FACULTY OF ELECTRICAL ENGINEERING
UNIVERSITY OF BANJA LUKA

ELECTRONICS

VOLUME 18, NUMBER 1, JUNE 2014

FACULTY OF ELECTRICAL ENGINEERING
UNIVERSITY OF BANJA LUKA

Address: Patre 5, 78000 Banja Luka, Bosnia and Herzegovina
Phone: +387 51 211824
Fax: +387 51 211408
Web: www.etfbl.net

ELECTRONICS

Web: www.electronics.etfbl.net
E-mail: electronics@etfbl.net

Editor-in-Chief:

Branko L. Dokić, Ph. D.
Faculty of Electrical Engineering, University of Banja Luka, Bosnia and Herzegovina
E-mail: branko.dokic@etfbl.net

Co Editor-In-Chief:

Prof. Tatjana Pešić-Brđanin, University of Banja Luka, Bosnia and Herzegovina
E-mail: tatjanapb@etfbl.net

International Editorial Board:

- Prof. Goce Arsov, St. Cyril and Methodius University, Macedonia
- Prof. Zdenka Babić, University of Banja Luka, Bosnia and Herzegovina
- Prof. Petar Biljanović, University of Zagreb, Croatia
- Prof. Milorad Božić, University of Banja Luka, Bosnia and Herzegovina
- Prof. Octavio Nieto-Taladriz Garcia, Polytechnic University of Madrid, Spain
- Dr Zoran Jakšić, IHTM, Serbia
- Prof. Vladimir Katić, University of Novi Sad, Serbia
- Prof. Tom J. Kazmierski, University of Southampton, United Kingdom
- Prof. Vančo Litovski, University of Niš, Serbia
- Dr Duško Lukač, University of Applied Sciences, Germany
- Prof. Danilo Mandić, Imperial College, London, United Kingdom
- Prof. Bratislav Milovanović, University of Niš, Serbia
- Prof. Vojin Oklobdžija, University of Texas at Austin, USA
- Prof. Predrag Pejović, University of Belgrade, Serbia
- Prof. Ninoslav Stojadinović, University of Niš, Serbia
- Prof. Robert Šobot, Western University, Canada
- Prof. Slobodan Vukosavić, University of Belgrade, Serbia
- Prof. Volker Zerbe, University of Applied Sciences of Erfurt, Germany

Secretary:

Mladen Knežić, M.Sc.
Željko Ivanović, M.Sc.

Publisher:

Faculty of Electrical Engineering, University of Banja Luka, Bosnia and Herzegovina

Number of printed copies: 100

Editorial

THE oldest and the most prestigious expert association in western Balkans, Serbia-based ETRAN Society has been organizing conferences dedicated to various aspects of electrical and electronics engineering since 1955. The conference is split into 16 separate symposia, each dedicated to a different aspect of electrical and electronics engineering. For its 57th conference held in 2013 in Zlatibor, Serbia the ETRAN Society introduced the best paper awards, where each section grants its award to the best paper picked by a jury of eminent members of the ETRAN Society, each of them expert for the field covered by the section. Each section also chose one outstanding presentation besides the awarded paper. For this issue of Electronics journal we picked three such presentations awarded at the ETRAN 2013 conference and asked the authors to submit significantly extended and rewritten versions of their manuscripts for this issue of Electronics. The papers were further subject to the standard reviewing procedure for Electronics. One of these presentations was the award recipient in Electronics Section (the paper by Miona Andrejević Stošović et al, dedicated to simulation of maximum power point tracking in photovoltaic systems), another in Automatics Section (the paper by Sanja Vujnović et al dealing with the use of Bayesian networks in thermal power plants), and one paper was chosen as the outstanding presentation in Electronics Section (the paper by Grigor Y. Zargaryan et al, dedicated to simulation and verification of USB controllers). In this way the ETRAN Society continues its already fruitful cooperation with the ETRAN Society.

We are pleased to recommend readers of this issue of Electronics journal to pay attention to the invited paper „Energy Efficient Multi-Core Processing“ by C. Leech and T. Kazmierski. In this paper, the present state of the art of energy-efficient embedded processor design techniques is evaluated, and it is demonstrated how small, variable architecture embedded processors may exploit a run-time minimal architectural synthesis technique to achieve greater energy and area efficiency whilst maintaining performance.

The paper „Measuring Capacitor Parameters Using Vector Network Analyzers“, by Deniss Stepins et al, is focused on the measurement accuracy of capacitor parameters using VNA (Vector Network Analyzers) and proper de-embedding of an experimental setup parasitics to get accurate measurement results.

Jasmin Igić and Milorad Božić in their paper present an improvement of the Approximate Internal Model-based Neural Control (AIMNC) structure. The authors also suggest an approach in which one can ensure satisfactory behavior of the AIMNC law for controlling slow industrial processes and provide the zero steady state error in the cases of constant reference signal and constant disturbances.

The two remaining papers, „Data-Driven Gradient Descent Direct Adaptive Control for Discrete-Time Nonlinear SISO Systems“ and „Integrated Cost-Benefit Assessment of Customer-Driven Distributed Generation“, are related to the research results published in the doctoral dissertations of Igor Krčmar and Čedomir Željković, respectively.

Prof. Bratislav Milovanović
Chair of ETRAN Society

Prof. Zoran Jakšić
Chair of ETRAN Program Committee

Prof. Tatjana Pešić-Brđanin
Electronics journal Co Editor-In-Chief

Energy Efficient Multi-Core Processing

Charles Leech and Tom J. Kazmierski

(Invited Paper)

***Abstract*—This paper evaluates the present state of the art of energy-efficient embedded processor design techniques and demonstrates, how small, variable-architecture embedded processors may exploit a run-time minimal architectural synthesis technique to achieve greater energy and area efficiency whilst maintaining performance. The picoMIPS architecture is presented, inspired by the MIPS, as an example of a minimal and energy efficient processor. The picoMIPS is a variable-architecture RISC microprocessor with an application-specific minimised instruction set. Each implementation will contain only the necessary datapath elements in order to maximise area efficiency. Due to the relationship between logic gate count and power consumption, energy efficiency is also maximised in the processor therefore the system is designed to perform a specific task in the most efficient processor-based form. The principles of the picoMIPS processor are illustrated with an example of the discrete cosine transform (DCT) and inverse DCT (IDCT) algorithms implemented in a multi-core context to demonstrate the concept of minimal architecture synthesis and how it can be used to produce an application specific, energy efficient processor.**

***Index Terms*—Embedded processors, application specific architectures, MIPS architecture, digital synthesis, energy efficient design, low power design.**

Review Paper
DOI: 10.7251/ELS1418003L

I. Introduction

THE energy efficiency of embedded processors is essential in mobile electronics where devices are powered by batteries. Processor performance has been increasing over the last few decades at a rate faster than the developments in battery technologies. This has led to a significant reduction of the battery life in mobile devices from days to hours. Also, new mobile applications demand higher performance and more graphically intensive processing. These demands are currently being addressed by many-core, high-frequency architectures which can deliver high-speed processing necessary to meet the tight execution deadlines. These two contradictory demands, the need to save energy and the requirement to deliver outstanding performance must be addressed by entirely new approaches. A number of research directions have appeared. Heterogeneous and reconfigurable embedded many-core systems can improve energy efficiency while maintaining high speed through judicious task scheduling and hardware adaptability. In a heterogeneous system, such as the ARM big.LITTLE architecture [1] smaller cores are employed to process simple and less demanding tasks to save energy while larger cores handle high performance and energy hungry processing when necessary. Reconfigurable architectures use flexible interconnect, power gating and software control within each core, thus achieving heterogeneity within the core. Reconfigurable cores can be configured in this way as either slower, but energy efficient processors, or faster, high-performance cores.. A number of approaches have been proposed to save energy within a core. Dynamic Voltage and Frequency Scaling (DVFS) [2] is a popular and well established technique where the supply voltage and the clock frequency are scaled to trade energy for performance and vice-versa. DVFS is typically implemented by including voltage regulators and phase-lock loop controlled clocks in the processor. The architecture is modified to allow the operating system to select a desired voltage and frequency through writing data to a DVFS control register. At any desired performance level, the operating system will put the processor into a minimum energy consumption mode. DVFS has proved very effective especially in applications where high performance is peaking only during a small fraction of the operating time as significant energy savings are achieved. Many other energy saving design techniques are currently being explored at the circuit, architecture and even system level. For example supply voltage in bus drivers can be reduced to extremely low levels to reduce bus energy consumption. New SRAM designs are being developed where energy consumption is reduced to extremely low levels in both the on-chip caches and the external memories. The architecture of processor cores are traditionally determined from a compromise of speed, power consumption, scalability, maintainability and extensibility. However, applications have different characteristics that require specific hardware implementations to enable optimal performance and therefore a system should be able to adapt its architecture to each application scenario. In this paper, we aim to demonstrate, through the evaluation of present technology, how small, variable-architecture embedded processors may exploit a run-time minimal architectural synthesis technique to achieve greater energy and area efficiency whilst maintaining performance.

Manuscript received 29 May 2014. Accepted for publication 12 June 2014.

C. Leech and T. J. Kazmierski are with the University of Southampton, UK (e-mail: {cl19g10, tjk}@ecs.soton.ac.uk).

II. Overview of Energy Efficiency Technologies

This section presents the current state of research in energy efficient technologies in multi-core systems for both traditional power saving techniques and novel technologies including heterogeneous and reconfigurable architectures. Through the analysis of present technology, we aim to demonstrate how a greater performance, energy efficiency and area efficiency balance can be achieved.

The introduction of multi-core structures to processor architectures has caused a significant increase in the power

consumption of these systems. In addition, the gap between average power and peak power has widened as the level of core integration increases [3]. A global power manager policy, such as that proposed by Isci et al, that has per-core control of parameters such as voltage and frequency levels is required in order to provide effective dynamic control of the power consumption [3]. Metrics such as performance-per-watt [4], [5], average and peak power or energy-delay product [6] are all used to quantify the power or energy efficiency of a system in order to evaluate it, however properties are prioritised differently depending on the application requirements. Modelling and simulation of multi-core processors is also an important process in order to better understand the complex interactions that occur inside a system and cause power and energy consumption [7]–[11]. For example, the model created by Basmadjian et al is tailored for multi-core architectures in that it accounts for resource sharing and power saving mechanisms [8].

A. Energy Efficiency techniques in Static Homogeneous Multi-core Architectures

Many energy efficiency and power saving technologies are already integrated into processor architectures in order to reduce power dissipation and extend battery life, especially in mobile devices. A combination of technologies is most commonly implemented to achieve the best energy efficiency whilst still allowing the system to meet performance targets [4]. Techniques to increase energy efficiency can be applied at many development levels from architecture co-design and code compilation to task scheduling, run-time management and application design [12]. A summary and analysis of these technologies is presented in the following section.

1) Dynamic Voltage and Frequency Scaling: Dynamic Voltage and Frequency Scaling (DVFS) is a technique used to control the power consumption of a processor through fine adjustment of the clock frequency and supply voltage levels [3], [4], [12], [13]. High levels are used when meeting performance targets is a priority and low levels (known as CPU throttling) are used when energy efficiency is most important or high performance is not required. When the supply voltage is lowered and the frequency reduced, the execution of instructions by the processor is slower but performed more energy efficiently due to the extension of delays in the pipeline stages. The rise and fall times for logic circuitry is increased along with the clock period meaning performance targets for applications must be relaxed. DVFS can be used in homogeneous multi-core architectures to emulate heterogeneity by controlling the frequency and supply to each core individually [14]. Each core therefore appears as though it has different delay properties however the architectures are still essentially identical. This per-core DVFS mechanism is investigated by Wonyoung et al who conclude that significant energy saving opportunities exist where on-chip integrated voltage regulators are used to provide nanosecond scale voltage switching [13]. DVFS can also be combined with thread migration to reduce energy consumption [4], [15]. Cai et al cite the problem that present DVFS energy saving techniques on multi-core systems assume one hardware context for each core whereas simultaneous multi-threading (SMT) is commonly implemented which causes these techniques to be less effective. Their novel technique, known as thread shuffling, uses concepts of thread criticality and thread criticality degree instead of thread-delay to maps together threads with similar criticality degree. This accounts for SMT when implementing DVFS and thread migration and achieves energy savings of up to 56% without impeding performance at all.

2) Clock Gating and Clock Distribution: Clock gating is a process, applied at the architectural design phase, to insert additional logic between the clock source and clock input of the processor's circuitry. During program execution, it reduces power consumption by logically disconnecting the clock of synchronous logic circuits to prevent unnecessary switching. Classed as a Dynamic Power Management (DPM) technique, as it is applied at run-time along with other techniques such as thread scheduling and DVFS to optimise the power/performance trade-off of a system [12]. The clock gating and distribution techniques implemented by Qualcomm in the Hexagon processor are analysed by Bassett et al on their ability to improve the energy efficiency of a digital signal processors (DSP) [16]. A low power clock network is implemented using multi-level clock gating strategies and spine-based clock distribution. The 4 levels of clock gating allows different size regions of the chip to be deactivated, from entire cores down to single logic cells. Further power reduction is achieved through a structured clock tree that aims to minimise the power consumed in distributing the clock signal across the chip whilst avoiding clock skew and delay. The clock tree structure (CTS) examined by Bassett et al is tested to give a 2 time reduction in skew over traditional CTS while power tests show reductions in power consumption by 8% for high-activity and over 35% for idle mode. Large portions of the chip will spend the majority of their time in idle more therefore high efficiency in this mode is critical.

3) Power Domains: Power domains are regions of a system or processor that are controlled from a single supply and can be completely powered down in order to minimise power consumption without entirely removing the power supply to the system. Power domains can be used dynamically and when used in conjunction with clock gating, lead to further improvements in energy efficiency. The ARM Cortex-A15 MPCore processor supports multiple power domains both for the core and for the surrounding logic [17]. Figure 1 shows these domains, labelled Processor and Non-Processor, that allow large parts of the processor to be deactivated. Smaller internal domains, such as CK_GCLKCR, are implemented to allow smaller sections to be deactivated for finer performance and power variations.

Power domains are often coupled with power modes as a means of switching on or off several power domains in order to enter low power, idle or shutdown states. The Cortex-A15 features multiple power modes with specific power domain configurations such as Dormant mode, where some Debug and L2 cache logic is powered down but the L2 cache RAMs are powered up to retain their state, or Powerdown mode where all power domains are powered down and the processor state

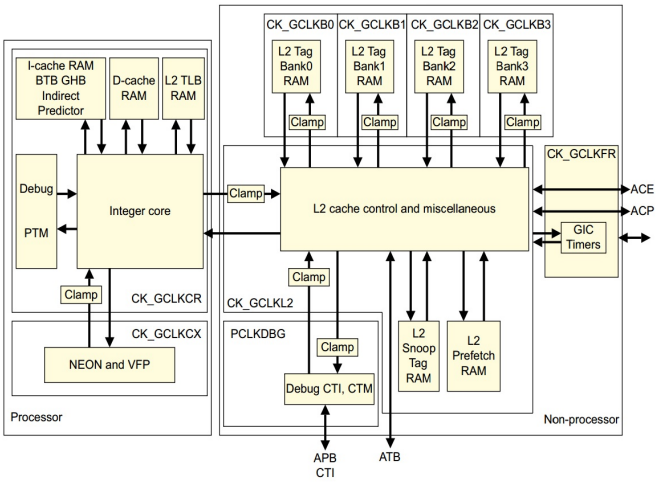


Fig. 1. The ARM Cortex-A15 features multiple power domains for the core and surrounding logic, reprinted from [17].

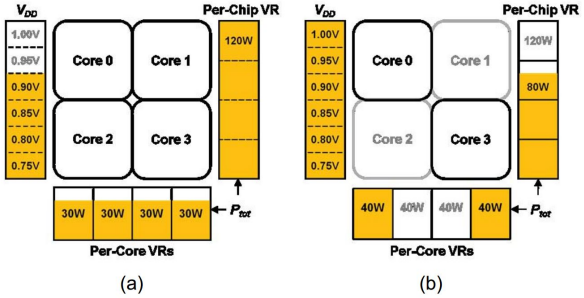


Fig. 2. Per-core power domains can provide reduce power consumption and higher performance levels, reprinted from [18].

is lost [17]. In multi-core architectures, power domains can be used to power down individual cores when idle or during non-parallel tasks in order to manage power consumption. Sinkar et al [18] and Ghasemi et al [19] present low-cost, per-core voltage domain solutions in order to improve performance in power-constrained processors. Figure 2 shows how Sinkar's mechanism can provide reduced chip power consumption but maintain per-core performance. In (a), a chip-wide power domain is shown with all cores active at the same level. In contrast, (b) shows a per-core power domain which allows unnecessary cores to be powered down and active cores to provide a higher performance level while still reducing the overall chip power level.

In the work by Rotem et al [20], topologies containing multiple clock, voltage and frequency domains are investigated in order to build a high-end chip multiprocessor (CMP) considering both power constraints and performance demands. These domains are controlled using DVFS techniques in connection with clock gating and are exercised by simulating power supply constrains. Results showed that multiple power domains can be beneficial for fully threaded applications whereas a single power domain is more suitable for light-threaded workloads. Power domains can be linked to DVFS techniques if the domain contains multiple voltage levels. Per-core power domains therefore enables per-core DVFS control such that each core can exploit run-time performance variations in multi-threaded applications [13], [18], [20].

4) Pipeline Balancing: Pipeline balancing (PLB) is a technique used to dynamically adjust the resources of the pipeline of a processor such that it retains performance while reducing power consumption [21]. The delay constraints on microarchitectural pipeline stages can be modified in order to make them more power efficient, in a similar way to DVFS, when the performance demand of the application is relaxed [22]. In work by Bahar et al, PLB operates in response to instruction per cycle (IPC) variations within a program [21]. The PLB mechanism dynamically reduces the issue width of the pipeline to save power or increases it to boost throughput. An 8-wide issue pipeline that has its unified issue queue divided into a left and right arbiter, separate left and right register files and functional units. A control unit is included can deactivate the right arbiter and its functional units to allow the pipeline to enter a low power mode. They show that this PLB technique can reduce power consumption of the issue queue and execution unit by up to 23% and 13% respectively with only an average performance loss of 1% to 2% [21]. Power Balanced pipelines is a concept in which the power disparity of pipeline stages is reduced by assigning different delays to each microarchitectural pipe stage while guaranteeing a certain level of performance/throughput [22]. In a similar fashion to [21], the concept uses cycle time stealing to redistribute cycle time from low power stages to stages that can perform more efficiently if given more time. Static power balancing is performed during design time to identify power heavy circuitry in pipe stages for which consumption remains fairly constant for different programs and reallocate cycle time accordingly. Dynamic power balancing is implemented on top of this to manage power fluctuations within each workload and further reduce the total power cost. Balancing of power rather than delay can result in a 46% power consumption reduction by the processor with no loss in throughput for a FabScalar processor over a baseline comparison. Power savings are also greater at lower frequencies.

