1994

[16] Yarga, S.; Sertel, K.; Volakis, J.L., "A Directive Resonator Antenna Using Degenerate Band Edge Crystals," Antennas and Propagation, IEEE Transactions on , vol.57, no.3, pp.799,803, March 2009

[17] St. Martin, J.T.H.; Antar, Y.M.M.; Kishk, A.A.; Ittipiboon, A.; Cuhaci, M., "Dielectric resonator antenna using aperture coupling," Electronics Letters , vol.26, no.24, pp.2015,2016, 22 Nov. 1990

[18] Zhang Cheng Hao; Wei Hong; AnDing Chen; Chen, Jixin; Ke Wu, "SIW fed dielectric resonator antennas (SIW-DRA)," Microwave Symposium Digest, 2006. IEEE MTT-S International , vol., no., pp.202,205, 11-16 June 2006

[19] Balanis, C., A. Antenna theory,analysis and design, 3/E. Wiley-Interscience,2005.

[20] Yong Mei Pan; Kwok Wa Leung; Kai Lu, "Omnidirectional Linearly and Circularly Polarized Rectangular Dielectric Resonator Antennas," Antennas and Propagation, IEEE Transactions on , vol.60, no.2, pp.751,759, Feb. 2012

[21] Avadanei, O.G.; Banciu, M.G.; Nedelcu, L., "Higher-Order Modes in High-Permittivity Cylindrical Dielectric Resonator Antenna Excited by an Off-Centered Rectangular Slot," Antennas and Wireless Propagation Letters, IEEE, vol.13, no., pp.1585,1588, 2014

[22] Avadanei, O. G.; Banciu, M. G.; Nicolaescu, I.; Nedelcu, L., "Superior Modes in High Permittivity Cylindrical Dielectric Resonator Antenna Excited by a Central Rectangular Slot," Antennas and Propagation, IEEE Transactions on, vol.60, no.11, pp.5032,5038, Nov. 2012

[23] Bin Li; Kwok Wa Leung, "Strip-fed rectangular dielectric resonator antennas with/without a parasitic patch," Antennas and Propagation, IEEE Transactions on , vol.53, no.7, pp.2200,2207, July 2005

[24] Yong Mei Pan; Shao Yong Zheng; Bin Jie Hu, "Design of Dual-Band Omnidirectional Cylindrical Dielectric Resonator Antenna," Antennas and Wireless Propagation Letters, IEEE , vol.13, no., pp.710,713, 2014

[25] FANG XIAOSHENG , Designs of Dualband and Wideband Dielectric Resonator Antennas using Higher-order Modes – dizertace

[26] Xiao Sheng Fang; Chi Kin Chow; Kwok Wa Leung; Eng Hock Lim, "New Single-/Dual-Mode Design Formulas of the Rectangular Dielectric Resonator Antenna Using Covariance Matrix Adaptation Evolutionary Strategy," Antennas and Wireless Propagation Letters, IEEE , vol.10, no., pp.734,737, 2011

[27] Sun, Y.X.; Leung, K.W., "Dual-Band and Wideband Dual-Polarized Cylindrical Dielectric Resonator Antennas," Antennas and Wireless Propagation Letters, IEEE , vol.12, no., pp.384,387, 2013

[28] Hady, L.K.; Kishk, A.A.; Kajfez, D., "Dual-Band Compact DRA With Circular and Monopole-Like Linear Polarizations as a Concept for GPS and WLAN Applications," Antennas and Propagation, IEEE Transactions on , vol.57, no.9, pp.2591,2598, Sept. 2009

[29] Yong-Xin Guo; Luk, Kwai-Man; Kwok-Wa Leung, "Mutual coupling between millimeter-wave dielectric-resonator antennas," Microwave Theory and Techniques, IEEE Transactions on , vol.47, no.11, pp.2164,2166, Nov 1999

[30] Kishk, Ahmed A., "Directive Yagi–Uda dielectric resonator antennas," Microwave and Optical Technology Letter , vol.44, no.5, pp.451,453, 2005

[31] Nasimuddin; Esselle, K.P., "Antennas with dielectric resonators and surface mounted short horns for high gain and large bandwidth,"

Microwaves, Antennas & Propagation, IET , vol.1, no.3, pp.723,728, June 2007

[32] Denidni, T.A.; Coulibaly, Y.; Boutayeb, H., "Hybrid Dielectric Resonator Antenna With Circular Mushroom-Like Structure for Gain Improvement," Antennas and Propagation, IEEE Transactions on , vol.57, no.4, pp.1043,1049, April 2009

[33] Coulibaly, Y.; Nedil, M.; Ben Mabrouk, I.; Talbi, L.; Denidni, T.A., "High gain rectangular dielectric resonator for broadband millimeter-waves underground communications," Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on , vol., no., pp.001088,001091, 8-11 May 2011

[34] Petosa, A.; Thirakoune, S.; Ittipiboon, A., "Higher-order modes in rectangular DRAs for gain enhancement," Antenna Technology and Applied Electromagnetics and the Canadian Radio Science Meeting, 2009. ANTEM/URSI 2009. 13th International Symposium on , vol., no., pp.1,4, 15-18 Feb. 2009

[35] Petosa, A.; Thirakoune, S., "Rectangular Dielectric Resonator Antennas With Enhanced Gain," Antennas and Propagation, IEEE Transactions on , vol.59, no.4, pp.1385,1389, April 2011

[36] Guha, D.; Banerjee, A.; Kumar, C.; Antar, Y.M.M., "Higher Order Mode Excitation for High-Gain Broadside Radiation From Cylindrical Dielectric Resonator Antennas," Antennas and Propagation, IEEE Transactions on , vol.60, no.1, pp.71,77, Jan. 2012

[37] Guha, D.; Banerjee, A.; Kumar, C.; Antar, Y.M.M., "New Technique to Excite Higher-Order Radiating Mode in a Cylindrical Dielectric Resonator Antenna," Antennas and Wireless Propagation Letters, IEEE , vol.13, no., pp.15,18, 2014

[38] Guha, D.; Banerjee, A.; Kumar, C.; Antar, Y.M.M.; Sebastian, M.T., "Design Guidelines for the Cylindrical Dielectric Resonator Antenna Using the Recently Proposed HEM12d Mode [Antenna Designer's Notebook]," Antennas and Propagation Magazine, IEEE , vol.56, no.4, pp.148,158, Aug. 2014

From 50 MHz to 4 GHz:
Powerful oscilloscopes from the T&M expert.

ROHDE&SCHWARZ

Digital Predistortion - from MATLAB Simulation to Fixed Point Implementation in USRP

Martin Pospisil, Roman Marsalek
Department of Radio electronics
Brno University of Technology
Technicka 12, Brno, Czech Republic
Email: xpospi29@stud.feec.vutbr.cz, marsaler@feec.vutbr.cz

I. Introduction

High efficiency of radio transmitters is required for current and future generation of wireless communication systems. On the contrary, the use of modulations with non-constant envelope reflects in the requirements on the transmitter (and thus mainly power amplifier) linearity. A huge family of methods for the compensation of power amplifier (PA) nonlinearities exists, such as feed-forward techniques [1], Cartesian feedback [2] and (usually Digital and adaptive) PreDistortion (DPD) [3]. The Look-Up-Table predistorters are often used for narrow-band signals because of their low complexity.

The basic principle of adaptive digital predistortion for linearization of power amplifiers is shown in Figure 1. Here we consider the baseband predistortion, where a transmitted baseband signal is upconverted (in digital or analog form) to desired carrier frequency. A part of the transmitted signal is fed back to the baseband through a directional coupler (and attenuator) to have a reference signal for a predistorter adaptation.

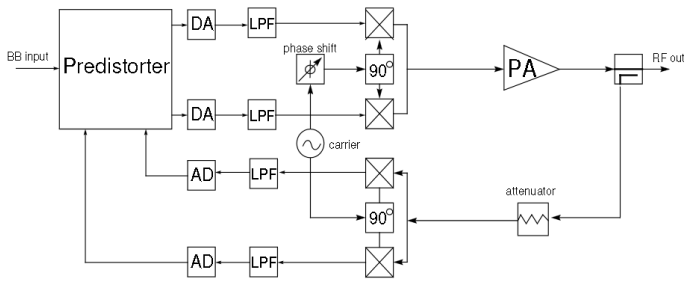


Fig. 1. Digital predistortion principle.

II. Look-Up-Table based DPD algorithm

We considered a complex gain LUT predistorter similarly as in [5], [3]. The predistorter output z_p at time instant n is

$$z_p(n) = z(n) f_{pre}(i), \tag{1}$$

where $f_{pre}(i)$ is a content of LUT entry i representing the inverse of the PA characteristics and $z(n)$ is a sample of modulated baseband signal. Note that in order to compensate for both amplitude and phase distortions, the LUT content is complex. The LUT can be addressed using either the instantaneous signal module or its instantaneous power. In order to get better linearization performance, the linear interpolation of the LUT entries is often done.

The adaptation of the predistorter has been done using an indirect-learning Recursive Least Squares (RLS) method in accordance with the block schematic in Fig. 2, [5]. In contrast to this sample-by-sample approach, we have slightly modified the algorithm to use it in the off-line block-by-block system. First the blocks of N data samples from the PA input and output are sampled, delay compensated and the predistorter LUT is computed from these blocks. Then the predistorter is applied to a new block of signal. In such a case we do not use any forgetting factor and the modulator output is used for the adaptation instead of the predistoter (that is not yet used in the adaptation phase) output. This leads to the adaptation formula:

$$f_{pre}(i) = \frac{\sum_{l=1}^{N} z(l) z_a(l)}{\sum_{l=1}^{N} |z_a(l)|^2}, \tag{2}$$

where z_a is the PA output (divided by the desired gain of the system after predistortion) and z the digital modulator output (i.e. PA input). You can notice that such a method is very simple and thus suitable for implementation.