5) Caches and Interconnects: It is not only the design of the processor's internal circuitry that is important in maintaining energy efficiency. Kumar et al conclude, from a paper examining the interconnections in multi-core architectures, that careful co-design of the interconnect, caches and the processor cores is required to achieve high performance and energy efficiency [23]. Through several theoretical examples, they examine parameters such as the area, power and bandwidth costs required to implement the processor-cache interconnect, showing that large caches sizes can constrain the interconnect's size and large interconnects can be power-hungry and inefficient. Zeng et al also recognise the high level of integration that is inherent in CMPs and attempt to reduce the interconnect power consumption by developing a novel cache coherence protocol [24]. Using their technique, results show that an average of 16.3% of L2 cache accesses could be optimised and as every access consumes time and power, an average 9.3% power reduction is recorded while increasing system performance by 1.4%.

B. Energy Efficiency techniques in Heterogeneous Multi-core Architectures

A heterogeneous or asymmetric multi-core architecture is composed of cores of varying size and complexity which are designed to compliment each other in terms of performance and energy efficiency [6]. A typical system will implement a small core to process simple tasks, in an energy efficient way, while a larger core provides higher performance processing for when computationally demanding tasks are presented. The cores represent different points in the power/performance design space and significant energy efficiency benefits can be achieved by dynamically allocating application execution to the most appropriate core [25]. A task matching or switching system is also implemented to intelligently assign tasks to cores; balancing a performance demand against maintaining system energy efficiency. These systems are particularly good at saving power whilst handling a diverse workload where fluctuations of high and low computational demand are common [26].

A heterogeneous architecture can be created in many different ways and many alternative have been developed due to the heavy research interest in this area. Modifications to general purpose processors, such as asymmetric core sizes [11], custom accelerators [27], varied caches sizes [14] and heterogeneity within each core [5], [28], have all been demonstrated to introduce heterogeneous features into a system.

One of the most prominent and successful heterogeneous architectures to date is the ARM big.LITTLE system. This is a production example of a heterogeneous multiprocessor system consisting of a compact and energy efficient "LITTLE" Cortex-A7 processor coupled with a higher performance "big" Cortex-A15 processor [26]. The system is designed with the dynamic usage patterns of modern smart phones in mind where there are typically periods of high intensity processing followed by longer periods of low intensity processing [29]. Low intensity tasks, such as texting and audio, can be handled by the A7 processor enabling a mobile device to save battery life. When a period of high intensity occurs, the A15 processor can be activated to increase the system's throughput and meet tighter performance deadlines. A power saving of up to 70% is advertised for a light workload, where the A7 processor can handle all of the tasks, and a 50% saving for medium workloads where some tasks will require allocation to the A15 processor.

Kumar et al present an alternative implementation where two architectures from the Alpha family, the EV5 and EV6, are combined to be more energy and area efficient than a homogeneous equivalent [6], [25]. They demonstrate that a much higher throughput can be achieved due to the ability of a heterogeneous multi-core architecture to better exploit changes in thread-level parallelism as well as inter- and intra- thread diversity [6]. In [25], they evaluate the system in terms of its power efficiency indicating a 39% average energy reduction for only a 3% performance drop [25].

Composite Cores is a microarchitectural design that reduces the migration overhead of task switching by bringing heterogeneity inside each individual core [28]. The design

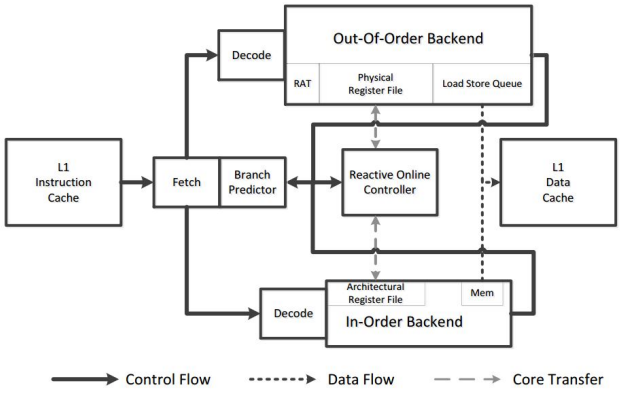


Fig. 3. The microarchitecture for Composite Cores, featuring two μEngines, reprinted from [28].

contains 2 separate backend modules, called μEngines, one of which features a deeper and more complex out-of-order pipeline, tailored for higher performance, while the other features a smaller, compact in-order pipeline designed with energy efficiency in mind. Figure Due to the high level of hardware resource sharing and the small μEngine state, the migration overhead is brought down from the order of 20,000 instructions to 2000 instructions. This greatly reduces the energy expenditure associated with migration and also allows more of the task to be run in an efficient mode. Their results show that the system can achieve an energy saving of 18% using dynamic task migration whilst only suffering a 5% performance loss.

Using both a heterogeneous architecture and hardware reconfiguration, a technique called Dynamic Core Morphing (DCM) is developed by Rodrigues et al to allow the shared hardware of a few tightly coupled cores to be morphed at runtime [5]. The cores all feature a baseline configuration but reconfiguration can trigger the re-assignment of high performance functional units to different cores to speed up execution. The efficiency of the system can lead to performance per watt gains of up to 43% and an average saving of 16% compared to a homogeneous static architecture.

The energy efficiency benefits of heterogeneity can only be exploited with the correct assignment of tasks or applications to each core [7], [10], [30]–[32]. Tasks must be assigned in order to maximise energy efficiency whilst ensuring performance deadlines are met. Awan et al perform scheduling in two phases to improve energy efficiency; task allocation to minimise active energy consumption and exchange of higher energy states to lower, more energy efficient sleep states [7]. Alternatively, Calcado et al propose division of tasks into m-threads to introduce fine-grain parallelism below thread level [33]. Moreover, Saha et al include power and temperature models into an adaptive task partitioning mechanism in order to allocate task according to their actual utilisations rather than based on a worst case execution time [10]. Simulation results confirm that the mechanism is effective in minimising energy consumption by 55% and reduces task migrations by 60% over

alternative task partitioning schemes.

Tasks assignment can also be performed in response to program phases which naturally occur during execution when the resource demands of the application change. Phase detection is used by Jooya and Analoui to dynamically re-assigning programs for each phase to improve the performance and power dissipation of heterogeneous multi-core processors [31]. Programs are profiled in dynamic time intervals in order to detect phase changes. Sawalha et al also propose an online scheduling technique that dynamically adjusts the program-to-core assignment as application behaviour changes between phases with an aim to maximise energy efficiency [32]. Simulated evaluation of the scheduler shows energy saving of 16% on average and up to 29% reductions in energy-delay product can be achieved as compared to static assignments.

C. Energy Efficiency techniques in Reconfigurable Multi-core Architectures

Reconfigurability is another property that has the potential to increase the energy and area efficiency of processors and systems on chip by introducing adaptability and hardware flexibility into the architecture. Building on the innovations that heterogeneous architectures bring, reconfigurable architectures aim to achieve both energy efficiency and high performance but within the same processor and therefore meet the requirements of many embedded systems. The flexible heterogeneous Multi-Core processor (FMC) is an example of the fusion of these two architectures that can deliver both a high throughput for uniform parallel applications and high performance for fluctuating general purpose workloads [34]. Reconfigurable architectures are dynamic, adjusting their complexity, speed and performance level in response to the currently executing application. With this property in mind, we disregard systems that are statically reconfigurable but fixed while operating, such as traditional FPGAs, considering only architectures that are run-time reconfigurable.

1) Dynamic Partial Reconfiguration: FPGA manufacturers such as Xilinx and Altera now offer a mechanism called Dynamic Partial Reconfiguration (DPR) [35] or Self-Reconfiguration (DPSR) [36] to enable reconfiguration during run-time of the circuits within an FPGA, allowing a region of the design to change dynamically while other areas remain active [37]. The FPGA's architecture is partitioned into a static region consisting of fixed logic, control circuits and an embedded processor that control and monitor the system. The rest of the design space is allocated to a dynamic/reconfigurable region containing a reconfigurable logic fabric that can be formed into any circuit whenever hardware acceleration is required.

PDR/PDSR presents energy efficiency opportunities over fixed architectures. PDR enables the system to react dynamically to changes in the structure or performance and power constraints of the application, allowing it to address inefficiencies in the allocation of resources and more accurately implement changing software routines as dynamic hardware accelerators [35]. These circuits can then be easily removed or gated when they are no longer required to reduce power consumption [38]. PDR can also increase the performance of an FPGA based system because it permits the continued operation of portions of the dynamic region unaffected by reconfiguration tasks. Therefore, it allows multiple applications to be run in parallel on a single FPGA [36]. This property also improves the hardware efficiency of the system as, where separate devices were required, different tasks can now be implemented on a single FPGA, reducing power consumption and board dimensions. In addition, PDR reduces reconfiguration times due to the fact that only small modification are made to the bitstream over time and the entire design does not need to be reloaded for each change.

A study into the power consumption patterns of DPSR programming was conducted by Bonamy et al [9] to investigate to what degree the sharing of silicon area between multiple accelerators will help to reduce power consumption. However, many parameters must be considered to assess whether the performance improvement outweighs preventative factors such as reconfiguration overhead, accelerator area and idle power consumption and as such any gain can be difficult to evaluate. Their results show complex variations in power usage at different stages during reconfiguration that is dependent on factors like the previous configuration and the contents of the configured circuit. In response to these experiments, three power models are proposed to help analyse the trade-off between implementing tasks as dynamically reconfigurable, in static configuration or in full software execution.

Despite clear benefits, several disadvantages become apparent with this form of reconfigurable technology. As was shown above, the power consumption overhead associated with programming new circuits can effectively imposed a minimum size or usage time on circuits for implementation to be validated. In addition, a baseline power and area cost is also always created due to the large static region which continuously consumes power and can contain unnecessary hardware. Finally, the FPGA interconnect reduces the speed and increases the power consumption of the circuit compared to an ASIC implementation because of an increased gate count required to give the system flexibility.

2) Composable and Partitionable Architectures: Partitioning and composition are techniques employed by some dynamically reconfigurable systems to provide adaptive parallel granularity [39]. Composition involves synthesising a larger logical processor from smaller processing elements when higher performance computation or greater instruction or thread level parallelism (ILP or TLP) is required. Partitioning on the other hand will divide up a large design in the most appropriate way and assign shared hardware resources to individual cores to meet the needs of an application.

Composable Lightweight Processors (CLP) is an example of a flexible architectural approach to designing a Chip Multiprocessor (CMP) where low-power processor cores can be aggregated together dynamically to form larger single-threaded processors [39]. The system has an advantage over other reconfigurable techniques in that there are no monolithic structure spanning the cores which instead communicate using a microarchitectural protocol. In tests against a fixed-granularity processor, the CLP has been shown to provide a

42% performance improvement whilst being on average 3.4 times as area efficient and 2 times as power efficient.

Core Fusion is a similar technique to CLP in that it allows multiple processors to be dynamically allocated to a single instruction window and operated as if there were one larger processor [40]. The main difference from CLP is that Core Fusion operates on conventional RISC or CISC ISAs giving it an advantage over CLP in terms of compatibility. However, this also requires that the standard structures in these ISAs are present and so can limit the scalability of the architecture.

3) Coarse Grained Reconfigurable Array Architectures: Coarse-Grained Reconfigurable Array (CGRA) architectures represent an important class of programmable system that act as an intermediate state between fixed general purpose processors and fine-grain reconfigurable FPGAs. They are designed to be reconfigurable at a module or block level rather than at the gate level in order to trade-off flexibility for reduced reconfiguration time [41].

One example of a CGRA designed with energy efficiency as the priority is the Ultra Low Power Samsung Reconfigurable Processor (ULP-SRP) presented by Changmoo et al [42]. Intended for biomedical applications as a mobile healthcare solution, the ULP-SRP is a variation of the ADRES processor [43] and uses 3 run-time switch-able power modes and automatic power gating to optimise the energy consumption of the device. Experimental results when running a low power monitoring application show a 46.1% energy consumption reduction compared to previous works.

III. Case Study - picoMIPS

In this section, we present the picoMIPS architecture as an example of a minimal and energy efficient processor implementation. The key points of the architecture will be described and evaluated, showing how it is a novel concept in minimal architecture synthesis. Developments are proposed to the architecture, that can improve performance and maintain energy efficiency, using the technologies described in the previous section.

The picoMIPS architecture is foremost a RISC microprocessor with a minimised instruction set architecture (ISA). Each implementation will contain only the necessary datapath elements in order to maximise area efficiency as the priority. For example, the instruction decoder will only recognise instructions that the user specifies and the ALU will only perform the required logic or arithmetic functions. Due to the proportionality between logic gate count and power consumption, energy efficiency is also maximised in the processor therefore the system is designed to perform a specific task in the most efficient processor-based form.

By synthesising the picoMIPS as a microprocessor, a baseline configuration is established upon which functionality can be added or removed, in the form of instructions or functions, while incurring only minimal changes to the area consumption of the design. If the task was implemented as a specific dedicated hardware circuit, any changes to the functionality could have a large influence on the area consumption of the design. Figure 4 shows an example configuration for the picoMIPS which can accommodate the majority of the simple RISC instructions. It is a Harvard architecture, with separate program and data memories, although the designer may choose to exclude a data memory entirely. The user can also specify the widths of each data bus to avoid unnecessary opcode bits from wasting logic gates.

The principles of the picoMIPS processor have been implemented in a few projects to demonstrate the concept of minimal architecture synthesis and how it can be used to produce an application specific, energy efficient processor. The discrete cosine transform (DCT) algorithm, a stage in JPEG compression, was synthesised into a processor architecture based on the picoMIPS concept. The resulting processor was more area efficient than a GPP due to the removal of unnecessary circuitry however its functionality was also reduced to performing only those functions which appear in the DCT algorithm. The processor can also be compared to a dedicated ASIC hardware implementation of the DCT algorithm. An ASIC implementations have a much higher performance and throughput of data however this is at the cost of area and energy efficiency. The picoMIPS therefore represents a balance between the two, sacrificing some performance for area and energy efficiency benefits.

The picoMIPS has also been implemented to perform the DCT and inverse DCT (IDCT) in a multi-core context [44]. A homogeneous architecture was deployed with the same single core structure, as in figure 4, being replicated 3 times. The cores are connected via a data bus to a distribution module as shown in figure 5 where block data is transferred to each core in turn. This structure theoretically triples the throughput of the system as it can process multiple data blocks in parallel.

As a microprocessor architecture, the picoMIPS can implement many of the technologies discussed in section II in order to improve energy efficiency. Clock gating, power domains and DVFS will all benefit the system however the area overhead of implementing them must first be considered as necessary. Pipeline balancing and caching can be integrated into more complex picoMIPS architectures however these are performance focused improvements and so are not priorities in the picoMIPS concept. The expansion of the system to multi-core is also one that can be employed to improve performance. Moreover, a heterogeneous architecture could be implemented to allow the picoMIPS to process multiple different applications simultaneously using several tailored ISAs. Reconfigurability can also be applied to picoMIPS to create an architecture which can be specific to each application that is executed, effectively creating a general purpose yet application specific processor. This property would require run-time synthesis algorithms to detect and develop the instructions and functional units that are required, before executing the application.

IV. Conclusion

A new concept of variable-architecture application-specific approach to embedded processor design has been presented. Over the past few decades, the trend in processor architectures, has evolved from single core to homogeneous multi-core and

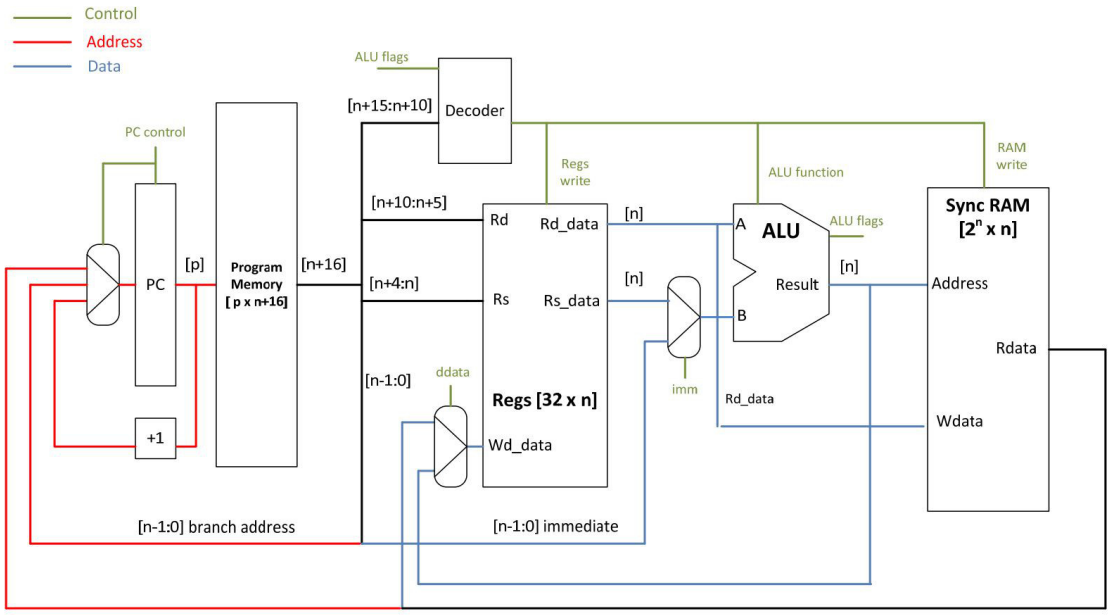


Fig. 4. An example implementation of the picoMIPS architecture.

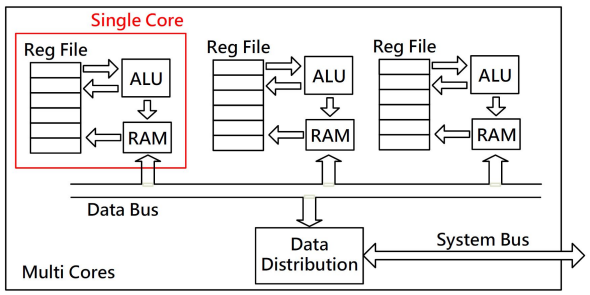


Fig. 5. A Multi-core implementation of the picoMIPS architecture, reprinted from [44].

into heterogeneous multi-core at the present. The proposed approach lends itself easily to a multi or many core architecture design where the performance is improved by enabling the simultaneous execution of threads independently on each core. The basic core design, of a core datapath and accelerators, may be replicated many times and integrated with some form of interconnect. This may form a homogeneous many-core processor, as all the cores are identical when no accelerators are connect. However, the variable-architecture processor is classed as a heterogeneous many-core processor due to the ability of run-time reconfiguration to make each core specific to the particular application that is currently executing on it during normal operation. Moreover, per-core DVFS controls can further differentiate each core using fine grain voltage and frequency adjustments that will affect the power consumption and performance of the core.

In addition to per-core DVFS control, an even finer granularity of control could be permitted through the use of per-accelerator DVFS controls. This links in with the per-accelerator power domains feature which is the first step towards DVFS control. This allows a range of operating modes for each accelerator to allow fine tuning of the systems energy consumption. Accelerators which feature in critical paths of the architecture could be run at higher DVFS levels in order to reduce their latency.