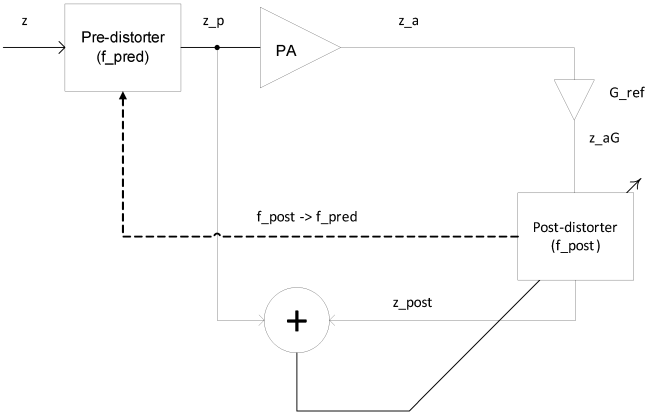


Fig. 2. Indirect learning architecture.

III. Software defined radio

A software defined radio (SDR) is defined [4] as *a radio in which the receive digitization is performed at some stage downstream from the antenna, typically after wideband filtering, low noise amplification, and down conversion to a lower frequency in subsequent stages with a reverse process occurring for the transmit digitization. Digital signal processing in flexible and reconfigurable functional blocks defines the characteristics of the radio.*

TABLE I. PARAMETERS OF SELECTED SDR DEVICES

Device type	FPGA	Bus	Bandwidth
USRP	Altera	USB 2.0	8 MHz
USRP2	Spartan 3A	Gig.ETH	25 MHz
USRP N210	Spartan 3A	Gig.ETH	25MHz
USRP E110	Spartan 3A+OMAP3	Gig.ETH	8MHz
USRP X310	Kintex 7	1/10Gig.ETH	120MHz
BladeRF x40	Cyclone 4E	USB 3.0	28 MHz
BladeRF x115	Cyclone 4E+ARM9	USB 3.0	28 MHz

Various software defined radios are currently available on the market. Table I provides the parameters of selected devices of two probably most-widespread families in the research community - Universal Software Radio Peripheral (USRP) from Ettus Research (now National instruments) and Blade RF designed and fabricated by NUAND company. Except USRP1, all devices of USRP family are MIMO capable and can be used with various front-ends such as LFTX/RX (covering frequency range from DC to 30 MHz), WBX (50-2200 MHz), TVRX2 (50-860 MHz) or the most recent CBX-120 (1200-6000 MHz). The BladeRF radios are MIMO expandable to 2x2 (x40) or 4x4 (x115) and work with a fixed front-end based on a LMS6002 chip for 300-3800 MHz frequency range (with extension to 60kHz-300 MHz) coverage.

From among the above mentioned SDRs, we have used USRP N200, together with WBX front-end for the implementation of DPD.

IV. DPD IMPLEMENTATION IN USRP

In our design, the direct path of DPD along with necessary components as memories are implemented in FPGA, while the adaptation part of the DPD algorithm is implemented in the in-built software processor ZPU. ZPU is a 32-bit stack-based soft processor core, implemented in FPGA. It is one of the smallest available 32-bit processors. Toolchain for ZPU utilizes GCC as a compiler. ZPU is not widely used processor and its toolchain offers only limited optimization capability. ZPU is very simple processor, suitable for implementation to low-density programmable devices as it requires less than 1000 LUT's, in minimal configuration even around 500 LUT's.

ZPU application controls high level functions of USRP, such as parsing payload from control packets. Some smaller tasks are implemented as interrupt routines, however bulk of code runs in main loop. We tested the simplest solution. Integer based code from MATLAB implementation was ported to C language, and called by adaptation routine. The computation went too slow, which delayed the main loop and caused USRP control timeouts.

For the communication of the modules in the general USRP N2x0 design, an open source Wishbone bus is used. In this particular case the bus has 32-bit data width and 32-bit address space. There is one wishbone master in the USRP design, which is the ZPU soft processor. The bus is configured to have 16 slaves, but some of the slave positions are not used. On originally unused address 4, the slave for DPD is implemented with the ZPU address space starting at 0x5800.

A. Integration of DPD into USRP

Overall schematic of DPD integration inside the USRP N200 is shown on Fig. 3. The direct path of DPD is implemented as the complex multiplier (IP core from Xilinx) addressed by the squared signal magnitude. The squared magnitude has been chosen in order to avoid the necessity for square root operation in magnitude calculation. The operation of direct path can be bypassed by a command from the PC.

The data can be transmitted either from the block RAM (TXRAM) of 4096 samples (32bit) or directly from the PC. Inversely, the received data can be either send directly to PC or stored to the block RAM RXRAM of the same length and width. Both TXRAM and RXRAM are implemented as dual-port memories in the FPGA with B port connected to ZPU (wishbone bus slave). A binary counter CNT is used to allow replay of data in TXRAM and storage in RXRAM. Another RAM (CoeffRAM) is used to store 32 complex LUT coefficients, each using 32-bits (16 bit for I, 16 bit for Q part, both in Q15 format). The CoeffRAM can either be filled by the adaptation algorithm, or alternatively directly from the PC, through the ZPU.

B. Operation of the DPD

Operation of the implemented DPD can be described as follows:

- Data to transmit are created (e.g. in MATLAB), stored to binary data file and uploaded to TXRAM.
- DPD is bypassed. Received data are stored in RXRAM
- The adaptation in ZPU is started. Note that ZPU needs to have a knowledge about the time delay between the samples in TXRAM and RXRAM caused by DUC/DDC chain.
- LUT coeficients are uploaded to CoeffRAM and DPD bypass is interrupted. Now the data are predistorted in the USRP!

C. Matlab fixed point implementation

Prior to implementation in programmable device, it is possible to simulate the algorithm performance in fixed point precision using MATLAB. Result of such simulation, in comparison with the standard MATLAB floating point simulation of predistortion adaptation is show in Fig. 4 for an example of the Long Term Evolution (LTE) adaptation signal. Here, the amplitude and phase corrections, i.e., the LUT entries are shown. You can notice only slight difference between the fixed and floating point results.

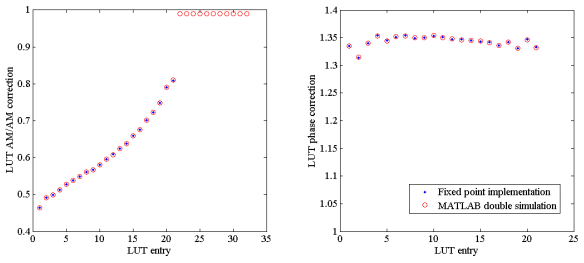


Fig. 4. Comparison of floating and fixed point simulation in MATLAB

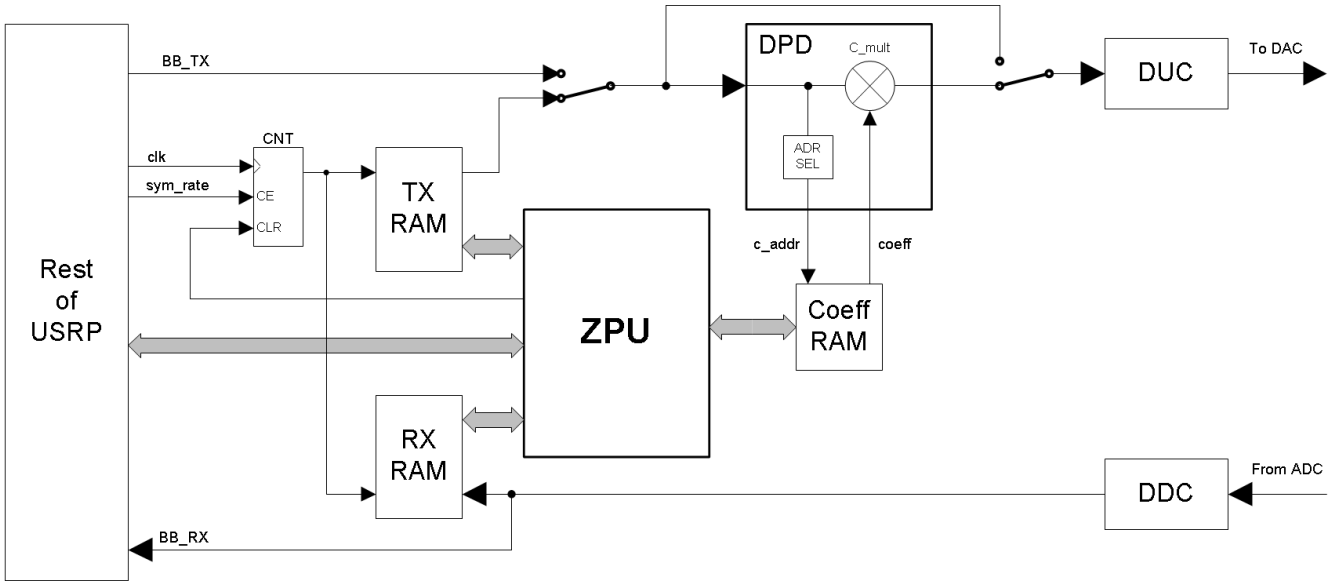


Fig. 3. Block diagram of USRP-DPD integration

TABLE II. RESOURCES FOR DPD AND ITS ADAPTATION IN FPGA

	Slices	Slice F-F	DSP48	Block RAM
direct path	70	33	5	17
RLS in FPGA	1154	1682	29	22
USRP free/tot.	1548/16640	17796/33280	56/84	30/84

D. Notes on adaptation implementation in USRP

For each signal sample, an equation 2 has to be performed. From that, the most time consuming operation on the ZPU is the division. We modified the equation and scaling of operands, to be able to use divisor values of power of two. These divisions were then implemented as arithmetic bit shifts. Even though bit shifting isnt equivalent to division, we found that the effect of negative rounding down is negligible in this application. That modification reduced computation time per signal sample significantly. Then we were able to compute 4096 signal samples per one main function loop, without disturbing rest of USRP functionality. At the end of the iterative part of algorithm, values of DPD coefficients are calculated. This uses two divisions per coefficient (real and imaginary part), and these need to be implemented as normal division.