The core level in a many-core system is the smallest duplicative region of the design featuring individual cores that contain their own core datapath and application specific accelerators. An intermediate layer of shared accelerators is implemented to allow neighbouring cores to share accelerators should they required additional hardware support. This approach is similar to core morphing, where cores are weak and strong in the execution of different instruction types. These levels can also form the basis for power domains so that a multi-level power control system can be implemented to allow fine grain control of the power consumption of the chip.

ACKNOWLEDGMENT

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), UK under grant number EP/K034448/1 " PRiME: Power-efficient, Reliable, Many-core Embedded systems" website: (*http://www.prime-project.org/*).

REFERENCES

[1] P. Greenhalgh, "big. LITTLE processing with ARM Cortex-A15 & Cortex-A7," *ARM White Paper*, 2011.

[2] P. Pillai and K. G. Shin, "Real-time dynamic voltage scaling for low-power embedded operating systems," *ACM SIGOPS Operating Systems Review*, vol. 35, no. 5, pp. 89–102, 2001.

[3] C. Isci, A. Buyuktosunoglu, C.-Y. Chen, P. Bose, and M. Martonosi, "An Analysis of Efficient Multi-Core Global Power Management Policies: Maximizing Performance for a Given Power Budget," in *Microarchitecture, 2006. MICRO-39. 39th Annual IEEE/ACM International Symposium on*, 2006, pp. 347–358.

[4] V. Hanumaiah and S. Vrudhula, "Energy-efficient Operation of Multi-core Processors by DVFS, Task Migration and Active Cooling," *Computers, IEEE Transactions on*, vol. 63, no. 2, pp. 349–360, 2012.

[5] R. Rodrigues, A. Annamalai, I. Koren, S. Kundu, and O. Khan, "Performance Per Watt Benefits of Dynamic Core Morphing in Asymmetric Multicores," in *Parallel Architectures and Compilation Techniques (PACT), 2011 International Conference on*, 2011, pp. 121–130.

[6] R. Kumar, D. Tullsen, P. Ranganathan, N. Jouppi, and K. Farkas, "Single-ISA heterogeneous multi-core architectures for multithreaded workload performance," in *Computer Architecture, 2004. Proceedings. 31st Annual International Symposium on*, 2004, pp. 64–75.

[7] M. Awan and S. Petters, "Energy-aware partitioning of tasks onto a heterogeneous multi-core platform," in *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2013 IEEE 19th*, 2013, pp. 205–214.

[8] R. Basmadjian and H. De Meer, "Evaluating and modeling power consumption of multi-core processors," in *Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy), 2012 Third International Conference on*, 2012, pp. 1–10. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6221107

[9] R. Bonamy, D. Chillet, S. Bilavarn, and O. Sentieys, "Power consumption model for partial and dynamic reconfiguration," in *Reconfigurable Computing and FPGAs (ReConFig), 2012 International Conference on*, 2012, pp. 1–8.

[10] S. Saha, J. Deogun, and Y. Lu, "Adaptive energy-efficient task partitioning for heterogeneous multi-core multiprocessor real-time systems," in *High Performance Computing and Simulation (HPCS), 2012 International Conference on*, 2012, pp. 147–153.

[11] D. . Woo and H.-H. Lee, "Extending Amdahl's Law for Energy-Efficient Computing in the Many-Core Era," *Computer*, vol. 41, no. 12, pp. 24–31, 2008.

[12] B. de Abreu Silva and V. Bonato, "Power/performance optimization in FPGA-based asymmetric multi-core systems," in *Field Programmable Logic and Applications (FPL), 2012 22nd International Conference on*, 2012, pp. 473–474.

[13] K. Wonyoung, M. Gupta, G.-Y. Wei, and D. Brooks, "System level analysis of fast, per-core DVFS using on-chip switching regulators," in *High Performance Computer Architecture, 2008. HPCA 2008. IEEE 14th International Symposium on*, 2008, pp. 123–134.

[14] B. de Abreu Silva, L. Cuminato, and V. Bonato, "Reducing the overall cache miss rate using different cache sizes for Heterogeneous Multi-core Processors," in *Reconfigurable Computing and FPGAs (ReConFig), 2012 International Conference on*, 2012, pp. 1–6.

[15] Q. Cai, J. Gonzalez, G. Magklis, P. Chaparro, and A. Gonzalez, "Thread shuffling: Combining DVFS and thread migration to reduce energy consumptions for multi-core systems," in *Low Power Electronics and Design (ISLPED) 2011 International Symposium on*, 2011, pp. 379–384.

[16] P. Bassett and M. Saint-Laurent, "Energy efficient design techniques for a digital signal processor," in *IC Design Technology (ICICDT), 2012 IEEE International Conference on*, 2012, pp. 1–4.

[17] ARM, *ARM Cortex-A15 MPCore Processor Technical Reference Manual*, ARM, June 2013, pages 53 - 63.

[18] A. Sinkar, H. Ghasemi, M. Schulte, U. Karpuzcu, and N. Kim, "Low-Cost Per-Core Voltage Domain Support for Power-Constrained High-Performance Processors," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 22, no. 4, pp. 747–758, 2013.

[19] H. Ghasemi, A. Sinkar, M. Schulte, and N. S. Kim, "Cost-effective power delivery to support per-core voltage domains for power-constrained processors," in *Design Automation Conference (DAC), 2012 49th ACM/EDAC/IEEE*, 2012, pp. 56–61.

[20] E. Rotem, A. Mendelson, R. Ginosar, and U. Weiser, "Multiple clock and Voltage Domains for chip multi processors," in *Microarchitecture, 2009. MICRO-42. 42nd Annual IEEE/ACM International Symposium on*, 2009, pp. 459–468. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5375435

[21] R. Bahar and S. Manne, "Power and energy reduction via pipeline balancing," in *Computer Architecture, 2001. Proceedings. 28th Annual International Symposium on*, 2001, pp. 218–229.

[22] J. Sartori, B. Ahrens, and R. Kumar, "Power balanced pipelines," in *High Performance Computer Architecture (HPCA), 2012 IEEE 18th International Symposium on*, 2012, pp. 1–12.

[23] R. Kumar, V. Zyuban, and D. Tullsen, "Interconnections in multi-core architectures: understanding mechanisms, overheads and scaling," in *Computer Architecture, 2005. ISCA '05. Proceedings. 32nd International Symposium on*, 2005, pp. 408–419.

[24] H. Zeng, J. Wang, G. Zhang, and W. Hu, "An interconnect-aware power efficient cache coherence protocol for CMPs," in *Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on*, 2008, pp. 1–11.

[25] R. Kumar, K. Farkas, N. Jouppi, P. Ranganathan, and D. Tullsen, "Single-ISA heterogeneous multi-core architectures: the potential for processor power reduction," in *Microarchitecture, 2003. MICRO-36. Proceedings. 36th Annual IEEE/ACM International Symposium on*, 2003, pp. 81–92.

[26] P. Greenhalgh, "big.LITTLE Processing with ARM Cortex-A15 & Cortex-A7," ARM, Tech. Rep., September 2011.

[27] H. M. Waidyasooriya, Y. Takei, M. Hariyama, and M. Kameyama, "FPGA implementation of heterogeneous multicore platform with SIMD/MIMD custom accelerators," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, 2012, pp. 1339–1342.

[28] A. Lukefahr, S. Padmanabha, R. Das, F. Sleiman, R. Dreslinski, T. Wenisch, and S. Mahlke, "Composite Cores: Pushing Heterogeneity Into a Core," in *Microarchitecture (MICRO), 2012 45th Annual IEEE/ACM International Symposium on*, 2012, pp. 317–328.

[29] B. Jeff, "Advances in big.LITTLE Technology for Power and Energy Savings," ARM, Tech. Rep., September 2012.

[30] S. Zhang and K. Chatha, "Automated techniques for energy efficient scheduling on homogeneous and heterogeneous chip multi-processor architectures," in *Design Automation Conference, 2008. ASPDAC 2008. Asia and South Pacific*, 2008, pp. 61–66.

[31] A. Z. Jooya and M. Analoui, "Program phase detection in heterogeneous multi-core processors," in *Computer Conference, 2009. CSICC 2009. 14th International CSI*, 2009, pp. 219–224.

[32] L. Sawalha and R. Barnes, "Energy-Efficient Phase-Aware Scheduling for Heterogeneous Multicore Processors," in *Green Technologies Conference, 2012 IEEE*, 2012, pp. 1–6.

[33] F. Calcado, S. Louise, V. David, and A. Merigot, "Efficient Use of Processing Cores on Heterogeneous Multicore Architecture," in *Complex, Intelligent and Software Intensive Systems, 2009. CISIS '09. International Conference on*, 2009, pp. 669–674.

[34] M. Pericas, A. Cristal, F. Cazorla, R. Gonzalez, D. Jimenez, and M. Valero, "A Flexible Heterogeneous Multi-Core Architecture," in *Parallel Architecture and Compilation Techniques, 2007. PACT 2007. 16th International Conference on*, 2007, pp. 13–24.

[35] M. Santambrogio, "From Reconfigurable Architectures to Self-Adaptive Autonomic Systems," in *Computational Science and Engineering, 2009. CSE '09. International Conference on*, vol. 2, 2009, pp. 926–931.

[36] J. Zalke and S. Pandey, "Dynamic Partial Reconfigurable Embedded System to Achieve Hardware Flexibility Using 8051 Based RTOS on Xilinx FPGA," in *Advances in Computing, Control, Telecommunication Technologies, 2009. ACT '09. International Conference on*, 2009, pp. 684–686.

[37] S. Bhandari, S. Subbaraman, S. Pujari, F. Cancare, F. Bruschi, M. Santambrogio, and P. Grassi, "High Speed Dynamic Partial Reconfiguration for Real Time Multimedia Signal Processing," in *Digital System Design (DSD), 2012 15th Euromicro Conference on*, 2012, pp. 319–326.

[38] S. Liu, R. Pittman, A. Forin, and J.-L. Gaudiot, "On energy efficiency of reconfigurable systems with run-time partial reconfiguration," in *Application-specific Systems Architectures and Processors (ASAP), 2010 21st IEEE International Conference on*, 2010, pp. 265–272.

[39] K. Changkyu, S. Sethumadhavan, M. S. Govindan, N. Ranganathan, D. Gulati, D. Burger, and S. Keckler, "Composable Lightweight Processors," in *Microarchitecture, 2007. MICRO 2007. 40th Annual IEEE/ACM International Symposium on*, 2007, pp. 381–394.

[40] E. Ipek, M. Kirman, N. Kirman, and J. F. Martinez, "Core fusion: accommodating software diversity in chip multiprocessors," in *Proceedings of the 34th annual international symposium on Computer architecture*, ser. ISCA '07. New York, NY, USA: ACM, 2007, pp. 186–197. [Online]. Available: http://doi.acm.org/10.1145/1250662.1250686

[41] Z. Rakossy, T. Naphade, and A. Chattopadhyay, "Design and analysis of layered coarse-grained reconfigurable architecture," in *Reconfigurable Computing and FPGAs (ReConFig), 2012 International Conference on*, 2012, pp. 1–6.

[42] K. Changmoo, C. Mookyoung, C. Yeongon, M. Konijnenburg, R. Soojung, and K. Jeongwook, "ULP-SRP: Ultra low power Samsung Reconfigurable Processor for biomedical applications," in *Field-Programmable Technology (FPT), 2012 International Conference on*, 2012, pp. 329–334.

[43] F. J. Veredas, M. Scheppler, W. Moffat, and B. Mei, "Custom implementation of the coarse-grained reconfigurable ADRES architecture for multimedia purposes," in *Field Programmable Logic and Applications, 2005. International Conference on*, 2005, pp. 106–111.

[44] G. Liu, "Fpga implementation of 2d-dct/idct algorithm using multi-core picomips," Master's thesis, University of Southampton, School of Electronics and Computer Science, September 2013.

SPICE Modeling and Simulation of a MPPT Algorithm

Miona Andrejević Stošović, Marko Dimitrijević, Duško Lukač, and Vančo Litovski

***Abstract*—One among several equally important subsystems of a standalone photovoltaic (PV) system is the circuit for maximum power point tracking (MPPT). There are several algorithms that may be used for it. In this paper we choose such an algorithm based on the maximum simplicity criteria. Then we make some small modifications to it in order to make it more robust. We synthesize a circuit built out of elements from the list of elements recognized by SPICE. The inputs are the voltage and the current at the PV panel to DC-DC converter interface. Its task is to generate a pulse width modulated pulse train whose duty ratio is defined to keep the input impedance of the DC-DC converter at the optimal value.**

***Index Terms*—Modeling; Simulation; PV systems; MPPT algorithm.**

Original Research Paper
DOI: 10.7251/ELS1418011A

I. INTRODUCTION

A TYPICAL standalone PV system is depicted in Fig. 1. As seen from the figure, it contains the following main subsystems: the PV panel, the line capacitor C_L, the DC to DC converter, the maximum power point tracking (MPPT) subsystem, the battery, and the load which is here represented as a resistor R_L.

Each part of the system may be realized in different versions. For example there are several circuit architectures for the DC to DC converter; there are different technologies implemented for production of the batteries and, most frequently, the load is a specific electronic system, e.g. metrological measurement station, working remotely.

To design the complete system one has to have circuit models of all subsystems and that is not the case. Namely, to our knowledge there are no published circuit simulations of the whole system. The published simulations are most frequently behavioural using Matlab-Simulink [1,2] which is successful when looking at the system level but makes it difficult to implement models of real components due to lack of proper libraries.

Circuit simulation with SPICE [3] and SPICE-like software enables easier data transfer to a PCB layout design tool so making the design process continuous, reducing the format translation activities, and minimizing the risk of error during manipulation of data.

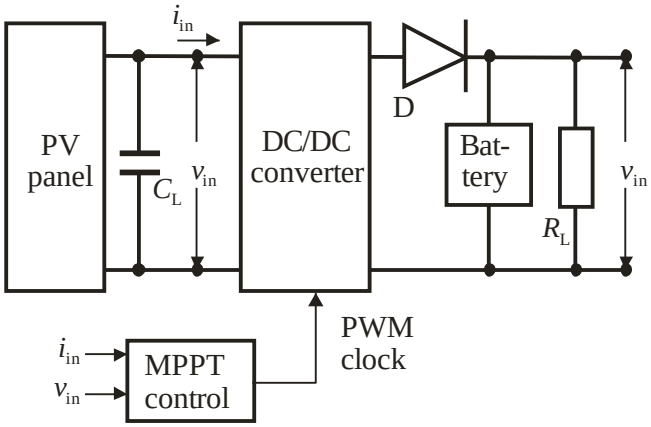


Fig. 1. System level schematic of a standalone PV system.

To our knowledge most of the parts of the PV system depicted in Fig. 1. are already modeled and simulated in SPICE e.g. [4,5,6]. However, there are no published results reporting SPICE modeling and simulation of the MPPT circuit. By creation of such a model, we expect, one will be capable to simulate and design the whole standalone system, hence the importance of this work.

The paper is organized as follows. First we will describe the basic properties of the PV panel from the sensitivity to illumination and temperature point of view in order to establish the feeling about the reason why the maximum power point is migrating during the operation of the PV system. Then, we will describe the most frequently used algorithm for MPPT named Perturb and Observe (P&O), as described in [7,8,9]. A minor improvement of the way how the algorithm is expressed will be introduced leading to a Modified Perturb and Observe (MP&O) algorithm. Finally, the SPICE model of the MPPT and simulation results verifying the model reported will be introduced.

II. MPPT ALGORITHM

Fig. 2. depicts the dependence of the photovoltaic power produced by a PV panel on the voltage on it with the illumination as a parameter [7]. As it may be seen the MPP

Manuscript received 13 April 2014. Accepted for publication 30 May 2014.

This research was partly funded by The Ministry of Education and Science of Republic of Serbia under contract No. TR32004.

Miona Andrejević Stošović and Marko Dimitrijević are with the University of Niš, Faculty of Electronic Engineering, 14 Aleksandra Medvedeva, 18000 Niš, Serbia (e-mail: miona.andrejevic@elfak.ni.ac.rs, marko.dimitrijevic@elfak.ni.ac.rs).

Duško Lukač is with the Rheinische Fachhochschule Köln, Germany (e-mail: lukac@rfh-koeln.de).

Vančo Litovski is with NiCAT - Niš Cluster of Advanced Technologies, 18000 Niš, Serbia (e-mail: vanco.litovski@elfak.ni.ac.rs).

migrates slightly due to the change of the output resistance of the panel [10].

As a counterpart to this dependence the migration of the MPP with temperature is shown in Fig. 3. [7]. Due to these reasons every PV system is equipped with a specific circuitry which controls the duty cycle of the pulse train controlling the switches within the converter (inverter). These changes lead to changes of the input resistance of the converter and if the control is properly tailored the PV panel will be kept in a position to deliver maximum power to the converter.

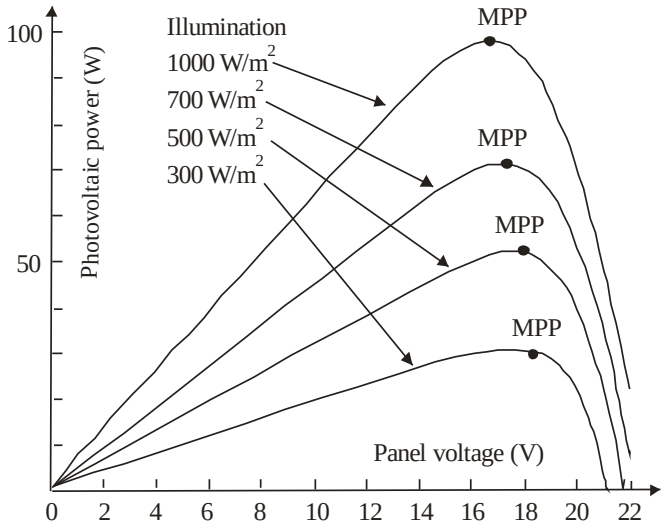


Fig. 2. Photovoltaic power versus panel voltage for different illumination intensities.

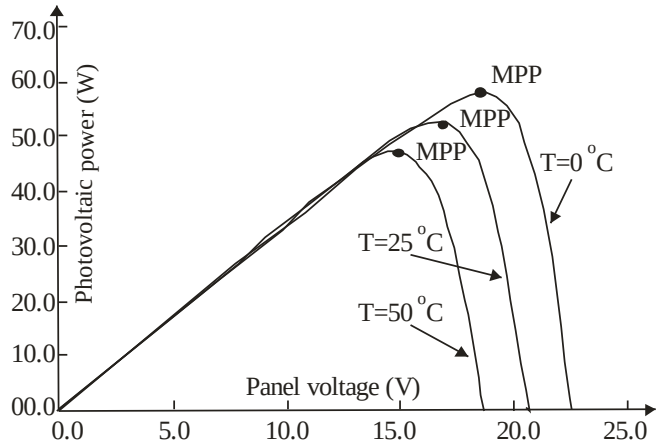


Fig. 3. Photovoltaic power versus cell voltage for different temperatures.

As already mentioned in Introduction there are several techniques and algorithms enabling implementation of the idea of tracking the MPP. Among them the most popular is the so called Perturb and Observe (P&O) which is depicted in Fig. 4. [8,9]. One may see from the figure that as a result a fixed increment (positive or negative) of the PV panel voltage is produced after each sampling period (of the voltage and the power of the PV panel). In real circuit this increment is used as information for a pulse width modulated (PWM) oscillator to control its duty ratio.

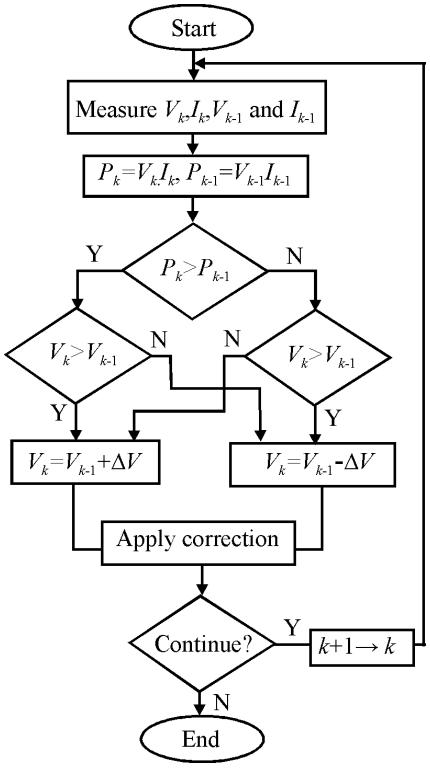


Fig. 4. The perturb and observe algorithm.

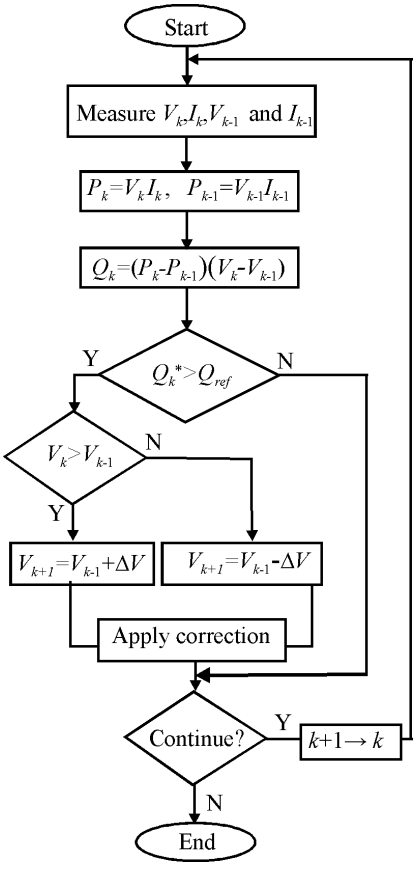


Fig. 5. The modified perturb and observe algorithm.

There are four branches in this algorithm that lead to two signs of the increment. If (P_k-P_{k-1}>0 and V_k-V_{k-1}>0) and if

(P_k-P_{k-1}<0 and V_k-V_{k-1}<0) the increment is positive. The value of the increment is negative if (P_k-P_{k-1}<0 and V_k-V_{k-1}>0) and if (P_k-P_{k-1}>0 and V_k-V_{k-1}<0). By careful inspection one may conclude that the sign of the increment is associated to the sign of Q_k=(P_k-P_{k-1})(V_k-V_{k-1}). If Q_k>0 the increment is positive and if Q_k<0 the increment is negative. This way of expression simplifies the whole diagram and probably the electronic circuitry implementing the algorithm.

Note that the algorithm is generating an increment no matter how large the differences (P_k-P_{k-1}) and (V_k-V_{k-1}) are, which, generally speaking, for small differences may lead to a duty cycle becoming much larger than necessary which may even be counterproductive i.e. it may leat to lead the quiescent point out of MPP region by "overshooting" it.

Having all that in mind we propose a modification of the algorithm as depicted in Fig. 5. In this algorithm, after calculation of Q_k one first makes a comparison with a threshold value Q_{ref}. If $|Q_k|$ is not larger than Q_{ref} there is no need for perturbation i.e. for a change of the duty cycle. If it is larger, only the sign of Q_k is used to create the sign of the voltage increment.

Note that the implementation of this idea may lead to an additional benefit. Namely, by proper choice of Q_{ref} and ΔV, larger increments (ΔV) may be implemented since overshooting is disabled. In other words one may expect faster recovery of the MPPT.

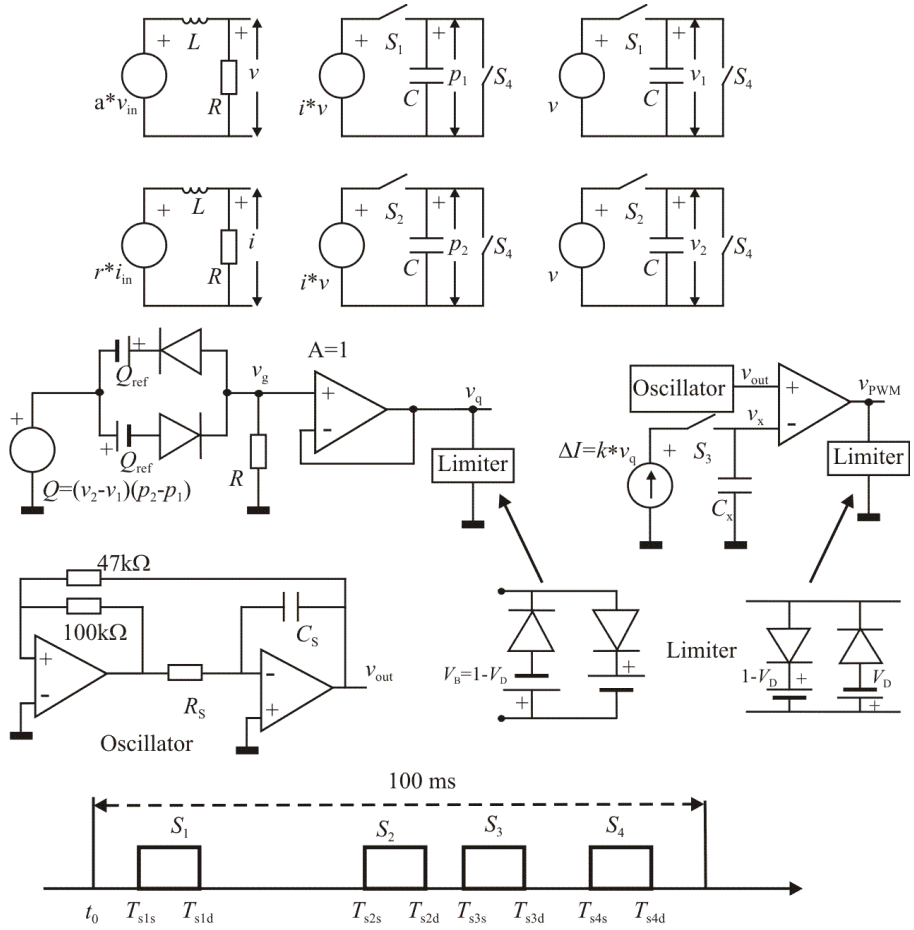


Fig. 6. The SPICE schematic modelling the MPPT control circuit.

III. SPICE MODEL OF THE MPPT ALGORITHM AND SIMULATION RESULTS

The SPICE implementation of the algorithm in Fig. 5. is depicted in Fig. 6. Further we shortly describe its units.

First, since the measured output voltage (v) and current (i) of the PV panel contain components that may be constant, slowly time-varying (due to change of temperature, illumination and the load) and very fast (due to the commutation within the converter), we implement a filtering function to eliminate the high frequency component. That is done by a L-R low-pass filter as shown at the top left part of the figure.

Next, we sample both the voltage and the power. It is done by a sample and hold circuit that starts at T_{s1s} and ends at T_{s1d}. The sampling circuitry is depicted in the top part of the figure (associated with S_1) while the timing diagram is given at the bottom. After this operation the capacitors in these two circuits memorize the sampled values of p_1 and v_1.

The next sampling takes place after a relatively long period

which is adjustable based on how fast changes of temperature and illuminations are expected. Here the sampling instant (associated with S_2) is denoted by T_{s2s}. So after T_{s2d} we have two new samples: p_2 and v_2.

Having these four quantities we compute $Q=(p_2-p_1)(v_2-v_1)$ which is represented in Fig. 6. as a controlled voltage source. That signal is contrasted to Q_{ref} and the result is brought to a limiter so that v_q takes (approximately) values 0, 1V and -1V, only.

The resulting v_q is used to create the voltage increment at C_x. It is $\Delta v_x=k\cdot v_q\cdot\Delta t/C_x$, where $\Delta t=T_{s3d}-T_{s3s}$. This increment is brought to one of the inputs of a comparator. The second input of the comparator is excited by the output (voltage) signal of an oscillator that produces symmetrical alternatively linearly rising and linearly falling signal. The oscillation period is equal to the switching period of the converter so that at the output of the limiter connected to the comparator`s output one gets pulse width modulated signal.

After completion of the sampling at S_3, using S_4 the sampled values of v_1, v_2, p_1, and p_2, are erased and a new measurement and control cycle is enabled. Here, to shorten the computer time, a measurement/control period of 100 ms was used. In real world that interval is supposed to be longer.

The following set of parameter values was used to enable the simulation of the circuit in Fig. 6: a=1, r=1Ω, C=1μF, L=1H, R=1kΩ, Q_{ref}=0.5V, R_s=10kΩ, C_s=0.9nF, k=10^{-3}, C_x=14.3mF, T_{s1s}=5ms, T_{s1d}=10ms, T_{s2s}=70ms, T_{s2d}=75ms, T_{s3s}=85ms, T_{s3d}=90ms, T_{s4s}=95ms, T_{s4d}=100ms. No battery was included in the system.

To illustrate the implementation of the algorithm an example will be shown. Instead of getting the signals directly from the PV panel, for verification purposes, we created two signals as follows (Figs. 7 and 8):

$$v(t)=2+\sin(2\pi\cdot 50000\cdot t)+\sin(2\pi\cdot 0.5\cdot t)\ [\mathrm{V}] \quad (1a)$$

$$i(t)=0.6+\sin(2\pi\cdot 50000\cdot t)+\sin(2\pi\cdot 0.5\cdot t+\pi/2)\ [\mathrm{A}]. \quad (1b)$$

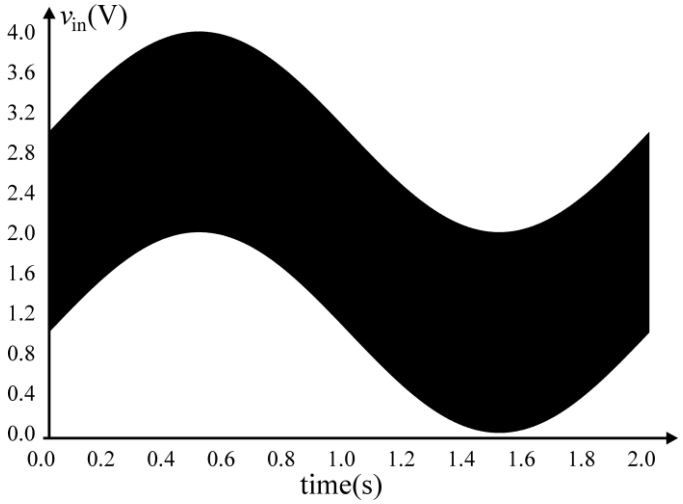


Fig. 7. Input voltage signal.

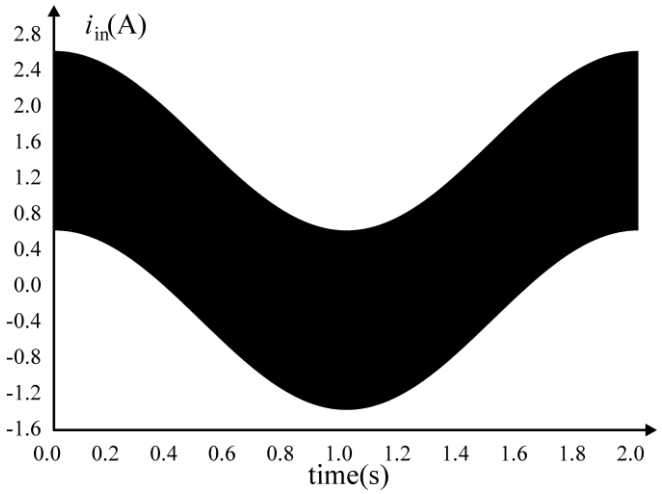


Fig. 8. Input current signal.

These were used as excitation for the circuit in Fig. 6. Fig. 9. depicts the signal Q of Fig. 6. Finally, Fig. 10 illustrates the dependence of the duty cycle at the output of the circuit in Fig. 6. It is clear that it follows the shape of Q. It also implements the new version of the MPPT algorithm presented in Fig. 5.

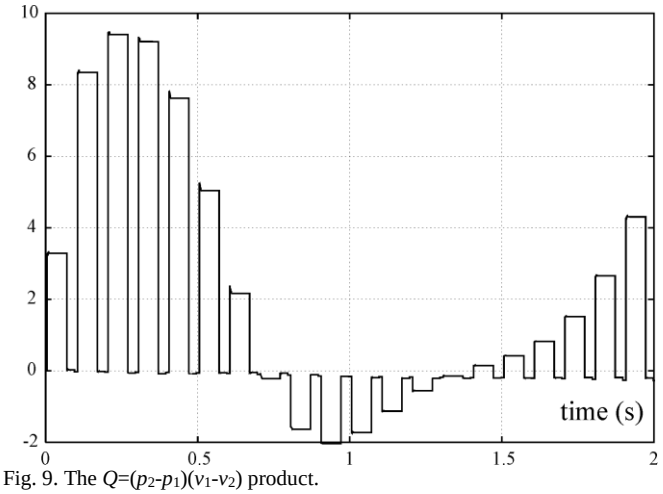


Fig. 9. The $Q=(p_2-p_1)(v_1-v_2)$ product.

In Figs. 11 and 12 pulses at the circuit output are given for two different time instances. We can see that pulse width dramatically changes in time, which was in fact the goal.

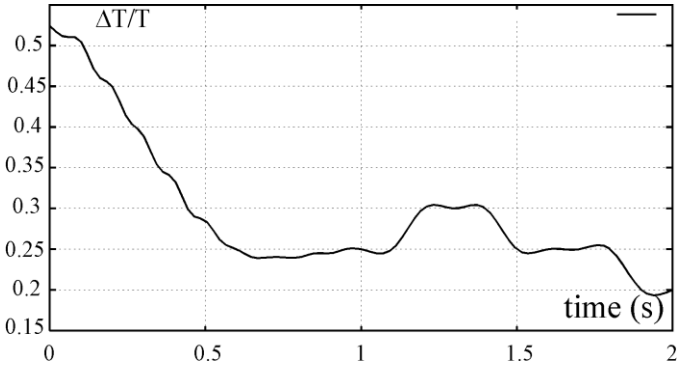


Fig. 10. Dependence of the duty cycle at the output of the circuit of Fig. 6.

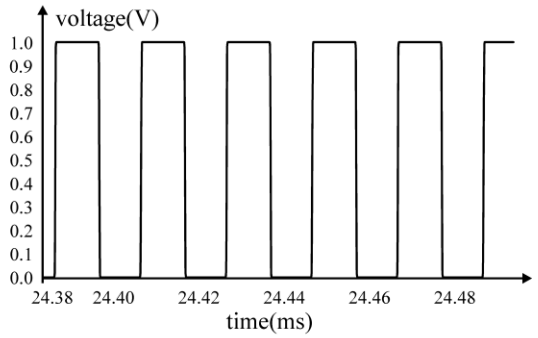


Fig. 11. Output of the circuit of Fig. 6- case 1.

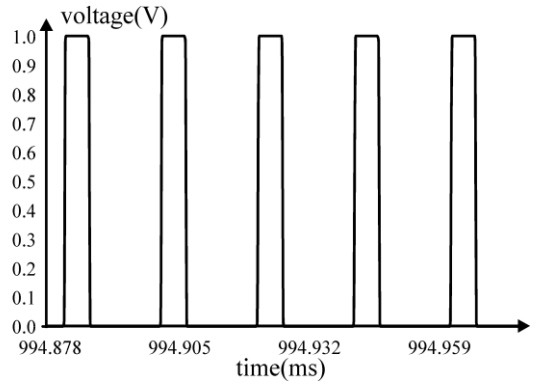


Fig. 12. Output of the circuit of Fig. 6- case 2.

CONCLUSION

An improvement was proposed to the existing MPPT algorithm since a drawback was noticed in it. To verify the new idea SPICE simulation of the whole system containing the PV panel, the line capacitor, the converter, and a resistive load, was performed. A specific contribution of this paper is an original SPICE model of the subsystem performing the MPPT algorithm. Future implementations of these achievements are expected in the simulation of the system performance under dynamic changes of the load extended with a super- capacitor charging system.

REFERENCES

[1] A. D. Hansen, P. Sorensen, L. H. Hansen, and H. Bindner, "Models for a stand-alone PV system," Riso, National Laboratories, Roskilde, Denmark, 2000, http://orbit.dtu.dk/fedora/objects/orbit:91233/datastreams/file_7727175/content.

[2] B. Sree Manju, R. Ramaprabha R., B. L. Mathur, "Design and Modeling of Standalone Solar Photovoltaic Charging System", *Int. J. of Computer Applications*, vol. 18, no. 2, pp.41-45, March 2011.

[3] L. W. Nagel, D. O. Pederson, SPICE (Simulation Program with Integrated Circuit Emphasis), University of California, Berkeley, Memorandum ERL-M 382, April 1973.

[4] V. Litovski, M. Savić, Ž. Mrčarica, "Electronic circuit simulation with ideal switches," *HAIT Journal of Science and Engineering B*, pp. 476-495, July 2005.

[5] E. W. Enlow, D. R. Alexander, "Photocurrent modeling of modern microcircuit pn junctions," *IEEE Trans. on Nuclear Science*, vol. 35, no. 6, pp. 1467–1474, Dec 1988.

[6] K. H. Edelmoser, F. A. Himmelstoss, "Bi-directional DC-to-DC Converter for Solar Battery Backup Applications," *35th Annual IEEE Power Electronics Specialists Conf.*, Aachen, Germany, pp. 2070-74, 2004.

[7] V. Salas, E. Olías, A. Barrado, A. Lázaro, "Review of the maximum power point tracking algorithms for stand-alone photovoltaic systems," *Solar Energy Materials & Solar Cells*, no. 90, pp. 1555–78, 2006.

[8] T. Esram, and P. L. Chapman, "Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques," *IEEE Transactions on Energy Conversion*, vol. 22, no. 2, pp. 439-449, June 2007.

[9] C. S. Chin, M. K. Tan, P. Neelakantan, B. L. Chua, and K. T. K. Teo, "Optimization of Partially Shaded PV Array using Fuzzy MPPT", *IEEE Colloquium on Humanities, Science and Engineering Research (CHUSER 2011)*, Penang, China, pp. 481–486, Dec. 2011.