ZPU contains only 1-byte instructions. After implementation of the DPD adaptation, the ZPU firmware contains 14180 bytes (i.e. instructions), while the original firmware we used as a starting point contains 12488 instructions. All the values are after optimizations of the code by the compiler. Note that no additional logic is necessary for DPD adaptation part and the only FPGA resources necessary are these for the direct path and RXRAM/TXRAM with counters. In order to highlight the advantage of adaptation implementation in ZPU, the estimation of resources necessary for RLS algorithm implementation purely in the FPGA logic are summarized in table II. This table also contains the total resources of FPGA in the USRP N200 device, the free resources and the resources necessary for the direct path of DPD.

The results of DPD applied to LTE signals can be observed by means of transmitted signal spectras shown on Fig. 5. You can see that DPD improves the adjacent channel emissions by approximately 10dB.

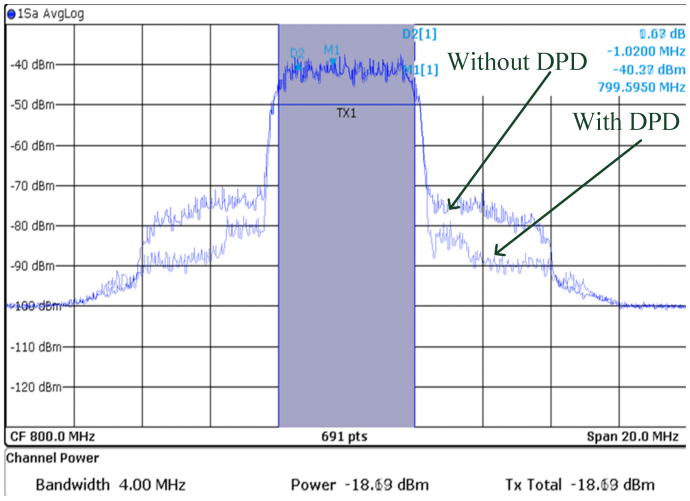


Fig. 5. Spectras of PA output without and with described DPD

V. CONCLUSIONS

In this paper we presented the implementation of digital predistortion algorithm in the USRP software definer radio. The direct path of the predistorter has been implemented in the FPGA logic and signal processing blocks, while the software processor ZPU has been used for the adaptation algorithm implementation. This is very advantageous as ZPU is already present in the original USRP design and such implementation saves FPGA logic for another applications. Moreover the chosen model of USRP (USRP N200) is equipped with smaller Spartan 3 FPGA and the free resources inside FPGA are very limited.

ACKNOWLEDGEMENT

The presented research has been supported by the BUT internal project FEKT-S-14-2177 (PEKOS). The experimental part was performed in the laboratories of the SIX research center, reg. no. CZ.1.05/2.1.00/03.0072 built with the support of project CZ.1.07/2.3.00/20.0007 WICOMT.

REFERENCES

[1] BURGLECHNER, S., GHADAM, A.S.H., SPRINGER, A., VALKAMA, M. and HUEBER, G.. DSP oriented implementation of a feedforward power amplifier linearizer. *2009 IEEE International Symposium on Circuits and Systems*, p. 409-425.

[2] JOHANSSON, M. and SUNDSTROM, L., Linearisation of RF multi-carrier amplifiers using Cartesian feedback. *Electronics Letters*. 1994, vol. 30, issue 14, p. 33-42.

[3] MARSALEK, R.; POSPISIL, M., Evaluation of digital predistortion using the USRP N200 software defined radio transceiver, In NORCHIP 2014 Conference proceedings, p.1,4, 27-28 Oct. 2014.

[4] TUTLEBEE, W. *Software defined radio: enabling technologies*. New York: John Wiley, 2002, ISBN 04-708-4318-7.

[5] Baudoin, G., Marsalek, R., Jardin, P., A new approach for LUT- based digital predistorters adaptation, In *proceedings of Confeence Electronic Devices and Systems EDS 2003*, available at http://www.esiee.fr/ baudoing/pdf/marsalek2003-eds.pdf.

Quality of Solder Joints in Automotive Electronics after Accelerated Aging Tests

Ján Schneider, Ján Gamec, Daniel Novák
Dept. of Electronics and Multimedia Communications
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Košice, Slovak Republic
Email: jan.schneider@tuke.sk, jan.gamec@tuke.sk

Alena Pietriková
Dept. of Technologies in Electronics
Faculty of Electrical Engineering and Informatics
Technical University of Košice,
Košice, Slovak Republic
Email: alena.pietrikova@tuke.sk

***Abstract*—This paper deals with testing quality of solder joints. An influence of repeated step change of temperature affecting to the solder joint is described here. It is produced in a professional manner and commonly used in automotive electronics as a backlight dashboard. Quality and adhesion to the substrate of solder joint is tested by shear test.**

It will be shown that the shear test can be used as a proxy to the drop test to evaluate quality of the components that are assembled not only in automotive electronics.

Keywords—Temperature cycling test; Shear test; Quality of solder joint.

I. Introduction

Solder joint reliability, which is directly related to the shape of solder joint, is vital to the SMD (surface mount devices) product. Quality of solder joint is very closely watched area of electronics devices. Solder joint quality under normal conditions does not have any significant reason to significantly weaken. Quality of solder joint depends on the processes of assembly design, material, techniques, equipment's, inspecting and quality control and so on [1].

In this paper, a quality of solder joints used in automotive dashboards after accelerated aging tests is presented. Next sections present tested sample, accelerated aging test, shear test and measured results.

II. Tested Sample

The main objective was to subject the samples with solder joints used in automotive industry to accelerated aging test. After accelerated aging test, the quality of solder joint was tested by shear test.

In this case, an automotive dashboard backlight was used as a testing sample. Production samples were carried out in the company engaged exclusively in assembly and installation of components on the PCB for the automotive industry. The sample is produced on fully automated manufacturing line. Final testing sample is depicted in Fig. 1.

As the substrate test samples was used FR4 dielectric material. It is a composite material composed of woven fiberglass cloth with an epoxy resin binder. Hot Air Leveling (HAL) was used as surface finish of tested samples. It is a substrate plating layer of tin. To create solder joints on the PCB was used solder paste type SAC305 (96.5% Sn / 3.0% Ag / 0.5% Cu) [2]. Parts used for our samples were type 3528 SMD LED with housing dimensions 3.5x2.8 mm. Pitch adjustment has been based on tin. Dimensions pad for one side of LED are 2.6x1.8 mm.

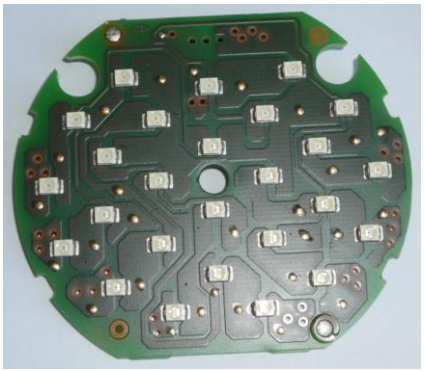


Fig. 1 Test sample.

III. Accelerated Aging Test

For our case was used one type of accelerated aging test. Changes of mechanical properties of solder joint are the indicators of changes in the material internal structure caused by the influence of a load. In Fig. 2 is shown temperature profile of temperature cycling test. Accelerated aging test was provided by thermal shock chamber.

A. Temperature cycling test

Standard JEDEC JESD22-A104-B and soak mode 3 (soak time 10 minutes) was used for temperature cycling tests. Quick temperature changes, from -55 °C to +125 °C were applied on whole PCB system with LED diodes in thermal shock chamber. These devices were exposed temperature, -55 °C and +125 °C, for 10 minutes. The force required to breach soldered joint was measured after 20, 50, 100 and 400 cycles [3][4].

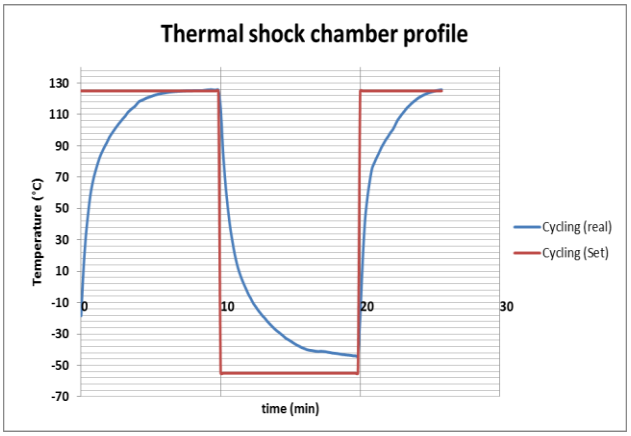


Fig. 2 Temperature profile of temperature cycling test.

IV. Shear Test

In our case, mechanical properties of solder joint was realized by shear test. Although shear test technology is more demanding to define dimensions of solder joint, but it appears to be the simplest testing technology. The resulting mechanical properties of the shear diagram are thorough for the type and size of components used in this case.