[10] M. Andrejević Stošović, M. Dimitrijević, D. Lukač, and V. Litovski, "Quantification of Power Quality Issues at the PV Panel-Converter Interface," *Proc. of the IX Symp. on Industrial Electronics, INDEL 2012*, Banja Luka, B&H, pp. 256-262, Nov. 2012.

The use of Bayesian Networks in Detecting the States of Ventilation Mills in Power Plants

Sanja Vujnović, Predrag Todorov, Željko Đurović, and Aleksandra Marjanović

***Abstract*—The main objective of this paper is to present a new method of predictive maintenance which can detect the states of coal grinding mills in thermal power plants using Bayesian networks. Several possible structures of Bayesian networks are proposed for solving this problem and one of them is implemented and tested on an actual system. This method uses acoustic signals and statistical signal pre-processing tools to compute the inputs of the Bayesian network. After that the network is trained and tested using signals measured in the vicinity of the mill in the period of 2 months. The goal of this algorithm is to increase the efficiency of the coal grinding process and reduce the maintenance cost by eliminating the unnecessary maintenance checks of the system**

***Index Terms*—Acoustic Signals, Bayesian Networks, Predictive Maintenance, Ventilation Mills.**

Original Research Paper
DOI: 10.7251/ELS1418016V

I. Introduction

Condition based maintenance (CBM) is a maintenance technique widely used in the industry today. It uses data collected through condition monitoring to advise whether maintenance is necessary, and thus reduces maintenance cost of the system [1]. Many of the maintenance techniques that are currently implemented are using a certain type of time-based preventive maintenance, where the components of the system are checked regularly after a certain period of time to ensure that no serious fault has occurred. This kind of maintenance is being conducted on ventilation mills in thermal power plants where, due to the inability to predict the state of the coal grinding plate within the mill, the entire subsystem needs to be periodically stopped and inspected. This paper proposes a new method of predictive maintenance which can detect the state of the plates within the mills based on the acoustic signals measured in the vicinity of the mill, using Bayesian networks (BN).

Bayesian networks have been widely used for the purpose of predictive maintenance in the last few years. Since their introduction in the mid 80's [2] they have experienced a rapid development in many areas of research. Also, due to their ability to represent complex systems in which high amount of uncertainty exists, they have proven themselves to be superior to other methods such as neural networks, support vector machines, etc. In this paper several BN structures will be proposed which serve to model the states and the behaviors of ventilation mills, and one of them will be tested on actual acoustic measurements.

There are many examples of fault detection algorithms which are based on the measurements of vibration signals of the machines ([3] and [4], to name a few). However, it has long since been shown that acoustic measurements can be as informative as vibration signals when it comes to detecting the faults in machines [5]. The usage of acoustic signals in fault detection has expanded during the last decade [6], however they are still considered to be a lesser alternative to vibration signals, due to their high susceptibility to surrounding noise. In this paper acoustic signals acquired in a very noisy environment will be used to test the efficiency of a simple Bayesian network for the detection of states in the ventilation mills. More complex realizations of the solution to the problem will be proposed as well.

This paper is structured as follows. In Chapter 2 a short introduction to Bayesian Networks is presented, while in Chapter 3 some practical realizations of BNs which can be used for fault detection are proposed. Chapter 4 describes the system on which the algorithms have been tested, as well as the process of acquiring acoustic signals. In Chapter 5 the results of the algorithm are presented and in Chapter 6 the final conclusions to this paper are stated.

II. Bayesian Belief Networks

Bayesian Belief Network (BBN) is a term introduced in 1980s by Jude Pearl [2] in an attempt to create a mathematical probabilistic tool capable of modeling and reproducing the process of human reasoning. The basic idea was to reproduce the way in which people accumulate information from different sources and use it to develop conclusions about certain ideas. During the last decade Bayesian networks have become a very powerful tool for representation of complex

Manuscript received 13 April 2014. Accepted for publication 30 May 2014.

S. Vujnović is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (phone: +381-11-3218-370; e-mail: svujnovic@etf.bg.ac.rs).

P. Todorov is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: pedjatodorov@etf.bg.ac.rs).

Ž. Đurović is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: zdjurovic@etf.bg.ac.rs).

A. Marjanović is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: amarjanovic@etf.bg.ac.rs).

systems and it has been used in many areas of research including fault detection and fault isolation [7], [8].

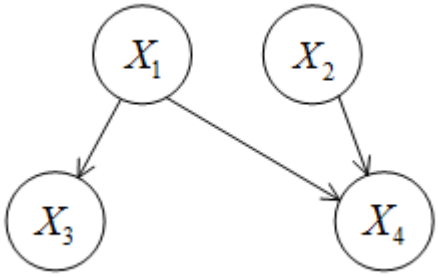


Fig. 1. A simple Bayesian network which consists of 4 nodes (X_1, X_2, X_3 and X_4) and three connections between them.

A. Bayesian Theorem

Bayesian networks apply Bayesian theorem on complex systems to calculate conditional probabilities of certain events or hypothesis when a new evidence is obtained. Bayesian theorem was coined by Thomas Bayes in the 18th century and it can be formulated as

$$P(H_n \mid E) = \frac{P(H_n)P(E \mid H_n)}{P(E)}. \quad (1)$$

Here $P(H_n \mid E)$ is a posterior probability of hypothesis H_n when evidence E is known. $P(H_n)$ and $P(E)$ are prior probabilities of hypothesis and evidence, respectively, and $P(E \mid H_n)$ is the probability of evidence E given the hypothesis H_n. The great advantage of Bayes' theorem is that it allows us to easily calculate conditional probabilities based on the corresponding prior probabilities. In this case, if H_n is, for example, a certain fault, and E is evidence or a symptom of a fault, than $P(H_n)$ and $P(E \mid H_n)$ can be more easily obtained from the survey or maintenance data, than the conditional probability of a fault given the evidence.

If we assume that i denotes a specific hypothesis H_i, then (1) can be rewritten using the rule of total probability, where the summation is taken over all hypotheses H_i which are mutually exclusive and their prior probabilities sum up to 1. The final form is given as

$$P(H_n \mid E) = \frac{P(H_n)P(E \mid H_n)}{\sum_i P(H_i)P(E \mid H_i)}. \quad (2)$$

This is called inference and it represents the basic idea of how the inserted evidence spreads throughout Bayesian networks.

B. Topology of Bayesian Network

Bayesian network is an acyclic probabilistic graph which contains a set of nodes and directed connections between them and it is used to represent the knowledge about uncertain events. Nods represent the probabilistic variables which can be continuous or discrete. Connections between nodes represent the probabilistic dependencies among certain variables. A simple type of Bayesian network is shown in Fig. 1. Here an edge from node X_1 to node X_3 and from nodes X_1 and X_2 to node X_4 represent a statistical dependence between the corresponding variables. Therefore a value taken by a variable X_4 depends on the values taken by variables X_1 and X_2. Nodes X_1 and X_2 are then referred to as the parents of X_4 and, similarly, X_4 is referred to as the child of X_1 and X_2. In general, for the network with n nodes: X_1, X_2,..., X_n, the joint probability can be expressed as

$$P(X_1, X_2, ..., X_n) = \sum_{i=1}^{n} P(X_i \mid pa(X_i)), \quad (3)$$

where $pa(X_i)$ is a set of parents of the node X_i.

There are many types of nodes which can be chosen for any given Bayesian network. However, in practice only two types of variables are used: discrete and continuous Gaussian. This is because only for these two types of nodes can the exact computation of conditional probabilities be done [9]. Similarly, the discrete variables cannot have continuous parents if the exact computation is required. Therefore, there are three different combinations of nodes for the parent-child relationship and three different methods for calculating conditional probabilities. If discrete variables have discrete parents, their conditional probabilities are expressed via conditional probability table. For example, if both X_1 and X_3 from Fig. 1 are discrete variables, with n and m different possible states respectively, their conditional probability table would have n x m entries of conditional probabilities $P(X_3 = x_{3i} \mid X_1 = x_{1k})$, where x_{3i} is the i-th state of node X_3 and x_{1k} the k-th state of the node X_1. On the other hand, if X_3 was a Gaussian continuous variable and X_1 its discrete parent, the conditional probability distribution would be Gaussian as well [9].

C. The Learning Process

There are two ways of calculating parameters of Bayesian network, as well as the structure. One is based on an expert knowledge of the system and the other on machine learning using experimental results. These two approaches can be used jointly or individually.

Assuming that the expert knowledge of the system is unavailable and that Bayesian network needs to be taught its structure and parameters, there are several ways in which this can be achieved. Structure learning itself is much more complicated than the parameter learning and we will not dwell on it further in this paper. For all intents and purposes we will assume that the structure of the Bayesian network is known and that only the parameters need to be taught.

In the case when all the nodes (variables) of the system are observable, the most common algorithm for machine learning of the Bayesian network is the *log likelihood* approach. Here

the goal is to learn all of the conditional probability densities which maximize the likelihood of the training data. That is to

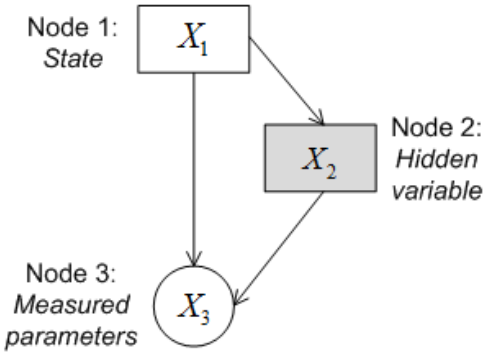


Fig. 2. Bayesian network with one measured continuous node (X_3), one hidden discrete node (X_2) and one discrete node which represents the state of impact plates and which cannot be directly measured (X_1).

say that if we have Bayesian network $B = (S, p)$ with defined structure S and parameters p, and where D is a set of training data with the values of all the parameters p of network B, then the goal is to maximize the likelihood, i.e. the probability of the data d given the Bayesian network B:

$$L = \prod_{d \in D} P(d \mid B), \tag{4}$$

or as it is more commonly expressed in logarithmic form:

$$LL = \sum_{d \in D} \log_2 P(d \mid B). \tag{5}$$

If by any chance there is a choice of several Bayesian network models, it is wise to adopt the model which gives the maximum likelihood given the data obtained.

Usually however, not all the parameters can be measured, and some connections between the nodes must be modeled with unobservable variables. In such cases the exact log likelihood approach cannot be used and the approximate parameter estimation must be conducted. The most popular approximate algorithm for teaching the Bayesian network with hidden nodes is the *Expectations maximization* (EM) algorithm. This algorithm consists of two steps. In the first step, called the *expectation* step, the current parameter estimations $\hat{p}$ are used to calculate expectations for the missing values of hidden variables. Then, in the second step called the *maximization* step, a new maximum likelihood estimate of the parameters is calculated after which the algorithm goes back to the previous step. This continues throughout the entire training set.

III. Application of Bayesian Network in Predictive Maintenance

The fundamental problem with Bayesian networks is deciding which model adequately represents the system which is observed. First of all, the variables used in the system need to be chosen. They can represent some physical, measurable property of the system or some unknown interference which

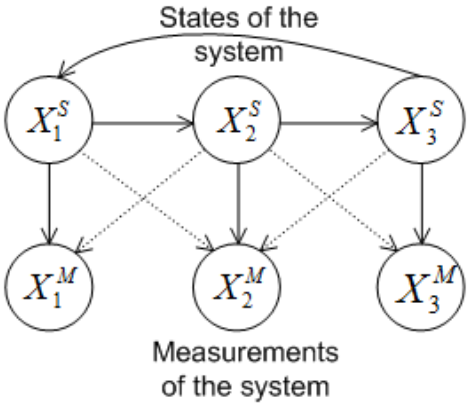


Fig. 3. Hidden Markov model represented with Bayesian networks. X_i^S are the states of the system and X_i^M are the measurements of the system. All the variables are continuous Gaussian and visible.

cannot be measured. Then, the decision needs to be made concerning the type of the chosen variables. They can be discrete or continuous, visible or hidden. Secondly, the question of the structure of the Bayesian network should be decided upon. This is the most important and the most complex question. The types of structures which can be represented with Bayesian networks are vast, and the only limit one faces when choosing them is the knowledge of the causal relationships within the physical system and the complexity sufficient to solve the given problem. In this chapter several structures will be proposed for solving the problem of detecting the states of ventilation mills.

A. The Simple Static Model

The simplest Bayesian network that can be used for detecting the state of the mill is shown in Fig. 1. Circles represent continuous-valued random variables, squares are discrete random variables, clear boxes represent observed nodes and gray boxes represent hidden nodes. This network contains only 3 nodes (variables), two of which are discrete and one of which is continuous Gaussian.

The basic idea is that the measurements taken outside of the mill are directly influenced by the state of the impact plates within the mill and by a hidden variable which is modeled as a discrete node. This hidden, immeasurable node can represent an outside noise, unknown component of the system or some other feature which does not necessarily need to have a physical interpretation. The state of the plates X_1 is modeled as a discrete node with 2 states: healthy plates and worn plates. The measured signal which is represented by the node X_3 is a two-dimensional Gaussian variable.

B. Hidden Markov Model

Hidden Markov model (HMM) can be modeled as a simple dynamic Bayesian network, where the next state of the system depends on the previous state and the state transition probabilities, and where the states themselves are not directly measured. However, even though the states cannot be

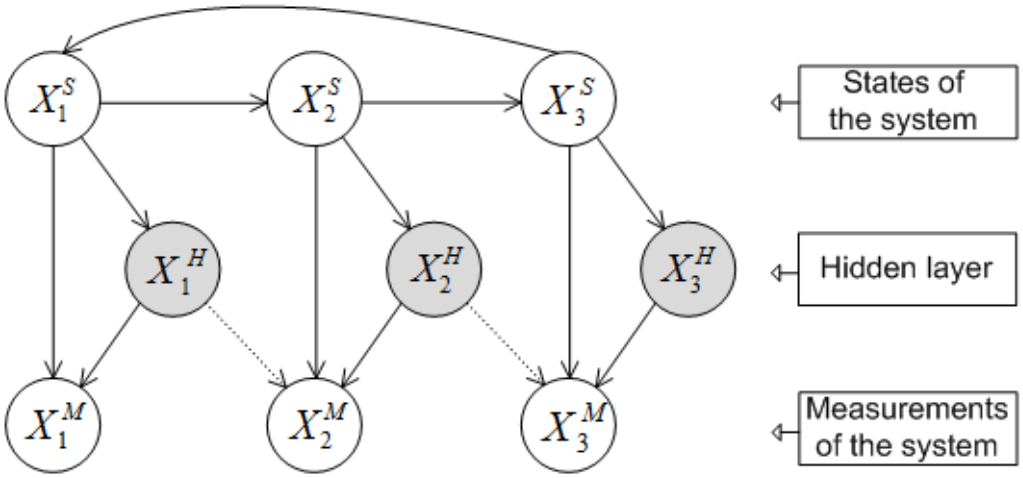


Fig. 4. Complex dynamical BN model. X_i^S are the states of the system, X_i^M are the measurements of the system and X_i^H. are the hidden variables of the system which cannot be observed or detected. All the nodes are continuous and Gaussian.

measured, they influence the measurements of the system which are then used to infer the state in which the system is in. Hidden Markov models are used for a very broad range of systems and one such model (Fig. 3) can be used to infer the states of the impact plates in ventilation mills.

In Fig. 3, X_i^S represents a state node and X_i^M a measurement node. All of the variables are chosen to be continuous and visible (that is to say, information on states and measurements can be used when training the network). In this case 3 continuous states are chosen: the state of healthy plates (X_1^S), the state when the plates are used but not damaged enough, so the repair is not necessary (X_2^S) and the state which represents worn plates (X_3^S) and which indicates that the repair should be done as soon as possible.

Note that the transition between states is modeled so that the state deteriorates up until a point when the plates have to be replaced, and only then can the state return to the initial value of healthy plates. Also, each state influences the appropriate measurement of the system, but due to noise and inconsistencies it can appear as a measurement of the adjacent state as well (presented by dotted lines in Fig. 3).

A. The Complex Dynamical Model

As we have noted, HMM are broadly used to model all sorts of systems in many areas of research. However, they represent only a simple usage of dynamic Bayesian network which can easily be modified to include more complex demands of the system. One such modification is proposed in Fig 4, where the mixture of first two solutions to Bayesian network models is implemented. Here the one-way transition between the states is included as well as the influence of the unknown, hidden variable on the measurements of the system.

IV. Case Study

In the previous chapter some possible structures of Bayesian networks were proposed that can be used for predictive maintenance. The actual process on which one of these solutions will be tested is a ventilation mill in thermal power plant Kostolac A (Serbia) and it will be described in this chapter. Also, the process of acquiring and pre-processing of acoustic signals used in simulations will be depicted.

A. Ventilation Mill

Coal grinding mills are a very important subsystem of thermal power plants. Their main purpose is to pulverize the coal into fine powder before it can get into the furnace. The mill itself consists of coal grinding plates (also called impact plates) which rotate within it with the frequency of f_r=12.5Hz and are used to crush the coal into smaller pieces until it gets sufficiently fragmented. After that it drifts into the furnace system where it is used as a fuel.

During the grinding process the impact plates get slowly depleted and their efficiency reduces. In the early stages this causes them to pulverize the coal more slowly, but if the plates are not changed for a long period of time it can cause a complete malfunction of the milling subsystem in thermal power plants. Some algorithms for solving this problem have already been proposed [10]. Currently this problem is being solved by employing a type of time based maintenance, which implies stopping the entire subsystem once every 8 or 9 weeks, conducting visual inspection of the plates and, if it is deemed necessary, replacement of the plates with the new ones. This can be very costly if the overhaul was concluded not to be necessary.

In this paper a new algorithm will be tested which uses acoustic signals recorded in the vicinity of the mill, while the mill is operating mode, and Bayesian networks to determine the state of the impact plates, i.e. whether the replacement is necessary or not.

B. Acoustic Signals

Acoustic signals were recorded in the vicinity of the mill while it was in operating mode. The measurements were conducted every couple of weeks in the period of two months, thus encompassing the entire life cycle of the plates, from the moment they were replaced to the moment when they were completely depleted. These signals were recorded with the

sampling frequency of f_s=48kHz and 24bit encryption. Approximately once every two weeks an acoustic signal has been measured, which has afterward been divided into 20-30 shorter signals.

For the training and testing of the Bayesian network the sampling period was downsized to 4.8kHz, Also, certain features were extracted from the signal which were later used in the decision making process. These features are the basic properties of the signal in time and frequency domain. Eight features in time domain are chosen as characteristic parameters of the signal [3]:

- Arithmetic mean value: $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x(i)$
- Root mean square value: $x_{rms} = \left(\frac{1}{N}\sum_{i=1}^{N} x^2(i)\right)^{1/2}$
- Square mean root value: $x_r = \left(\frac{1}{N}\sum_{i=1}^{N} |x(i)|^{1/2}\right)^2$
- Skewness index: $x_{ske} = \left(\frac{1}{N}\sum_{i=1}^{N} (x(i)-\bar{x})^3\right) / x_v^{3/2}$
- Kurtosis index: $x_{kur} = \left(\frac{1}{N}\sum_{i=1}^{N} (x(i)-\bar{x})4\right) / x_v^2$
- C factor: $C = PP / x_{rms}$
- L factor: $L = PP / x_r$
- S factor: $S = x_{rms} / \bar{x}$

where x_v is the variance of the signal x and PP is peak-to-peak value of the signal. Also, eight features in the frequency domain were chosen as the most dominant components in the frequency spectrum of the signal, in the same way it has been done in [10]. They represent the maximum value of the amplitude frequency spectrum of the signal at the frequencies:

$$f = [25 \quad 37.5 \quad 50 \quad 62.5 \quad 125 \quad 250 \quad 375 \quad 500]\,\text{Hz}.$$

After extracting these features a vector of 16 elements is acquired for each of the measured signals. To reduce the dimension of this vector and therefore reduce the amount of variables needed for the Bayesian network, a linear dimension reduction algorithm is applied which transforms a 16-dimentional vector X into a two-dimentional vector Y [11]:

$$Y_{2x1} = A^T_{16x2} X_{2x1}\,. \tag{6}$$

The matrix A is chosen so to minimize the loss of information by maximizing the distance of two extreme classes - the class of healthy and the class of worn plates. This matrix is acquired by maximizing the criteria $J = tr(S_W^{-1} S_B)$ where S_W is a within-class scatter matrix and S_B is a between-class scatter matrix, as defined in [11], and applied in [3]. After this procedure has been implemented, the 2-dimensional vector of features Y is used as an input to a Bayesian network.

V. Results

For the testing of this algorithm the Bayesian Network depicted in Fig. 2 is used. The first variable (X_1) represents the true state of the system and is chosen to be a discrete variable with two possible states - healthy and worn plates. The second variable (X_2) is a hidden variable and it does not have a specific physical meaning. It represents the noise in the surroundings that can influence the measurements, as well as unknown parameters of the system that were not otherwise included in the model. It is also chosen to be a discrete variable with two possible states. Finally, the third variable (X_3) represents the measured output of the system, or in this case, a parameter Y gained by signal pre-processing techniques described in the previous chapter. This variable is chosen to be continuous and Gaussian. With this in mind, the joint probability of this Bayesian network is:

$$P(X_1, X_2, X_3) = P(X_3 \mid X_1, X_2) P(X_2 \mid X_1) P(X_1)\,. \tag{7}$$

The idea is to calculate the probability of each of the states in X_1 given the measurements X_3.

Learning and testing procedures for this Bayesian network are done in a programming package Matlab, using toolbox for Bayesian networks created by [9]. For a training set, the signals taken slightly before and slightly after the change of plates were chosen, representing the two extreme states of the mill. Since there is a hidden node within the system, the EM learning algorithm is applied. The testing is done on all the signals. No initial assumptions about the conditional probabilities were made, so only after the learning process has been completed, the conditional probability table for all the nodes had been constructed. The testing has been done in two ways - real time and offline testing.

In average of every two weeks, a recording has been conducted, enabling us to acquire around 30 acoustic signals, thus giving us a complete database of a over 120 signals throughout the entire lifespan of the grinding plates. In order to test this algorithm in real time, all these signals have been artificially merged so to appear as though the seconds have passed between different measurement sessions, when indeed it has actually been weeks. However, despite the artificially created sudden changes in the state of the mill (which are realistically much slower) the algorithm tested in real-time has shown some very good results.

From Fig. 5 it can be concluded that this simple Bayesian network can successfully detect the states of the grinding plates. Two weeks after they have been changed the algorithm detects that the plates are completely healthy, 5 weeks after the change the deterioration of the plates becomes noticeable, and the Bayesian network detects 25% probability that the plates are worn. After 6 weeks this percentage increases to 50%, whereas after 8 weeks the algorithm reports with 100% certainty that the plates need to be replaced. In order to prevent sudden changes in real time testing environment, the output of the Bayesian network is averaged over the last 10 time periods (each of which lasts 15 seconds). It can be noted

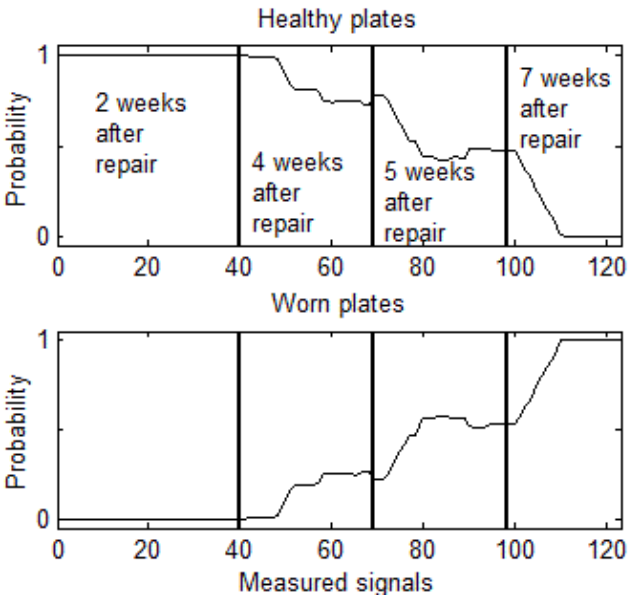


Fig. 5. Results of a real-time test. Bayesian network has managed to calculate the probability for grinding plates to belong to each of the defined states, for each of the measured signals.

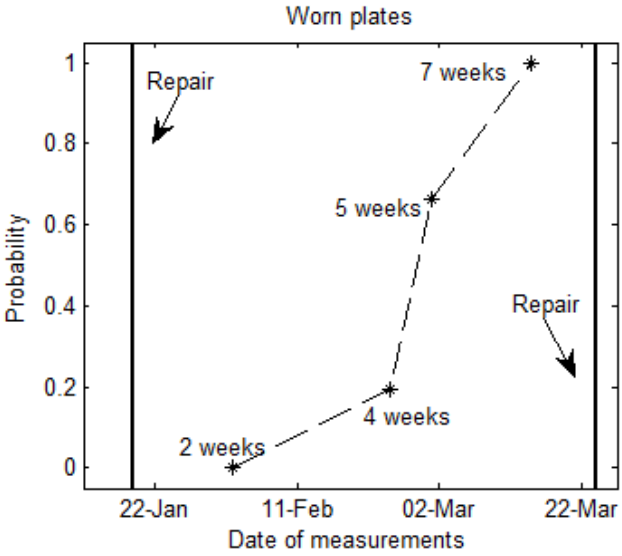


Fig. 6. Results of offline testing. The probability that the grinding plates are damaged is calculated for each day of the measurements. The further the measurements are from the last repair the higher is the probability that the plates need to be changed.

that a sudden change in the state of the plates cannot be detected in a very short period of time. It is needed approximately 4 minutes (15 time periods) for the algorithm to converge to the right state.

Another way of testing this algorithm, which may be more useful when taking into account the nature of the problem, is in an offline simulation. In this case the simulation has been run only after all the signals that day have been recorded. Seeing that the sudden changes in the state of the plates cannot happen, this is a more appropriate way to use this algorithm. It does not require the constant presence of the recording device which can be used once every week or so for a couple of hours to predict whether the maintenance check is necessary. The results are given in Fig. 6 and they are similar to those obtained with a real-time test.

VI. Conclusion

In this paper an application of Bayesian networks on predictive maintenance of ventilation mills in thermal power plants has been investigated. The basic principles of BN have been described and several solutions to the predictive maintenance problem have been proposed. Using statistical signal processing and one of the proposed structures of Bayesian networks, the algorithm has been proposed which manages to detect the state of the plates of the coal grinding mills based on acoustic signals, and therefore predict the time when it is necessary to conduct the replacement of the plates. This algorithm has been successfully tested on a real system during a time period of several months.

The final purpose of this algorithm is the development of the commercially accessible device which would be able to detect the state of ventilation mills and predict the time when it is necessary to repair them. The great advantage of this method is that it uses acoustic signals which can be acquired without disturbing the operation of the mill. Unfortunately, this also means that the algorithm is highly susceptible to noise and acoustic disturbances. However, even with these drawbacks the algorithm developed here yields very good results despite the fact that the simplest BN model has been adopted and the measured signals had a noticeable presence of noise

In future works more complex models of Bayesian networks that include the dynamic behavior of the transition between the states of the system should be tested. Also a better database of the signals should be acquired. Finally it would be interesting to test the algorithm on acoustic signals gained with a cheaper microphone which is commercially more available.

References

[1] A. K. S. Jardine, D. Lin, D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol 20, pp. 1483–1510, 2006.

[2] J. Pearl, "Bayesian networks: a model of self-activated memory for evidential reasoning," *7th Annu. Conf. of the Cognitive Science Society*, Irvine, 1985.

[3] P. Stepanic, I. V. Latinovic, Z. Djurovic, "A new approach to detection of defects in rolling element bearings based on statistical pattern recognition," *Int J Adv Manuf Technol*, vol 45, pp. 91–100, 2009.

[4] S. A. McInerny, Y. Dai, "Basic vibration signal processing for fault detection," *IEEE Transactions on Education*, vol 46, no. 1, pp. 149–156, 2003.

[5] N. Baydar, A. Ball, "A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution," *Mechanical Systems and Signal Processing*, vol 15, no. 6. pp. 1091–1107, 2001.

[6] A. Albarbar, F. Gu, A. D. Ball, A. Starr, "Acoustic monitoring of engine fuel injection based on adaptive filtering techniques," *Applied Acoustics*, vol 71, pp. 1132–1141, 2010.

[7] Y. Zhao, F. Xiao, S. Wang, "An intelligent chiller fault detection and diagnosis methodology using Bayesian belief nettwork," *Energy and Buildings*, vol 57, pp. 278–288, 2013.

[8] S. Verron, J. Li, T. Tiplica, "Fault detection and isolation of faults in a multivariate process with Bayesian network," *Journal of Process Control*, vol 20, pp. 902–911, 2010.

[9] K. P. Murphy, "Dynamic Bayesian networks: Representation, Inference and Learning," Ph.D. dissertation, Dept. Computer Science, Univ. of California, Berkeley, 2002.

[10] E. Kisić, V. Petrović, S. Vujnović, Ž. Đurović, M. Ivezić, "Analysis of the condition of coal grinding mills in thermal power plants based on the T2 multivatiate control chart applied on acoustic measurements," *Facta Universitatis Series: Automatic Control and Robotics*, vol 11, no. 2, pp. 141–151, 2012.

[11] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. Boston: Academic Press, 1980.

USB HW/SW Co-Simulation Environment with Custom Test Tool Integration

Zargaryan Y. Grigor, Aharonyan K. Vahram, Melikyan V. Nazeli and Marko A. Dimitrijević

***Abstract*—This paper describes a new verification environment for USB 2.0 controller. New methodology is presented, where a co-simulation environment is used as one of the starting points for the embedded hardware/software development and as an accelerator of the overall design process. The verification environment is based on the device emulation/virtualization technique, using USB controller's real register transfer level (RTL) instead of models. This approach is functionally very close to the corresponding real-world devices and allows wider opportunities for hardware debugging. The new software utilities for USB host and device functionality testing are also presented. This tool allows generating custom tests by including various transfer types and modifying parameters such as data payload, interval, number of pipes, etc. It can be used for both hardware (HW) and software (SW) limitations characterization, as well as debugging.**

***Index Terms*—Co-simulation, FPGA, System C, QEMU, USB**

Original Research Paper
DOI: 10.7251/ELS1418023G

I. INTRODUCTION

THE Universal Serial Bus (USB) is a fast, bidirectional, low-cost, dynamically attachable communication interface consistent with the requirements of the present microelectronic platforms [1]. It is widely used in the huge number of instruments and devices, which include personal computers, digital cameras, scanners, image devices, printers, keyboards, mice, telephones, embedded systems, systems on a chip (SoC), etc. Fig. 1 shows the high-level architecture diagram of USB controller within a typical system.

Field programmable gate array (FPGA) based techniques can be used for verification and functional debuging in design flow of USB IPs. In this approach, hardware core is mapped onto an emulation platform, based on an FPGA that mimics the behavior of the final chip, and the software modules are loaded into the memory of the emulation platform. Once programmed, the emulation platform enables testing and debugging of hardware (HW) and software (SW) components, close to system's full operational speed. In fact, FPGAs are used primarily to speed up testing and make some parts of the verification easier. Various components of USB system can be prototyped and verified using FPGAs. In [2] the development of USB peripheral core on the FPGA is described. USB protocol analyzer and the core were implemented for verification. For example, only UTMI logic can be designed using VHDL code, then simulated, synthesized and programmed to the targeted FPGA [3]. But the drawback of both approaches [2], [3] is that designed components were not fully validated in complete USB system (including host, USB SW stack, etc.).

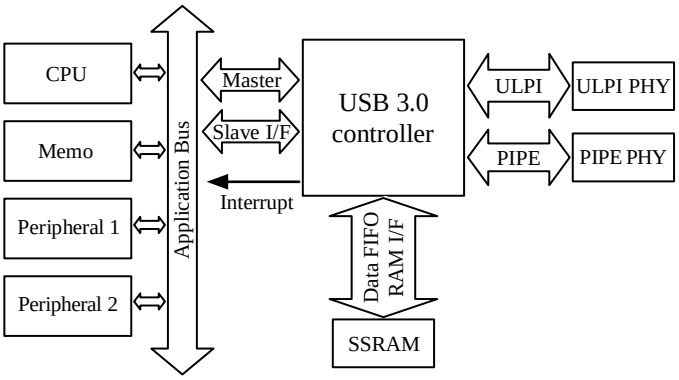


Fig.1. System-Level Block diagram including USB controller.

Another drawback of FPGA based HW verification can be relatively slow debugging of signals, when some problem in register transfer level (RTL) emerges. If the developer wants to look at some signals generated in the core, he needs to perform some modifications in RTL, connect appropriate wires to the debug pins of chip, resynthesize logic, connect analyzing osciloscope tools like Logic Analyzer and trigger the signals or use JTAG cable to connect to FPGA and store necessary signals in RAM memory supported (which requires more resources) and finaly study signals from this store point. If the developer is trying to proceed directly with RTL validation in real system using FPGAs, problems can occur with the absence of other USB system's components (USB PHY chip, programmable processor of SoC, memory access unit, etc.).

On the other hand, in order to actually validate and use the implemented USB hardware in applications running by SoC or some other system, the software layer that encapsulates the underlying hardware and provides the application developer defined API (at minimum it is the so called hardware driver or the hardware abstraction layer; often it contains more software layers like OS modules) should be developed. Thus, only hardware subcomponents (in this case USB IP core licensed by suppliers) become a combination of hardware and software.

Manuscript received 13 April 2014. Accepted for publication 30 May 2014.

G. Y. Zargaryan, V. K. Aharonyan, N. V. Melikyan are with the Synopsys Armenia CJSC (e-mail: {grigorz, vahrama, nazeli}@synopsys.com).

M. A. Dimitrijević is with Faculty of Electrical Engineering, University of Niš, Niš, Serbia, (+381-18-529-321; e-mail: marko@venus.elfak.ni.ac.rs).

In this case, bare USB hardware IP is not enough. Development of this software usually starts when the hardware part is finished (or almost finished) and can be prototyped in an FPGA device. This may considerably extend the project timeframe. Early software development based on virtual prototyping is getting more popular at system level, but in the IP development it has not become a common practice yet. To overcome the above mentioned problems, the necessity arises to have a complete emulation and simulation environment for USB coreduring controller design flow, .

II. Background of known methods for simulation and validation of USB controller

SystemC [4], [5] has solved many of the software-hardware co-simulation issues today, and software based verification perfectly supports the new IP packaging requirements (i.e. requirement to deliver hardware dependent software along with the hardware component). Therefore, to design a USB controller HW in a faster manner and to begin the development of necessary drivers in the first stage of the design, various modeling and verification methods have been suggested. In [6] USB device simulation using Microsoft Device Simulation Framework (DSF) is presented, which allows quick construction of a virtual mass storage USB device having final RTL architecture, but DSF does not provide sufficient capabilities for USB simulation and validation, and offers only smoke tests (i.e. file copy from and to). Paper [7] describes the way of simulation and validation system design for USB peripheral device's Verilog HDL code using SmartModel tools which include functions of USB 2.0 host and UTMI. This environment tests main functions of device peripheral without SW stack (Start of Frame packet generation, handling of Split transfers, ping protocol, etc.). Paper [8] presents the USB host controller verification environment using Transaction Level modeling (TLM), but this environment actually does not cooperate with RTL, nor USB host software, delivered as the final package to the user. Only in [9] an example is given of design flow that includes real HW (described in RTL language) validation with SW stack. Here, SystemC TLM is used for hardware architecture definition and preliminary SW development, in the first step. RTL verification and prototyping is the next step. User-level test is done only after USB RTL prototyping to the FPGA.

In this paper, a new approach to HW/SW co-simulation and verification technique is proposed, using two entries of USB Super-Speed (SS) core acting as a host and device. It allows testing and validation of Verilog HDL code of USB core together with depending software (SW) drivers, whole Linux USB stack, running real user level applications (for either smoke or stress tests, or specific protocol testing means [10]), which makes simulation very close to real testing and debugging. The main role of the simulation environment is to make the USB controller work in real physical world conditions prior to implementation on FPGA or ASIC. The USB host and device controllers are attached to the emulated x86 PCs running 3.6.3 Linux kernel using PCI bus. The simulation environment provides full access to all signals of RTL, during and after simulation. The possibility of capturing and analyzing data packets sent and received over USB bus is also provided. PC systems are emulated using QEMU open source virtualization software which works straightforwardly with verification environment, considering it as a combination of two real-world Linux systems connected to each other via USB. Another important advantage of this method is that the other USB controller (EHCI, OHCI, USB 3.0 xHCI) RTL codes can be easily integrated into verification environment as well, after specific changes in Verilog source code of USB PHY model. As Linux OS runs on emulated PC, the validation of these controllers with corresponding drivers included in Linux kernel becomes suitable and does not require any specific changes in the SW stack.

The testing of complete USB hardware and software environment is performed using certification tools – USB CV, USB GoldenTree environment, WHCK, etc. [11], [12]. A majority of these tools run on the final product and cannot be used during early stages of the USB controller design. For Linux systems there is a combination of *ustbtest* class and *g_zero* gadget drivers [10], which allow generating traffic on the various pipes of the host and device controllers either on final product or using co-simulation environment. Actually, this tool is for USB 2.0 mode only and does not allow testing various configurations and interfaces, changes of test parameters, etc. This paper also presents new testing tool integrated within suggested co-simulation environment that provides flexible interface for generating custom tests running in USB 2.0 and 3.0 modes, gathering statistics on the test results and observing failures.