Shear test provides resistance to force that can act on the soldered joint already during the production equipment, when handling the device, transportation and even the use of the finished device. Because of the benefits outlined above, we chose just this test. The principle of the chosen shear test is shown in Fig. 3.

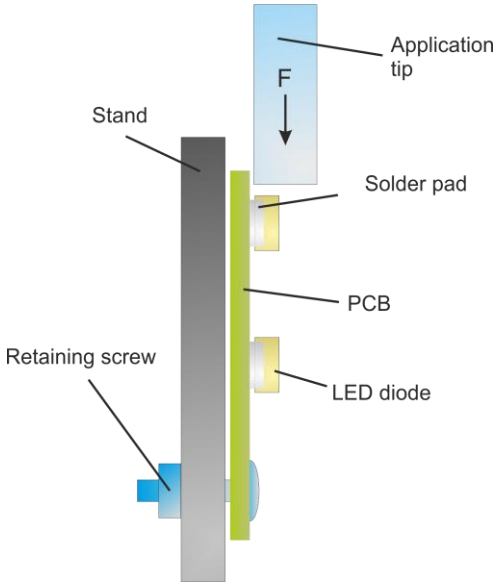


Fig. 3 Principle of shear test.

Mechanical shear test was performed at a constant rate of displacement tip 2 mm/min (0.3333 mm/s). Testing sample (PCB with LED diodes) was fixed in a special holding device tailor-made for our test only. Rebound tip from PCB has been set at a maximum of ¼ of the amount of component (LED diode). From the measured test our sophisticated software Wintest Analysis [5] provides a graphical output value.

V. Measured Results

Measurements were done by shear test device Testometric M250-2.5CT [5]. Based on the value of output we get diagram mechanical properties of shear test see in Fig. 4.

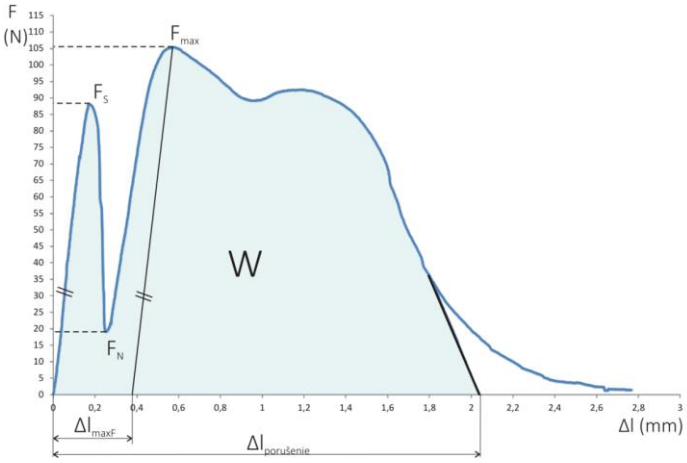


Fig. 4 Diagram mechanical properties of shear test.

Diagram provides us following parameters:

- Fs [N] - The force at which the detachment parts,
- Fmax [N] - The maximum solder joint strength, are degraded joint,
- W [J] - Work needed to break the solder joint.

Force F begins to increase when tip starts act on component (LED diode). Dependence force to change of the deformation to the point Fs represents elastic deformation of the material. From the point Fs there is plastic deformation. Plastic deformation is irreversible. To total avulsion of component there is a change on curve between points Fs and Fn. At force Fn component does not hold any solder. From this point the tip act only on remaining pads of component and solder to the maximum force Fmax there is complete degradation of the joint. The complete destruction of the adhesion between PCB copper pads and solder with the contact of component (LED) will not happen, because this joint is very durable. When the tip act on solder (without component) this causes that the tip degrades and damages the solder - the tip cropping solder. Fig. **5** shows solder pads after shear test. During this operation, the force decreases until the moment until the tip does not pass the full path of soldering pads. Tip stop act on a solder when curve starts chamfer to x axis. Work needed to break the solder joint is acquired by interpolation down curve from point Fmax and calculating area under the curve. Area under the curve is calculated by trapezoidal method. This method is reaching the error of calculating only maximal 0.05 %. Values of work W and forces Fmax, Fs with percentage change are in Table I.

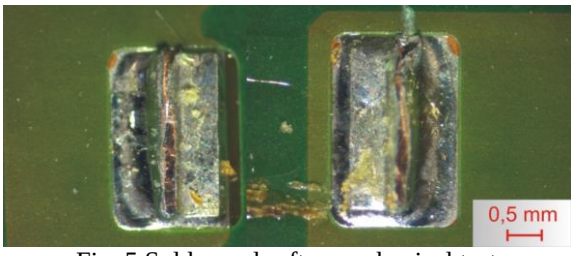


Fig. 5 Solder pads after mechanical test

Table I

Thermal cycles	F_S (N)	F_{max} (N)	W (J)	% change W (%)	% change Fmax (%)
0	87.88	114.8	173.63		
20	80.24	107.2	157.69	9.18	6.62
50	79.63	96.9	144.13	16.99	15.59
100	77.81	88.1	133.58	23.7	23.26
400	76.74	73.77	121.42	30.7	35.74

For a better interpretation we show the influence of thermal shock to the solder joints in this bar graph in Fig. 6.

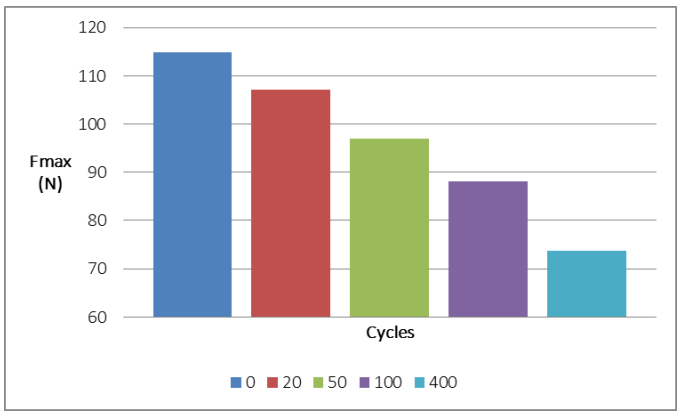


Fig. 6 Influence of thermal shock to the solder joints.

VI. Conclusion

Quality of solder joint used in automotive electronics was presented in this paper. Based on completed test applied on sample, the shear test has proven appropriate methodology for estimating the mechanical properties of solder joints. From shear test was determined work to breaking the solder joint (W), force at which the detachment parts (Fs) and maximum solder joint strength (Fmax). Effect of temperature cycling test seems to be most appropriate on solder joints in automotive electronics. The impact of these shocks is the closest to real conditions in practice. It is plausible in view of the technical evaluation. All named parameters (W, Fs, Fmax) after accelerated aging test were evaluated in Table . The Table . and Fig. 6 show the percentage change in the evaluated parameters raises increasing number of thermal shocks. The solder joint is strong enough come to a complete destruction of the adhesion between the substrate and the contact of LED diode.

Acknowledgment

This work was supported by Slovak Research and Development Agency under the contract No. APVV-0404-12.

References

[1] ZHOU Dejian,HUANG Chunyue, WU Zhaohua, et. “SMT solder joints quality assurance based on solder joints virtual evolving technology”, Computer Integrated Manufacturing SystemsSystems, 2006.

[2] ALPHA® OM-338-PT FINE FEATURE, PIN-TESTABLE LEAD-FREE SOLDER PASTE, Technical Bulletin, avalaible: http://alpha.alent.com/~/media/Files/CooksonElectronics/Products/Solder%20Paste/TB-TECHNICAL%20BULLETINS/OM-338%20SERIES/OM-338-PT/ALPHA%20OM-338-PT%20CNP%20TB%20SM893-21%20%20ENGLISH%2020141125.pdf , 2015.

[3] Rovensky, T.; Pietrikova, A.; Ruman, K.; Vehec, I., "Stability of LTCC substrates in high frequency area after accelerated aging tests," Electronics Technology (ISSE), Proceedings of the 2014 37th International Spring Seminar on , vol., no., pp.221,224, 7-11 May 2014.

[4] JEDEC STANDARD, High Temperature Optimum Life, JEDEC JES22-A108.

[5] Testometric M250-2.5CT, Data Sheet, avalaible: http://www.hartech.nl/sites/all/files/files/ct_m250_2-5kn.pdf, 2015

Combined System for Testing of AC Traction Motors

Lubos Struharnansky
Dpt. of Power Electrical Systems
Faculty of Electrical Engineering, University of Zilina
Zilina, Slovakia
lubos.struharnansky@fel.uniza.sk

Matej Pacha
Dpt. of Power Electrical Systems
Faculty of Electrical Engineering, University of Zilina
Zilina, Slovakia
matej.pacha@ieee.org

***Abstract*—The paper presents a design of a stand for testing of AC traction motors in combined system. Based on a set of requirements, parameters are designed and verified by simulation. The testing stand is introduced including future utilization.**

Keywords—induction machine; traction motor; combined system; converter;

I. INTRODUCTION

In recent years, electric traction drives became very popular in various types of transportation. Moreover, the segment of electric cars grows very quickly these days. Traction drive development requires a set of laboratory tests focused on functionality, efficiency, control algorithms, reliability and safety. Any high power traction drive test should be designed to handle the energy with minimum losses. If a dynamometer is available, a recuperation is a key feature. Otherwise a combined system should be designed.

II. DESIGN OF A STAND

A. Tested Traction Motors

Traction motors for testing are AC induction motors designed for public transportation (trams) and manufactured by SLOVRES Košice (Fig. 1). Parameters are listed in Table I. Two motors are connected together with a torque sensor between them, where one machine acts as a motor and the other acts as a generator (load).

TABLE I. PARAMETERS OF TRACTION MOTOR

Traction Motor ATR 200L-4	
Parameter	***Value***
Nominal Power	55 kW
Nominal Voltage	300 V AC
Nominal Currennt	150 A
Nominal Speed (@45 Hz)	1310 RPM
Frequency Range	0,1 – 140 Hz
Nominal Power Factor	cos φ = 0,76
Cooling System	Self

B. Power Converters

Power converters are 3phase full-bridge IGBT voltage converters for small traction systems or auxiliaries up to 72 kVA (modular type EVPU SN72-800, Fig. 1). External power supply 3x400 V AC or 3x690 V AC is connected to the DC bus via input rectifier. Auxiliary circuits provide switching commands (with pre-charging of the DC bus), overcurrent protection and a fan control. DC bus protection is made of a single IGBT driven resistive brake.

The converters introduces the whole set of necessary hardware and software protection systems. Debug and communication interfaces enable both low and high level software development and future connectivity with a computer.

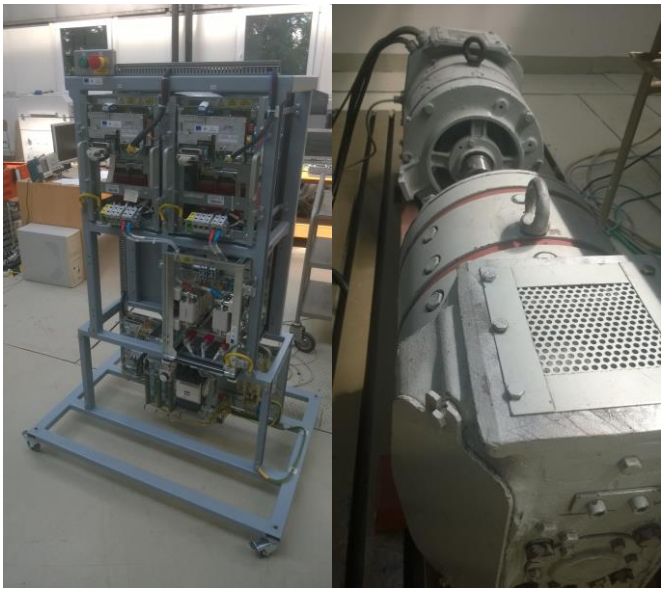


Fig. 1. Power converters (left) and traction motors (right)

C. Combined System

Combined system is defined in [1] as a unit of a power converter, a motor (motors), a control system, power cables and a cooling system. To minimize power losses, combined system introduces two units connected together. The system is usually supplied from an external power source. Power requirements on the external power source are limited to the power losses of the system and some dynamic power for

acceleration. Therefore, even high powered traction drives can be tested with much smaller power source available.

Our system is designed to handle two AC induction traction machines of 55 kW, where the input power is 59.2 kW in motoric mode. Thus, 4.2 kW per one machine is needed to be covered from the external power source. In addition, 10 kW are estimated for the acceleration. Overall power input is designed to 20 kW (32 A power slot, thus up to 22 kW).

Since the traction motors require relatively high reactive power (cos φ = 0,76), the complex power is slightly above the converters' maximum (79 kVA over 72 kVA). However, lowering a switching frequency the converters can be used in continuous mode up to 80 kVA.

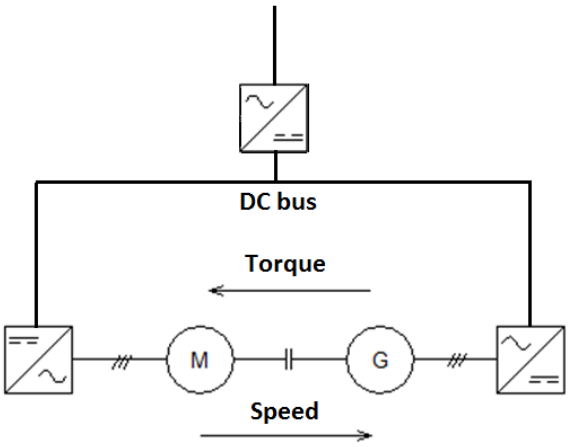


Fig. 2. Combined system

III. Test Description

The main standard [1] describes several tests (type tests, investigation tests and, in ČSN/STN national standard, a routine test). Type tests are very complex and include e.g. temperature-rise tests, torque characteristics, efficiency, power line cut-off, input harmonic distortion, etc. However, the basic test is a torque characteristics of a motor and torque characteristics of a traction drive.

A. *Torque characteristics of a motor*

Natural characteristics of a motor is defined for a fixed frequency and voltage, usually for a nominal frequency [2] or a set of frequencies. Nevertheless, traction motors are usually well utilized and maximal theoretical ratings are higher than allowed maximums. Therefore, the characteristics are usually limited to the nominal and short-time ratings [3].

In combined system, one converter-motor set acts as a motor and the other set acts as a generator. The control system of the motor drives the converter at constant voltage and frequency. The control system of the generator produces a load torque to the motor according to e.g. slope or step commands, usually in a vector control in a torque (current) loop.

Fig. 3 shows simulated load curves of a torque slope command test. The motor (blue curves) is slowing down as the load torque rises. The advantage of this method is a very small impact of a rotor inertia. Tests can be stopped when reached maximal ratings. Fig. 4 introduces the torque characteristics obtained from the simulation.

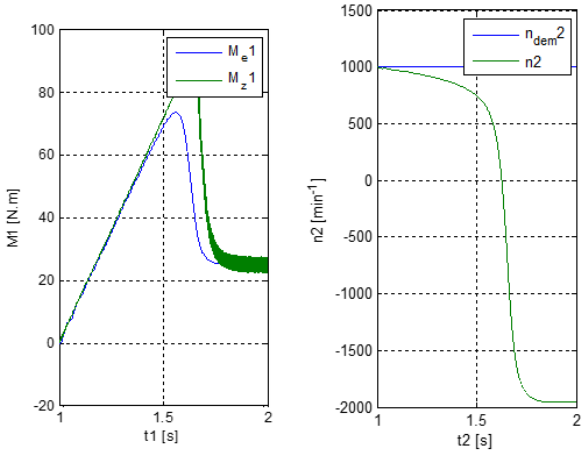


Fig. 3. Slope load curves of a combined system

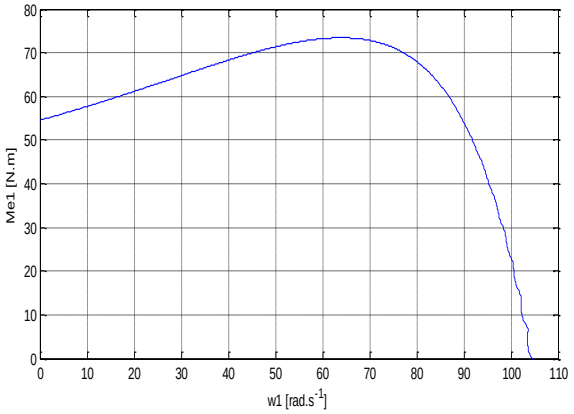


Fig. 4. Torque characteristics of the motor

B. *Torque characteristics of a drive*

Traction drive characteristics are given by design according to the application. In some cases, standard (or mass produced) traction motors are used instead of developing a new traction motor. In the case, traction drive characteristics can be set bellow the motor ratings. [3]

The traction characteristics are usually limited by adhesion limit, mechanical limits of the transmission (gearbox, axles), rated power and a torque limit of a motor. All these limits can be "floating limits", according to temperatures, adhesion conditions, etc. However, the traction drive should be tested to its absolute limits.

Fig. 5 shows traction drive characteristics and expected input voltage and current curves. There are other techniques of shaping the characteristics (e.g. voltage boost by different modulation index), but they must be designed in cooperation with a motor manufacturer.

When measuring the characteristics, the thermal conditions should be constant. Therefore, the torque characteristics should be measured after the temperature-rise test. The measurement should start at the lowest speed that can be measured correctly. The rest of the test should be done as fast as possible to avoid temperature changes. Minimal number of required points is marked in Fig. 5. [1]

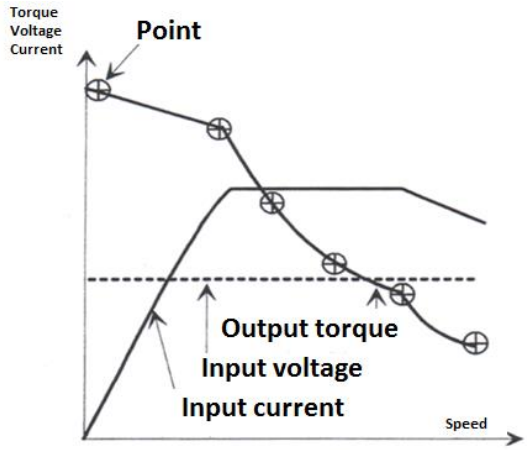


Fig. 5. Traction drive characteristics [1]

C. *Efficiency test*

Efficiency test is an optional type test. The efficiency of a combined system is given by a mechanical power measured on a shaft of a motor relative to the DC power on terminals of the converter [1]. The combined system stand designed has a set of terminals for a measurement of electrical signals. Mechanical data is collected by torque sensor mounted between the machines and the speed is measured by internal speed sensor built in the motors.

IV. Future Utilization of the Stand

The design of the testing stand is a part of a PhD thesis of the first author. The main advantage of the design is to have a small/middle sized traction drive testing environment with real components.

Traction converters SN72-800 are widely used in practice and provide a fully supported platform for control algorithms development. In addition, future connection with a computer will introduce a set of automated tests and evaluation software.

Need for a high efficiency traction drive is very actual these days in general. Therefore, high efficiency control techniques will be implemented and tested, for example “maximum torque per ampere” [4] or “loss minimization control” [5].