III. QEMU virtualization software

QEMU [11] is a processor emulator that relies on dynamic binary translation in order to achieve a reasonable speed. QEMU is easy to port on new host CPU architectures. In conjunction with CPU emulation, it also provides a set of device models, ensuring possibility to run a variety of unmodified guest operating systems. It can thus be viewed as a hosted virtual machine monitor. It also provides an accelerated mode for supporting a mixture of binary translation (for kernel code) and native execution (for user code). It has two main modes of operation [12], [13]:

- User mode emulation - QEMU can launch processes on CPUs that are compiled on another CPU. It is used to ease cross compilation or cross-debugging.
- Full system emulation - QEMU emulates a full system (for example PC), including processors and several peripherals.

The second mode is used to launch different kinds of operating systems on the same PC, as virtual machines. This mode is used to build co-simulation environment. One of the main advantages of QEMU is the fact that it runs in user space. There is no kernel module and therefore no adherence with the host system. It is executed like any other application. QEMU

is a layer oriented application. The highest layer in QEMU is the lowest layer for the system inside the emulator (i.e. the hardware).

Basically, QEMU supports various processor architectures but in terms of this task Intel x86 is used. QEMU can emulate several hardware controllers such as CD-ROM, IDE hard disk interface, parallel and serial port, sound card, USB controller, etc. QEMU also emulates a PCI UHCI USB controller. User is able to virtually plug virtual USB devices or real host USB devices and emulator will automatically create and connect virtual USB hubs which is necessary to connect multiple USB devices. This is based on QEMU internal models of UHCI controller, and it allows to validate real USB controller.

IV. Proposed USB RTL/SW co-simulation approach

The main difference and novelty of this approach from known verification/simulation methods is that the real RTL implementation of the core (not TLM or any other model of the controller) is co-simulated using virtualization, giving a chance to perform complete functionality tests on USB IP.

The simplified block diagram of co-simulation environment is presented in the Fig 2. Two entries of SS RTL core are interconnected via PHY model, written in Verilog. Advanced High-performance Bus Verification IP (AHB VIP) performs memory read/write actions between the system and the USB controller (handling any action regarding data reception/sending, direct memory access (DMA) actions, control and status registers (CSR) read write, etc.). Using QEMU, two x86 based PC systems are emulated as host and device, and both work with implemented Direct Programming Interface (DPI) which allows higher level Linux/QEMU C code to communicate with Verilog/Vera implementation and vice versa. User-level application, device and host drivers as well as Linux OS make the software layer of environment. User level USB testing application is registered on the driver and installed in Linux.

One of the main components of designed verification environment is the DPI layer, serving as a bridge between SVerilog/Vera and C code domains. DPI consists of two layers: A SVerilog/Vera Layer and a C language layer. Both layers are isolated from each other. DPI allows direct inter-language function calls between the SVerilog/Vera and C language. The functions implemented in C language are called from SVerilog/Vera and such functions are called Import functions. Similarly, functions implemented in SVerilog/Vera code domain can be called from C language, such functions are called Export functions. DPIs allow data transfer between two domains through function arguments. The main functions of DPI are forward DMA reads and writes initiated by RTL to memory in 'C' domain, report interrupt status to 'C' domain, execute processor ('C' domain) Read/Write to registers inside the core, control advancing simulation time while performing one of these actions. In Fig. 3 the common mechanism of DPI implementation is shown. For example *dma_read* function that is implemented in C code domain should be imported to SVerilog/Vera code, and correspondingly after implementing *ahb_read_vr* task in SVerilog/Vera file, it should be exported there and declared as extern function in C code. The main specific thing here is that requests from SVerilog/Vera domain to 'C' domain have highest priority and are implemented as blocking statements. They must be executed in '0' simulation time.

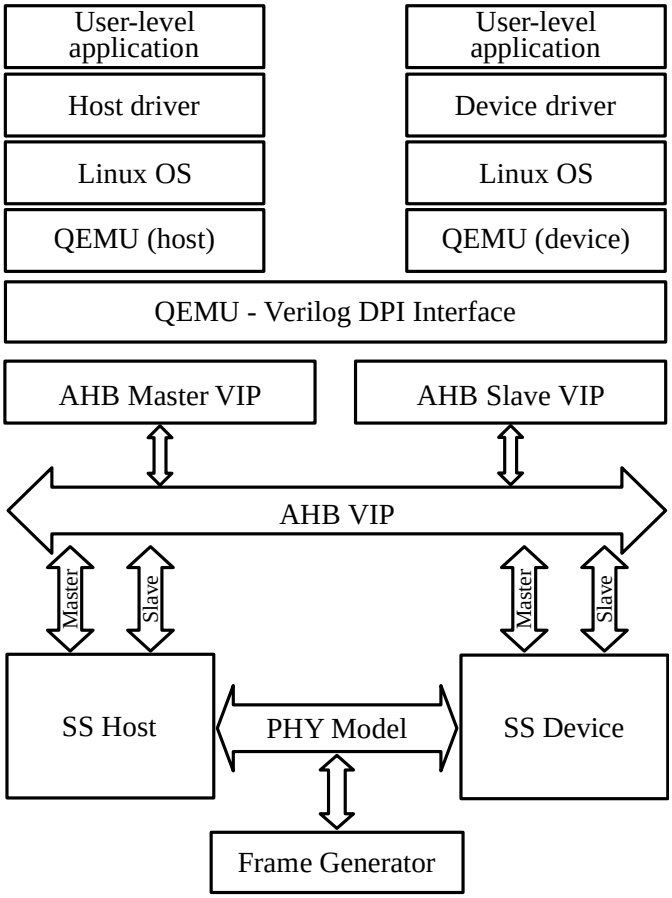


Fig. 2 Block diagram of co-simulation environment.

C domain

- void dma_read(int addr, int* value) { };
- extern void ahb_read_vr(int*data, int addr, int length)

SVerilog/Vera domain

- import "DPI" context task dma_read(integer addr, var integer value);
- export "DPI" task ahb_read_vr { };

Fig. 3 Direct Programming Interface.

Co-simulation environment runs in three processes simultaneously: VCS *simv*, QEMU host, and QEMU device.

The flow of building and running developed environment is shown in Fig. 4. It consists of the following major steps:

• Compiling Linux kernel, QEMU C-sources using gcc. Compile RTL sources of USB controller, SS PHY model using VCS tool, integrate DPI (result will be *simv* simulation executable).

• Compiling and installing the developed driver for USB controller, setup USB user level testing utilities.

• Running three threads on Linux machine: QEMU Host, QEMU Device and *simv* executable (HW simulation)

• Loading SW modules on QEMU emulated devices, run user-specified tests,

• Debugging with the use of VPD file (signal diagrams from RTL simulation are shown in Fig. 5), console logs.

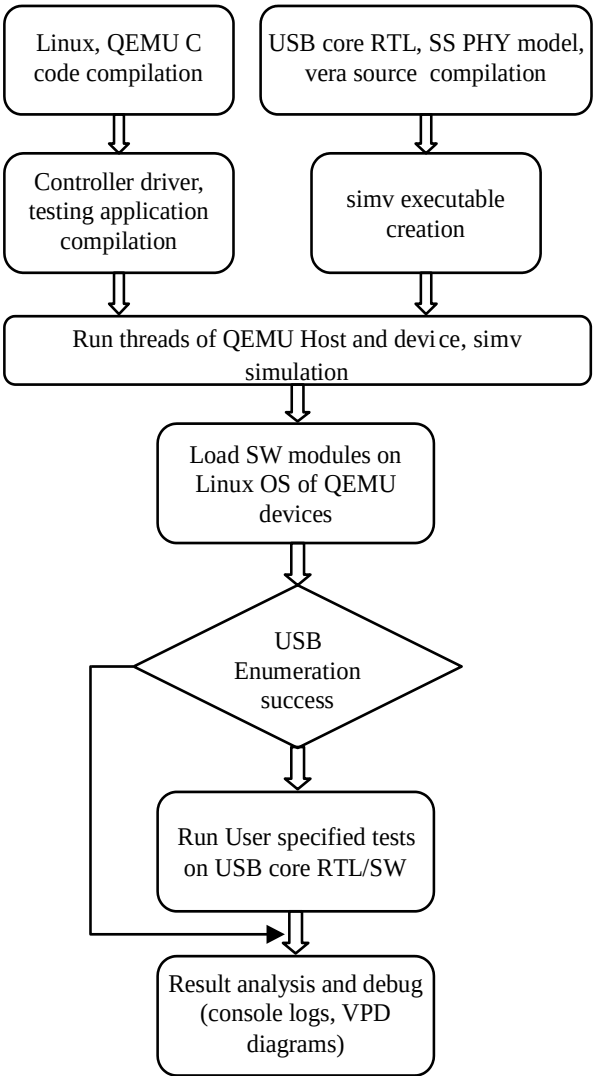


Fig. 4 Co-simulation environment setup and run flow.

It is important to note that after running QEMU host and device threads, third thread responsible for running *simv* executes simulation of the whole Verilog/Vera code domain including USB core RTL. On both systems SS USB core is registered on PCI bus and system times are synchronized with simulation time to ensure absence of concurrencies between them and RTL simulation time. At this point, virtually emulated PCs have fully functional SS USB controller on their disposal. While loading core drivers on both sides, OS are requesting recourses using communication interface with PCI registered devices (mapping memory, interrupt line allocation, driver registering in OS, etc.) and proceeds with USB cores initialization. Afterwards USB host detects device connection, notifies upper layers of Linux USB stack and initiates USB bus reset to begin enumeration. After successfully passing enumeration, user level USB testing application can be executed in this environment to validate USB core RTL/SW combination.

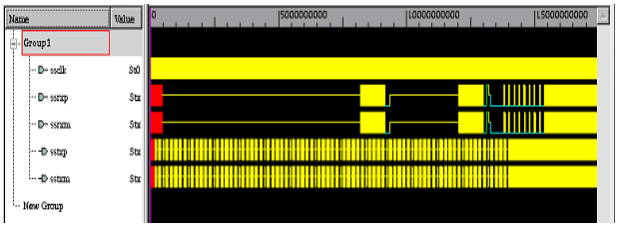


Fig.5. Signal diagrams from RTL simulation on USB core.

V. User level USB testing application

Implemented test tool consists of three modules – test application, class driver and gadget driver. The last one is loaded on the device side of the USB environment and the first two are executed on the host side. The architecture of the implemented test environment is shown in Fig. 6.

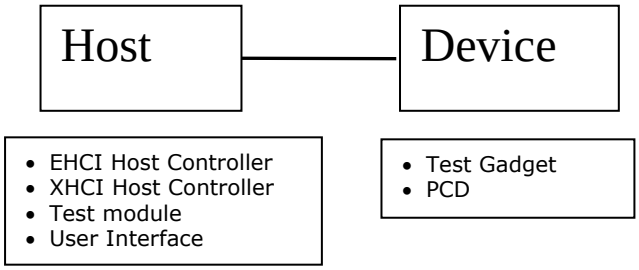


Fig. 6 Testing application architecture.

Test application provides the command line user interface for test cases execution. Test application parses the command line options, analyzes the test parameter values, checks availability of the corresponding pipes and sends appropriate commands to the Linux test class module for further test execution and final results reporting.. The mentioned class driver is the interface between test application and USB Host controller driver. Class driver creates USB transfer blocks, submits USB requests to the USB core with defined communication protocol, and report results to test application after completion. The implemented testing gadget driver submits transfer requests to the underlying Peripheral Controller Driver (PCD) and handles their completion. The gadget driver is responsible for data verification for OUT transfers and for data generation in IN transfer mode. There are also loop back mode tests, when the host writes the data to

the device first, then reads it back and verifies its accuracy.

For communication between host and device, vendor specific protocol is developed (Table 1) in addition to standard USB requests.

TABLE 1 VENDOR SPECIFIC PROTOCOLS

Request Mnemonic	Description/Purpose
DEV_CAP	The device will return the device capabilities.
CONF_EP_IF	Allows the host to change the parameters of endpoints and configure an interface.
TEST_SETUP	The host sets the test parameters.
TEST_STAT	The device will return the Test Case status.
CONTROL_WRITE	Writes data to the control pipe.
CONTROL_READ	Reads data from the control pipe.
SET_TR	Provides to device for setting up a transfer for an Endpoint.
RET_RES	The device will return the results of the most recent transfer.
DEV_RESET	Resets the Device or Interface to a known state.

User interface module is a command line interface. Each test case (or group of test cases) needs to be configured individually with flexible input arguments through user interface module. Each configuration parameters can be divided into the following groups: endpoints parameters which represent endpoint descriptors main fields and endpoint's pair number (interface, number, direction, type, and max packet size), test parameters (mode, transfer size and iteration count).

A special algorithm is developed to handle errors. The transfer buffers are filled with data using implemented algorithm. This algorithm is common for the test gadget and the host test module. Depending on the direction of the test, the test module or test gadget will fill the transfer buffer. The other part will verify the received data.

The USB test gadget provides functionality for implementing the Communication protocol. The test gadget handles all vendor specific requests, controlling endpoints and interface.

VI. STATISTICS AND TEST RESULTS

Co-simulation environment requires significant resources from host Linux machine. The analysis shows that major CPU load is caused by *simv* executable running. Simulation profiling shows that about 50% of CPU utilization by *simv* is due to Verilog modules (core RTL design itself). The basic timing statistics of co-simulation and CPU utilization are follows:

- Compilation done – time 7.2 min (performing once during setup)
- QEMU Host and Device booting – 2.5 min
- Driver loading done – time 1.3 min
- Core Initialization by driver done – time 1.2 min
- Simulating 1mSec – time 3.6 sec
- CPU Utilization Profile [%]: total time 20 min

Direct Programming Interface (DPI)	1.53%
Programming language interface (PLI)	24.45%
VCS for writing VCD and VPD files	6.82%
VCS for internal operations (KERNEL)	15.86%
Verilog Modules	50.69%
A SystemVerilog testbench program block	0.38%
Garbage Collector(PROGRAM GC)	0.27%

As it is shown in statistics analysis, it takes about 3–4s of real time to simulate 1 ms of hardware work. Hence basic file copy test between host and device systems via USB, with data size, for example, 1MB takes about 3 minutes of real-time, but based on simulation time copy is performed with 46-60MB/s speed.

After driver loading phase and USB cores' initialization by SW on emulated host and device systems, various tests are performed using the suggested test utility – control, bulk, isochronous and interrupt traffic generation in both IN and OUT directions. Also, test parameters like maximum packet size, interval for periodic endpoints, burst for USB 3.0 transfers, stream count, etc. were modified during tests. Obtained results are captured in Table 2.

TABLE 2. LINUX USBTEST UTILITY TEST RESULTS IN CO-SIMULATION ENVIRONMENT

Test Description	Max packet size	Max burst size	Tr. size	Interval	Duration	Result
Bulk in	1024	0	0.1M	-	0.328s	Pass
Bulk in	1024	10	0.1M	-	0.18s	Pass
Bulk out	1024	0	0.1M	-	0.331s	Pass
Bulk out	1024	10	0.1M	-	0.19s	Pass
Control in	512	-	1M	-	2.083s	Pass
Control out	512	-	1M	-	2.063s	Pass
Isochronous in	1024	0	10M	1	1.5s	Pass
Isochronous out	1024	0	16K	14	19.09s	Pass

VII. CONCLUSIONS

In this paper a new approach to USB controller RTL and driver co-simulation and verification method is proposed. The suggested method is based on virtualization techniques (QEMU emulator) using USB SS core. The method is based on the idea of using a real USB controller RTL model, instead of TLM or other abstract models. The designed environment allows identifying/debugging USB HW design easily, debugging device driver before real tests on FPGA or SoC,

reproducing and debugging issues emerging in the real circumstances using signal diagrams obtained from simulation, beginning device driver development and testing when hardware is not available yet. A user level USB testing application is also developed, which allows running custom tests with various test parameters using different speeds. The implemented application has the ability to ensure four types of transfer running in virtual environment. The method is flexible and does not depend on the USB controller or any other USB IP's RTL, and can be easily integrated into the environment after minor changes in the USB PHY model.

REFERENCES

[1] Universal Serial Bus 3.0 Specification. Rev 1.0. - June 2011, www.usb.org

[2] De Maria, E.A.A., Gho, E., Maidana, C.E., Szklanny, F.I., Tantignone, H.R., A Low Cost FPGA based USB Device Core // IEEE 4th Southern Conference on Programmable Logic, 2008, pp. 149-154.

[3] Babulu, K., Rajan, K.S., FPGA Implementation of USB Transceiver Macrocell Interface with USB2.0 Specifications // ICETET '08. First International Conference on Emerging Trends in Engineering and Technology, July 2008, pp. 966-970.

[4] IEEE Computer Society, IEEE Standard System C Language Reference Manual, 2005, pp. 423.

[5] OSCI, Transaction Level Modeling(TLM) Library, Release 1.0, 2005

[6] Anderson, R.B., Borowczak, M., Wilsey, P.A., The Use of Device Simulation in Development of USB Storage Devices // IEEE 41st Annual symposium, 2008, pp. 220-226.

[7] Xiaoping, B.; Shenlei, The Study on System Verification of USB2.0 Interface Protocol Control Chip Hardware Design // ICEE '07. International Conference on Electrical Engineering, 2007, pp. 1-5.

[8] Puhar, P., Zemva, A., Functional Verification of a USB Host Controller // IEEE 11th EUROMICRO Conference on Digital System Design Architectures, Methods and Tools, 2008, pp. 735-740.

[9] Sobański, I.; Sakowski, W., Hardware/software co-design in USB 3.0 mass storage application // IEEE International Conference on Signals and Electronic Systems (ICSES), 2010, pp. 343-346..

[10] USB Testing on Linux, March 2007, http://www.linux-usb.org/usbtest/.

[11] SuperSpeed USB Compliance Testing http://www.usb.org/developers/ssusb/testing

[12] QEMU open source processor emulator, http://www.qemu.org, 2013.

[13] Ribière, A., Emulation of obsolete hardware in open source virtualization software // IEEE 8th IEEE International Conference on Industrial Informatics (INDIN), 2010, pp. 354-360.

Measuring Capacitor Parameters Using Vector Network Analyzers

Deniss Stepins, Gundars Asmanis, and Aivis Asmanis

***Abstract*—Vector network analyzer (VNA) is versatile measuring equipment which is primarily used for two-port device S parameters measurements. This paper addresses measurement of capacitor parameters using VNA in broad frequency range. The main attention is focused on the measurement accuracy of capacitors parameters using VNA and proper de-embedding of an experimental setup parasitics to get accurate results. Comparative measurement error analysis for different measurement techniques is presented. Suitability of each measurement technique for measurements of capacitor parameters using VNA is discussed and effect of the experimental setup parasitics on the measurement results is addressed. Moreover useful procedures for proper capacitor impedance measurement using VNA are developed.**

***Index Terms*—Capacitors, vector network analyzers, measurements, measurement accuracy.**

Original Research Paper
DOI: 10.7251/ELS1418029S

I. Introduction

Capacitors are used in many electronic devices, especially in radio-frequency equipment, such as transmitters, receivers, conducted electromagnetic interference filters, etc.