Another task is to develop sensor-less control techniques for traction purposes. Usually there is always a speed sensor for safety reasons [6], but in case of failure a backup information about state of the motor is crucial.

Additional purpose of the stand is dedicated to education, where students will run real tests on a real middle sized electric drive.

V. Conclusion

The paper presents a new design of a test stand for testing of AC traction drives. The concept introduces a combined system defined in IEC standard, including required tests.

Basic control schemes and requirements are introduced for a torque characteristics of a motor and of the whole drive.

Finally, future utilization of the stand is described and research directions are introduced. The main direction of high efficiency traction drives follows current trends in the world.

Acknowledgment

The paper has been supported by project KEGA No. 006ZU-4/2014. The test stand has been supported by project SFEU ITMS 26210120021 and EVPU, a.s. Nova Dubnica.

References

[1] EN 61377-1 “Railway applications – Rolling stock, Part 1: Combined testing of inverted-fed alternating current motors and their control system,” (IEC 61377-1:2006).

[2] EN 60349-2 “Rotating electrical machines for rail and road vehicles, Part 2: Electronic convertor-fed alternating current motors,” (IEC 60349-2:2010.

[3] Danzer, J., “Electric Traction,” ZCU Plzen, 2007.

[4] Hrkel, M.; Vittek, J.; Biel, Z., “Maximum torque per ampere control strategy of induction motor with iron losses,” ELEKTRO, 2012 , vol., no., pp.185,190, 21-22 May 2012.

[5] Famouri, P.; Cathey, J.J., “Loss minimization control of an induction motor drive,” Industry Applications Society Annual Meeting, 1989., Conference Record of the 1989 IEEE , vol., no., pp.226,231 vol.1, 1-5 Oct. 1989.

[6] “Technical Specification for Interoperability – LOC&PAS-TSI specific part,” European Railway Agency, [online], http://www.era.europa.eu/, 05/2015.

The Influence of Objects on Surface Waves - the Determination of the Reference Obstacle

Jan Vélim, Zbyněk Raida
Department of Radio Electronics, Brno University of Technology
Email: velim@phd.feec.vutbr.cz

***Abstract*—The paper aims to specify the methodology of how to evaluate independence of a wireless link on the obstacles in surrounding environment. The transmission coefficient between two antennas is observed while various obstacles are positioned into defined positions. Several different materials are used for the guiding of the electromagnetic waves.**

I. Introduction

Nowadays, wireless links are still in progress and the attention moves to the higher frequency bands. It is known that the frequency band around 60 GHz is suitable for the short distance communication. For example, a wireless link along the roof inside car is developed in this frequency band. The important issue is to find out how big an interference of passengers which can have an effect on this wireless connection.

There is no established methodology of how to confirm the independence of the wireless connection. Since the electrical structure of the investigated environment is large, the methodology is therefore primarily based on experimental investigation.

II. Surface waves

As mentioned above the propagation of electromagnetic waves along the car's roof is the object of interest. This environment is composed of a conductive metal sheet and dielectric materials which are placed below it. From the theory [1] is known that in such structures the energy is not radiated into all directions but it is guided along the surface of the material. Such type of wave is called a surface wave. Because this name is used for different types of waves, a special attention should be given to the terminology [2]. In accordance with this article, the wave which propagates along the roof is called Zenneck surface wave (Fig. 1).

So Zenneck surface wave propagates along the interface between two materials with different properties (ϵ - permittivity, μ - permeability, σ - conductivity). The figure also shows that the electromagnetic field is composed of three components (H_z, E_x, E_y). The field distribution in medium 1 is described by equations 1, 2, 3 [3]:

$$H_z = Ae^{j\omega t}e^{-\gamma x}e^{-uy}, \tag{1}$$

$$E_x = -A\frac{u}{j\omega\epsilon}e^{j\omega t}e^{-\gamma x}e^{-uy}, \tag{2}$$

$$E_y = A\frac{\gamma}{j\omega\epsilon}e^{j\omega t}e^{-\gamma x}e^{-uy}, \tag{3}$$

$$\Gamma = \alpha + j\beta,$$
$$u = a + jb,$$

where A is a constant of the amplitude, Γ is propagation coefficient along x-axis with attenuation α and phase change β. In y direction the field changes according to the propagation constant u with attenuation a and phase change b.

III. Experimental investigations

The wireless links are variously susceptible to objects in surrounding environment. They depend on the manner of the excitation of surface waves and on the properties of a medium through which the waves propagate. Real obstacles, such as a human head, have a variety of shapes, sizes and unknown electrical properties. It is not possible to compare the wireless links with one another by these real obstacles. It is necessary to establish a reference obstacle for the comparison.

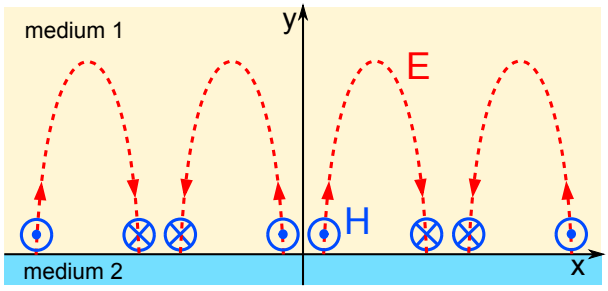


Fig. 1. The field distribution of the Zenneck surface wave.

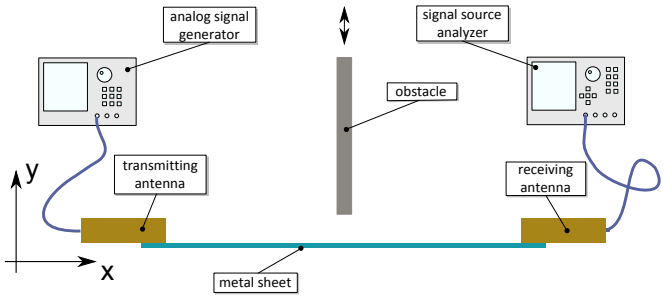


Fig. 2. Reference measuring workplace.

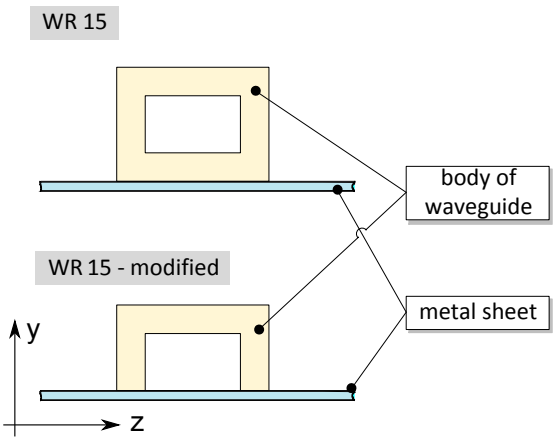


Fig. 3. Rectangular waveguide WR15 (top) and his modified version (below) used for excitation of the surface waves.

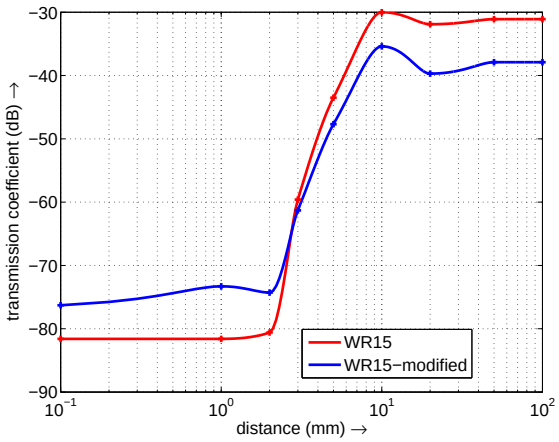


Fig. 4. Dependency of the transmission coefficient on the position of the obstacle for two different antennas.

The measurement workplace (Fig. 2) has been built for the determination of the reference obstacle.

Dependency of the transmission coefficient on the position of the obstacle is investigated by this measurement workplace. The obstacle moves in the direction of the y-axis and the transmission coefficient is measured between two antennas that are spaced 180 mm.

A. *Excitation of surface waves*

In the past the excitations of the surface waves were mainly solved by wire antennas, [4], [5]. As this research is made in the frequency band at 60 GHz, the waveguides are used for excitation of the surface waves. It turned out that in a real environment, it is not easy to excite a surface wave equal to the ideal case described by Fig. 1. Therefore several tests with two different types of waveguides (Fig. 3) should prove that the imperfections of excitation do not have a big impact on this investigation.

The rectangular waveguide WR15 which is placed on the metal sheet has maximum gain 14.46 dBi. Because of the waveguide's body has width 1.03 mm, the significant reflections appear. It means that such antenna has an undesirable sidelobe. Because the sidelobe could generate inaccurate results, the modified type of antenna is used for comparison. The modified WR15 has no bottom side of the body. It prevents creating the sidelobes. The smaller directivity is the disadvantage of this antenna therefore the gain reduces to 12.15 dBi.

In the Fig. 4 two curves are depicted. The curves present the dependencies of the transmission coefficient of the two different antennas on the position of the obstacle. Although on the start the obstacle lies on the metal sheet in the figure the initial value on the horizontal axis is 0.1 mm. It is because of a rough surface of the obstacle. Next the obstacle moves up until the distance is 100 mm from the metal sheet. In this position, the value of the transmission coefficient equals the value when the obstacle is not present. It is possible to notice the similarity of both curves. The subject of interest is especially the right half-plane of the chart. Only the difference between the curves is the level of the transmission coefficient. That means that no reflection from the sidelobe has an affect on this measuring. The next interesting point is that the maximum value of the transmission coefficient is for the distance 10 mm of the obstacle from the metal sheet. This applies to both antennas even if the antennas have different heights (WR15 - 3.94 mm, WR15-modified - 2.88 mm) and different directivities. It is possible to affirm from the results that no object should be located in the zone high up to 10 mm from the surface of the metal sheet.

B. *Obstacles*

The most important part of this research is to define a suitable reference obstacle. In the introduction it was said that the point of interest is to know the influence of passengers on an in-car wireless link. However, it is not a good solution to have a reference obstacle which is similar to human body because such materials do not have a long shelf life and the manufacturing of them is demanding. The two kinds of obstacles are chosen for the tests. The first material is the aluminium sheet 1 mm thick. The second one is a 50 mm thick anechoic rectangle with flat walls. This material is commonly used in anechoic chambers. The dimensions of the obstacles in y-z plane (with agreement with Fig. 2) have to present an infinite proportion for the radiation pattern of antennas.

In the first test, the antennas lie on the metal sheet again and the obstacles move up from the point zero. The zero point means that an obstacle lies on the metal sheet too. The results are shown in the Fig. 5. The interesting fact is the difference in the transmission coefficient at low altitudes of the obstacles. While in the case of the aluminium sheet the transmission coefficient grows immediately, in the case of the anechoic material the value of the transmission coefficient begins to grow up at a distance of 3 mm from the surface of the metal sheet. This is probably caused by the diffraction at the bottom edge of the aluminium sheet.

In the second test, the upholstery of the car is put between the antennas. Now the lowest position of the obstacles is the top surface of the upholstery. The upholstery is about 10 mm high. As we show in the Fig. 6, the smaller difference of the values of the transmission coefficient at a height of 0.1 mm and at a height of 100 mm is for aluminium sheet. It is again

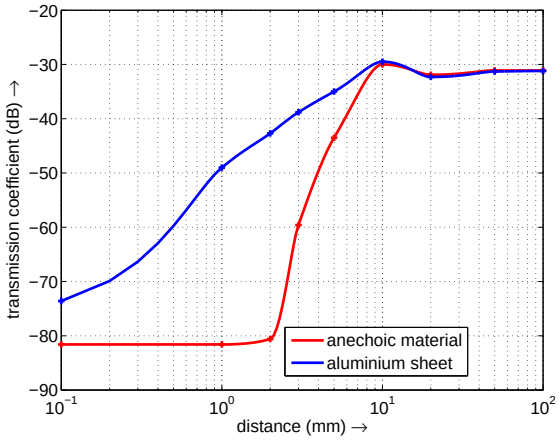


Fig. 5. Dependency of the transmission coefficient on the position of the obstacles for two different materials. Antennas WR15 are used.

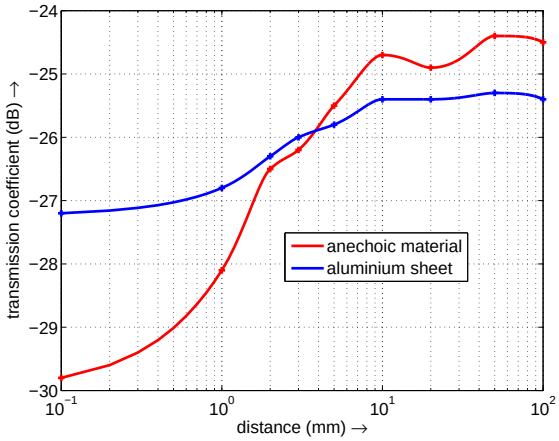


Fig. 6. Dependency of the transmission coefficient on the position of the obstacles for two different materials. Upholstery is put between the antennas WR15.

due to the diffraction on the bottom edge of the aluminium sheet.

IV. Conclusion

The experimental investigation shows that the propagation of surface waves is sensitive to obstacles in the path, which is located in the zone about 10 mm above the surface, irrespective of antenna's directivity.

Furthermore, from the results of testing of the materials is considered that it is preferable to use the anechoic material as a reference obstacle, since it does not show diffraction consequences, as in the case of the usage of an aluminium plate. Diffraction causes less resolution of the obstacle presence in the case that the upholstery is inserted between antennas.

Acknowledgment

The presented research was supported by the Czech Grant Agency project no. P102/12/1274, by the Czech Ministry of Education in frame of National Sustainability Program under grant LO1401, and by the Internal Grant Agency of Brno University of Technology project no. FEKT-S-14-2483. The research is the part of the COST Action IC1301 which is financially supported by the grant of the Czech Ministry of Education no. LD14057. For research, infrastructure of the SIX Center was used.

References

[1] Collin, R. E., *Field Theory of Guided Waves*, Wiley-IEEE Press, 1991.

[2] Schelkunoff, S.A., "Anatomy of "Surface waves"," *Antennas and Propagation, IRE Transactions on* , vol.7, no.5, pp.133,139, December 1959

[3] Barlow, H.M.; Cullen, A.L., "Surface waves," *Proceedings of the IEE - Part III: Radio and Communication Engineering* , vol.100, no.68, pp.329,341, November 1953

[4] Collen, A.L., "The excitation of plane surface waves," *Proceedings of the IEE - Part III: Radio and Communication Engineering* , vol.101, no.72, pp.276,277, July 1954

[5] Brick, D.B., "The Excitation of Surface Waves by a Vertical Antenna," Proceedings of the IRE , vol.43, no.6, pp.721,727, June 1955

Current Video Coding - Solutions and Challenges

Ondrej Zach
Dept. of Radio Electronics
Brno University of Technology
Technicka 12, 616 00 Brno, Czech Republic
Email: ondrej.zach@phd.feec.vutbr.cz

***Abstract*—Presented paper deals with current solutions for video coding. The main focus will be dedicated to state-of-the-art research in video coding area, mainly High Efficiency Video Coding (HEVC). Next part of the paper will focus on possibilities of multimedia streaming. Streaming multimedia content over the Internet brings many advantages, as well as new challenges.**

I. Introduction

Creating video content has never been easier than it is nowadays. Almost everyone has at least one device capable of taking moving pictures. Regular cell phones, smart phones or cameras with video recording functionality are widely spread and also specialized camcorders are commonly used. According to statistics of the video server YouTube [1], their users upload more than 300 hours of video data every minute. Such high amount of video content has significant demands on data storage. This is where video coding comes into the game.

The history of modern video coding started in 1988 with H.261 - the first usable video coding algorithm created by the ITU. It was H.261 [2] that introduced coding techniques used nowadays. Macroblock structure, 4:2:0 chroma subsampling, inter-picture processing and Discrete Cosine Transform (DCT) and its derivatives are used in majority of up-to-date video compression algorithms. As another milestone in video coding can be regarded the year 1995 where the H.262, commonly known as MPEG-2 Part 2 (ISO/IEC name) was firstly introduced, [3]. Although MPEG-2 Part 2 was not very suitable for lower bitrates, it found its use on DVDs and even in this time is still used in Digital Video Broadcasting (DVB) systems. However, its share will probably decline due to its successor H.264. H.264 (also known MPEG-4/AVC) serves as a golden standard of video coding nowadays, [4]. H.264 is used practically in all video distribution systems: Blu-ray discs, DVB systems of the second generations, mainly in HD or online video streaming services are using this video coding standard. However, as the progress in video technology leads to UHD, the H.264 begins to be insufficient.

The most recent coding algorithm presented as a joint collaboration of both ITU and MPEG is H.265/HEVC. This coding algorithm promises preserving the same visual quality as with H.264 with using just half of the bitrate. It is also well suited for encoding UHD videos, therefore it is considered to replace H.264 in near future.

The paper is organized as follows: Section II. is focused on High Efficiency Video Coding (HEVC) and the current research in this area. Section III. introduces new challenges of Adaptive Streaming with use of HEVC. Section IV. summarizes new opportunities and last section concludes.

II. Modern Video Coding - is HEVC the Future?

Current research regarding video coding is mainly focused on HEVC. There are also some studies on VP9. However, these papers usually only evaluate the performance of VP9 compared to HEVC or AVC, [7], or evaluate the Quality of Experience (QoE) when using VP9, [8]. In other words, VP9 does not seem to be an interesting topic in the scientific field and its massive use is uncertain.

On the other hand, H.265/HEVC has the full attention of researchers in the field. The researchers' community investigates all components of the standard. As in the case of its predecessor H.264/AVC, high amount of studies is focused on optimization of individual coding tools. Very hot topic is optimization of motion estimation and fast motion vector decision, [9]. Other studies focus on optimization of TU size decision, [10]. Optimization and speed-up of these essential parts of an encoder is an important task for painless adoption of the standard by the market.

Other part of research is studying the possibilities of hardware implementation of H.265/HEVC. One of the topics in this sub-field is a real-time H.265/HEVC encoder, [11]. The authors of this paper proposed and designed a FPGA-based device capable of coding FulHD 1080p 60fps video in real time using HEVC. Their study also mentions the future research in real-time UHD encoder. Development of such devices also plays a big role in further adoption of HEVC in real practice.

However, majority of studies is focusing on improvements of performance of the encoders in terms of computational resources. Studies of HEVC involving real applications are quite rare. In [12], the authors evaluate the behavior of HEVC in mobile environment (streaming over mobile network). However, they were using only SD sequences, considering mobile phone as their presentation device. A study utilizing both HEVC and UHD resolution is published in [13]. The authors held a subjective study and compared the behavior of AVC and HEVC. However, they were using only 5 s long sequences. There are almost no studies evaluating effect of longer sequences.

Apart from conventional video coding techniques, the research is also aiming to different possibilities. One of them seems to be *distributed video coding* (DVC). A study comparing current DVC architectures is presented in [14]. DVC is a technique where the computational demands are moved to decoder rather then to encoder and the process of encoding

the video can be compared with turbo-coding. The higher decoding load can be even considered as a benefit in case where on the decoding side, there is a PC or a video server with enough computational power. Thanks to this, the encoding device can be quite simple and cheap. From its principles, DVC is suited for example for video security systems or CCTV. Nevertheless, DVC does not seem to be widely adopted in near future.

III. Adaptive Streaming - Solution for Any Device

HTTP adaptive streaming is a technology for delivering to the user the best achievable quality of the streamed video sequence possible no matter of the actual state of his internet connection. This technology enables to adapt the bitrate of the video to avoid stalling during video playback. However, some changes have to be made in the video delivery system as both client and server have to support HAS.

A. Basic principles of HAS

The basic idea behind HAS is, that the video content is split into *segments*. These segments should be of the same length; the specific value depends on the concrete implementation of HAS and usually lies in the range from 2 s to 10 s. The length of the segments influences the frequency of the adaptation as the quality change can be made on the borders of the segments only. These segments are then independently encoded using different quality levels. Finally, we have one video content split into many segments, each segment available at all quality levels. Therefore, the quality of the video can vary over time.

To lower the bitrate of the video, we can use adaptation in following dimension:

- Quantization,
- spatial resolution,
- framerate.

Adaptation of a specific dimension has different impact on the QoE and is a part of the current research in the field of HAS.

An important part of HAS is the client software (player). The player has to maintain the measurements of the connection and plan which segment to download next. The decision logic of the client is determined by the specific implementation of HAS. Currently, there are several proprietary solution of HAS. A list of the implementations and corresponding companies follows:

- HTTP Live Streaming, Apple,
- HTTP Dynamic Streaming, Adobe,
- Silverlight Smooth Streaming, Microsoft.

All these implementations have their benefits and are used in practice. But they have also common limitations. All these implementations allow to use only specified coding algorithms for video and audio and usually have fixed segment length. The only currently available HAS implementation without these limitations is MPEG Dynamic Adaptive Streaming over HTTP (MPEG DASH). [15], therefore it is well suited for utilizing HEVC or other possible modern coding algorithms.

B. Current research in HAS

A comprehensive study of HTTP adaptive streaming is presented in [16]. The authors describe the basics of HAS and its implementations.The study also evaluates the influence of specific *adaptation dimension* on the QoE.

Another study on HAS is presented in [17]. The authors of this study evaluate the behavior of MPEG-4/AVC and scalable video coding (SVC) in HAS. A survey of previously published studies on subjective quality in HTTP adaptive streaming can be found in [18]. This paper summarizes results of known studies and tries to ask still unresolved questions regarding HAS. However, behavior of HEVC regarding HTTP adaptive streaming has not been not well described yet. Therefore, this is the reason, why should HEVC in HAS get our attention. Also the possibilities of scalableHEVC (scalable extension of H.265/HEVC) have not been fully investigated. This is the area where will be the focus of this dissertation thesis.

A study different from above mentioned is presented in [19]. The authors of this paper investigate the possibilities of bandwidth estimation in HAS. As HAS is based on adaptation the bitrate of the video to the actual conditions of the broadband connection, the estimation of available bandwidth is an obvious need. Similarly-based study is presented in [20]. In this paper, the authors evaluate the behavior of HAS in mobile networks on tablet devices. As the progress is aiming to absolute mobility and users already consume the content on mobile devices, this is also an interesting topic. However, this area seems to be studied with enough attention.

IV. New Challenges

As the behavior of the consumers is moving towards watching video on mobile devices, HTTP Adaptive Streaming seems to be a suitable approach to be used in video delivery systems. Other demand which is arising is the popularity of HD and beyond HD resolution. This might be achievable by using modern video coding mechanisms, such as HEVC. Currently, there are only a few studies considering HAS with HEVC. As the progress is aiming to 4K also, the application of HEVC will be more important.

New challenges are expected to be in the field of optimization of video delivery systems in such a way, that one system will be capable of delivering video content to any device connected using any type of connection (mobile, wireless, fixed) with wide variety of resolutions from sub SD and SD up to UHD.

V. Conclusion

In this paper, we briefly described the state-of-the-art research in the field of video coding and video delivery systems. We focused on two topics, which are currently the most popular in the academic circles. It seems that both HEVC and HAS are the most suitable solutions for delivering high quality video content to the wide variety of online connected devices. The mentioned challenges will serve as the basics of our further research.

ACKNOWLEDGMENT

This paper was supported by the project no. CZ.1.07/2.3.00/20.0007 and by the BUT project no. FEKT-S-14-2177. The described research was performed in laboratories supported by the SIX project; no. CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation. Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

REFERENCES

[1] YouTube, LLC. *Statistics* [online]. 2015. [cit. 2015-04-02]. Available at: <https://www.youtube.com/t/contact_us>.

[2] ITU-T Recommendation H.261. *Video codec for audiovisual services at p 64 kbits*. The International Telecommunication Union, Geneva, Switzerland, 1993.

[3] ITU-T Recommendation H.262. *Information technology - generic coding of moving pictures and associated audio information : video*. The International Telecommunication Union, Geneva, Switzerland, 1995.

[4] ITU-T Recommendation H.264. *Advanced video coding for generic audiovisual services*. The International Telecommunication Union, Geneva, Switzerland, 2014.

[5] ITU-T Recommendation H.265. *High efficiency video coding*. The International Telecommunication Union, Geneva, Switzerland, 2014.

[6] WebM: an open web media project. *VP9 video codec*. [online]. Available at: <http://www.webmproject.org/vp9/>.

[7] GROIS, Dan, MARPE, Detlef, MULAYOFF, Amit, ITZHAKY Benaya and Ofer HADAR. Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders. *2013 Picture Coding Symposium (PCS)*. IEEE, 2013, p. 394-397. DOI: 10.1109/PCS.2013.6737766.

[8] RAMZAN, Naeem, PERVEZ Zeeshan and Abbes AMIRA. Quality of experience evaluation of H.265/MPEG-HEVC and VP9 comparison efficiency. 2014 26th International Conference on Microelectronics (ICM). IEEE, 2014, p. 220-223. DOI: 10.1109/ICM.2014.7071846.

[9] ZHANG, Hao and Zhan MA. Fast Intra Mode Decision for High Efficiency Video Coding (HEVC). *IEEE Transactions on Circuits and Systems for Video Technology*. 2014, vol. 24, issue 4, p. 660-668. DOI: 10.1109/TCSVT.2013.2290578.

[10] WANG, Chou Chen, LIAO, Yen Czu, WANG, Jing Wein and Chi Wei TUNG. An Effective TU Size Decision Method for Fast HEVC Encoders. *2014 International Symposium on Computer, Consumer and Control*. IEEE, 2014, p. 1195-1198. DOI: 10.1109/IS3C.2014.310.

[11] MIYAZAWA, Kazuyuki, SAKATE, Hiroharu, SEKIGUCHI Shun-ichi et al. Real-time hardware implementation of HEVC video encoder for 1080p HD video. *2013 Picture Coding Symposium (PCS)*. IEEE, 2013, p. 225-228. DOI: 10.1109/PCS.2013.6737724.

[12] NIGHTINGALE, James, WANG, Qi, GRECOS, Christos and Sergio GOMA. The impact of network impairment on quality of experience (QoE) in H.265/HEVC video streaming. *IEEE Transactions on Consumer Electronics*. 2014, vol. 60, issue 2, p. 242-250. DOI: 10.1109/TCE.2014.6852000.

[13] HANHART, Philippe, RERABEK, Martin, DE SIMONE, Francesca, EBRAHIMI, Touradj and Andrew G. TESCHER. Subjective quality evaluation of the upcoming HEVC video compression standard. *SPIE 8499, Applications of Digital Image Processing XXXV*. DOI: 10.1117/12.946036.

[14] LEI, Ted Chih Wei and Fan Shuo TSENG. Study for Distributed Video Coding Architectures. *2014 International Symposium on Computer, Consumer and Control*. IEEE, 2014, p. 380-383. DOI: 10.1109/IS3C.2014.105.

[15] ISO/IEC. *23009-1:2012 Information Technology Dynamic Adaptive Streaming over HTTP (DASH) Part 1: Media Presentation Description and Segment Formats*. 2012.

[16] SEUFERT, Michael, EGGER, Sebastian, SLANINA, Martin, et al. A Survey on Quality of Experience of HTTP Adaptive Streaming. *IEEE Communications Surveys & Tutorials*. 2014, vol. PP, issue. 99. DOI: 10.1109/comst.2014.2360940.

[17] KALVA, Hari, ADZIC, Velibor and Borko FURHT. Comparing MPEG AVC and SVC for adaptive HTTP streaming. *2012 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2012, p. 158-159. DOI: 10.1109/ICCE.2012.6161787.

[18] GARCIA, M.-N, DE SIMONE,F., TAVAKOLI, S., STAELENS, N., EGGER, S., BRUNNSTROM, K. and A. RAAKE. Quality of experience and HTTP adaptive streaming: A review of subjective studies. *2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2014, p. 141-146. DOI: 10.1109/QoMEX.2014.6982310.

[19] YUN, Dooyeol, CHUNG, Kwangsue and Jinpyo HONG. Efficient Bandwidth Estimation for HTTP adaptive streaming. *The International Conference on Information Networking 2014 (ICOIN2014)*. IEEE, 2014, p. 464-468. DOI: 10.1109/ICOIN.2014.6799725.

[20] STAELENS, N., DE MEULENAERE, J.; CLAEYS, M. et al. Subjective Quality Assessment of Longer Duration Video Sequences Delivered Over HTTP Adaptive Streaming to Tablet Devices. *Broadcasting, IEEE Transactions on* . vol.60, no.4, pp.707-714. DOI: 10.1109/TBC.2014.2359255