Capacitors capacitance is typically ranging from several pF to several thousands of µF. The main parameters of a capacitor are its capacitance C and impedance magnitude $|Z_c|$ (as a function of frequency). Real capacitors, of course, have parasitic parameters such as series resistance R_s (mainly lead wires resistance), parallel resistance R_p (due to losses in capacitor dielectric), equivalent series inductance (ESL) and self-resonance frequency f_{res}. Real capacitors can be substituted by their equivalent circuit model shown in Fig.1 (a). However in practice when designing electronic circuits with capacitors, more simple capacitor equivalent circuit model is usually used as shown in Fig.1(b). In this model ESR accounts for both R_p and R_s [1]. Usually knowing ESR is much more important than R_p and R_s separately [2]. That is why capacitor manufacturers only give ESR values in their catalogs.

Manuscript received 15 April 2014. Received in revised form 7 June 2014. Accepted for publication 11 June 2014.

This work was supported by the Project "Competence Centre of Latvian Electric and Optical Equipment Production" co financed by the European Union (agreement no. L-KC-11-0006). D. Stepins, G. Asmanis and A. Asmanis are with the Latvian Electronic Equipment Testing Center, Riga, Latvia (phone: +37128327653; e-mail: stepin2@inbox.lv).

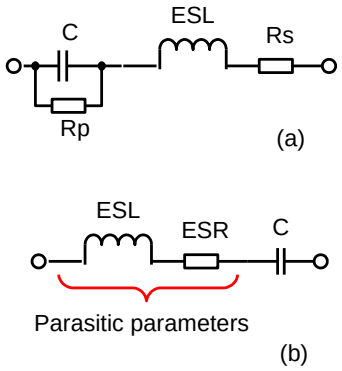


Fig. 1. Real capacitor equivalent circuit model (a) and its simplified version (b).

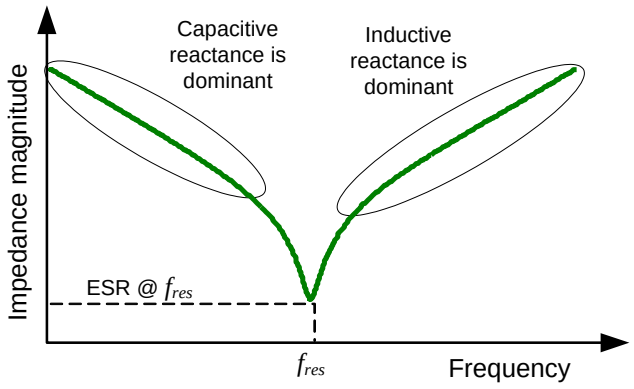


Fig. 2. Real capacitor impedance magnitude versus frequency.

At a self-resonance frequency f_{res} a capacitor has its minimum $|Z_c|$, which is equal to ESR (at f_{res}). For $f<f_{res}$, $|Z_c|$ is decreasing function of f, but for $f>f_{res}$, $|Z_c|$ increases as f increases, Fig.2.

Accurate measurement of capacitor parameters (including parasitic ones) is of major importance when designing electronic equipment. For capacitor and inductor parameters measurements in broad frequency range usually impedance analyzers are used [3] – [10]. However the impedance analyzers are expensive and the measurement frequency range is usually limited up to several hundreds of MHz [3], [6], [10]. In contrast to the impedance analyzers, vector network analyzers (VNA) are less expensive, more versatile and can be used for capacitor parameters measurements in much broader frequency range up to tens of GHz.

For two-terminal passive electronic components measurements using VNA, three measurement techniques can be used: reflection (based on reflection coefficient S_{11} measurement), shunt-through (based on transmission coefficient S_{21} measurements) and series-through techniques (based on S_{21} measurements) as shown in Fig.3 [3], [6]. Shunt-through technique is usually used for low-impedance passive

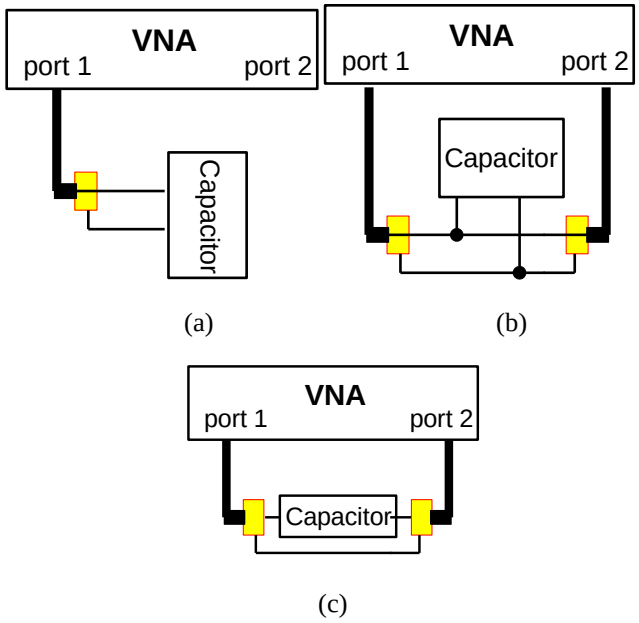


Fig. 3. Two terminal passive components measueremnt techniques using VNA: (a) reflection technique; (b) shunt-through technique; (c) series through technique.

electronic components, while series-through technique for high-impedance components [3].

Although different measurement techniques of two-terminal passive electronic components are investigated in [3], [6] - [10], no one of the papers consider measurement accuracy of VNA techniques. Moreover suitability of each VNA technique for capacitors parameters measurements is not addressed and proper de-embedding of experimental setup parasitic parameters is not considered in the papers. For example, in [10] the main attention is focused on auto-balancing bridge techniques for ESR measurements and VNA techniques are not investigated thoroughly, in [6] inductors and capacitors impedance measurement repeatability and temperature dependence for the radio-frequency impedance analyzers and VNA is comparatively investigated, but in [3] experimental measurements of low-impedance inductors and ferrite complex magnetic permeability using VNA are presented without addressing any measurement accuracy issues.

The main original contribution of this paper is focused on the measurement accuracy of capacitors parameters using VNA and proper de-embedding of an experimental setup parasitics to get accurate measurement results. Comparative measurement error analysis for all the three techniques will be presented. Suitability of each measurement technique for capacitors parameters measurements using VNA will be discussed and effect of the experimental setup parasitics on capacitor measurement results will be addressed. Moreover useful procedures for proper capacitor impedance measurement using VNA will be developed.

II. Measurement of Capacitor Parameters

A. Experimental setups

In order to get capacitor impedance and other capacitor parameters, *S* parameters should be measured first. When using the reflection technique S_{11} should only be measured; for the shunt-through technique S_{21} should only be measured; when using the series-through technique S_{21} should only be measured [3],[6]. For *S* parameters measurements VNA Rohde and Schwarz ZVRE is used. Necessary capacitor parameters are then extracted from the *S* parameters measured. The measurements are done in the frequency range 100 kHz – 500 MHz, with VNA intermediate frequency filter bandwidth of 300 Hz and 1600 points per sweep. After the warm-up time of one hour, full two-port TOSM (through-open-short-matched) calibration has been performed to significantly reduce systematic errors due to imperfect VNA cables and connectors.

Experimental setups used for the capacitor parameters measurements are shown in Fig.4 – Fig.6. To perform one-port reflection measurements, a capacitor is attached to one VNA cable using type N RF connector (Fig.4(a)). To perform two-port shunt-through capacitor measurement, two VNA cables and capacitor are connected in parallel as shown in Fig.5(a). To perform two-port series-through capacitor measurement, two VNA cables and capacitor are connected in series as

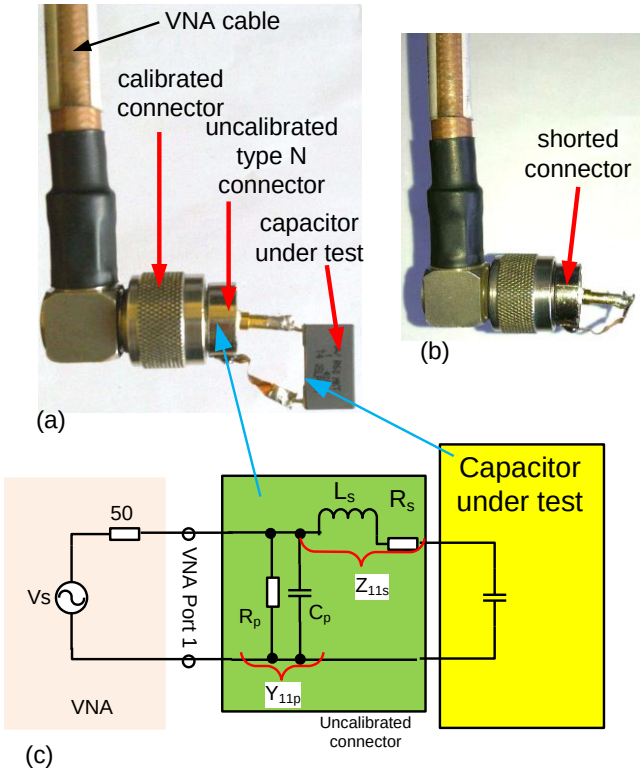


Fig. 4. (a) Experimental setup to perform the reflection technique; (b) shorted N type connector; (c) equivalent circuit model.

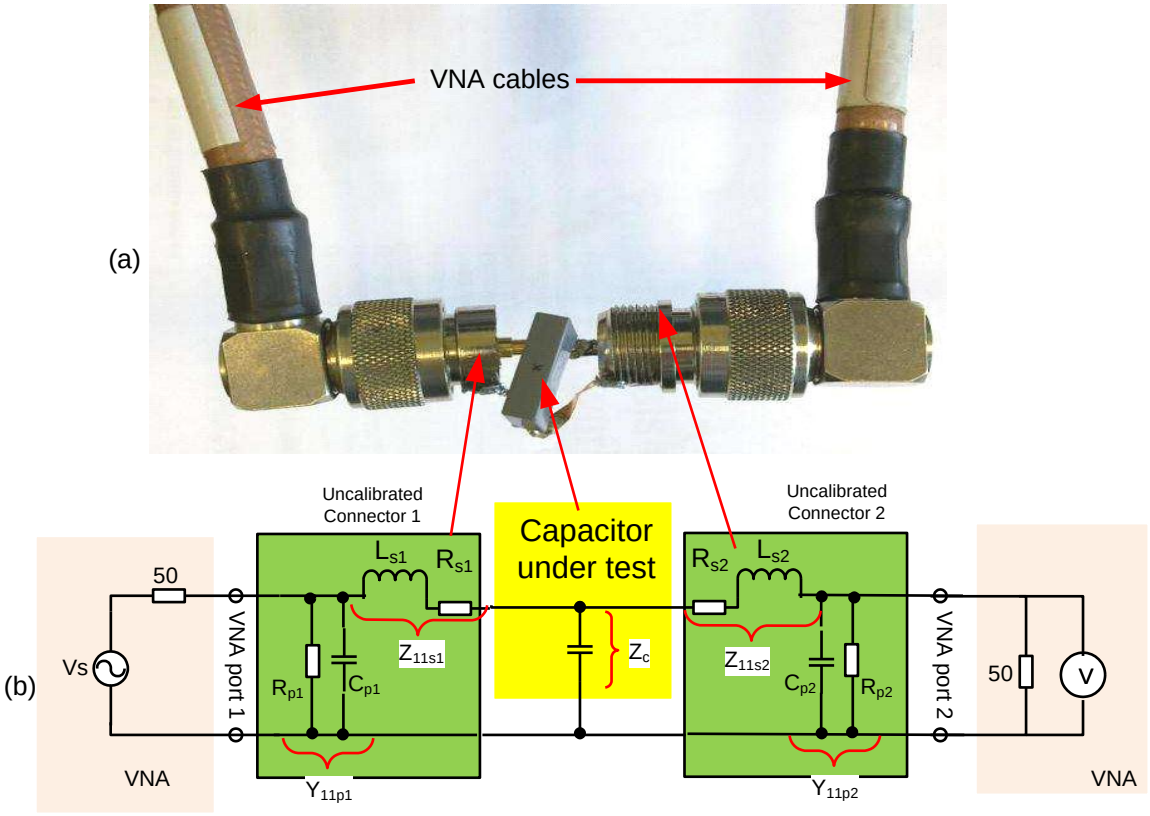


Fig. 5. Experimental setup to perform the shunt-through technique (a) and its equivalent circuit model (b).

shown in Fig.6(a).

Since it is impossible to make TOSM calibration for the N type RF connectors to which the capacitors are attached in the experiments, the parasitic parameters (parasitic equivalent series impedance Z_{11s} and parasitic equivalent conductance Y_{11p}) of the connectors should be taken into account and removed from the capacitors measurement results mathematically using de-embedding.

B. De-embedding of the connectors parasitics

To de-embed the connector's parasitics, simple equivalent circuit models of the experimental setups for all the three measurement techniques are developed in the paper as shown in Fig.4 – Fig.6. The equivalent circuit models can give us a possibility to derive expressions for capacitors complex impedance Z_c calculations.

- For the reflection technique it can be derived

$$Z_c = \frac{1}{1/Z_{11} - Y_{11p}} - Z_{11s} \quad , \qquad (1)$$

where Z_{11} is complex impedance which can be calculated from the measured reflection coefficient S_{11} as follows [3]

$$Z_{11} = 50\frac{1+S_{11}}{1-S_{11}} = 50\frac{1-|S_{11}|^2 + j2|S_{11}|\sin\varphi}{1-2|S_{11}|\cos\varphi+|S_{11}|^2}. \qquad (2)$$

If parasitics of the connectors are not high, then approximately Z_c can be calculated using (2).

- For the shunt-through technique it can be derived

$$Z_c = \frac{50Z_{11s1}Z_a + Z_{11s1}Z_{11s2} + 2500 + 50Z_{11s2}(50Y_{11p2}+1)}{\frac{100}{S_{21}} - (100 + Z_b + 50Y_{11p1}Z_{11s2} + Z_{11s2}(1+50Y_{11p2}))}, \qquad (3)$$

where $Z_a = 1 + Z_{11s2}Y_{11p2}(1+50Y_{11p1}) + Y_{11p1}(50+Z_{11s2})$

$$Z_b = Z_{11s2}(1+50Y_{11p2}) + 2500(Y_{11p1}+Y_{11p2})$$

Z_{11s1} and Z_{11s2} are parasitic equivalent series impedances of the 1st and the 2nd uncalibrated connectors respectively; Y_{11p1} and Y_{11p2} are parasitic equivalent conductances of the 1st and the 2nd uncalibrated connectors respectively.

If parasitics of the connectors are not high, then approximately Z_c can be calculated as follows [3]

$$Z_c \approx 25\frac{S_{21}}{1-S_{21}} = 25\frac{|S_{21}|\cos\varphi - |S_{21}|^2 + j|S_{21}|\sin\varphi}{1-2|S_{21}|\cos\varphi+|S_{21}|^2}. \qquad (4)$$

- For the series-through technique it can be derived

$$Z_c = \frac{\frac{2Z_1}{S_{21}} - 50 - Z_1 - Z_s - 50Z_1Y_{11p1}Z_s}{1+50Y_{11p1}} - Z_{11s} \quad , \qquad (5)$$

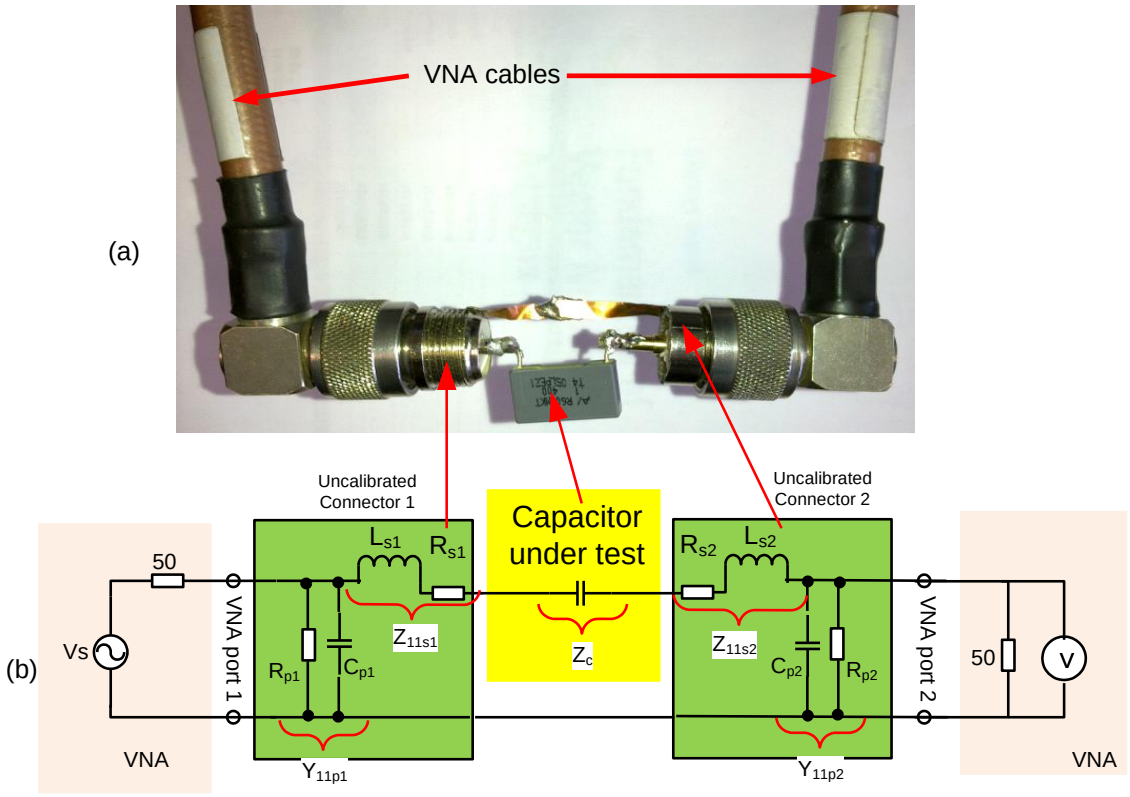


Fig. 6. Experimental setup to perform the series-through technique (a) and its equivalent circuit model (b).

where $Z_s=Z_{11s1}+Z_{11s2}$; $Z_1=1/(1/50+Y_{11p2})$.

If parasitics of the connectors are not high, then approximately Z_c can be calculated as follows

$$Z_c \approx 100\frac{1-S_{21}}{S_{21}} = 100\left[\frac{\cos\varphi}{|S_{21}|} - 1 - j\frac{\sin\varphi}{|S_{21}|}\right] \qquad (6)$$

C. *Determination of the connectors parasitics*

Typical RF N type connector parasitic capacitance is about 1 pF and inductance of shorted connector is about 10nH. This means that parasitic resonant frequency of the connectors is well above 1 GHz. Since all the measurements are done in the frequency range up to 500 MHz, then parasitic series impedances Z_{11s1} and Z_{11s2} can be determined by measuring impedance of shorted connectors, but parasitic parallel conductances Y_{11p1} and Y_{11p2} can be determined by measuring conductance of open connectors. The measurement results for the two connectors used in the experiments are shown in Fig.7. As it can be seen the shorted connector impedance magnitude increases almost linearly as frequency increases, because inductive reactance (due to L_s) is dominant, and conductance magnitude of the open connectors also increases almost linearly because capacitive conductance (due to C_p) dominates.

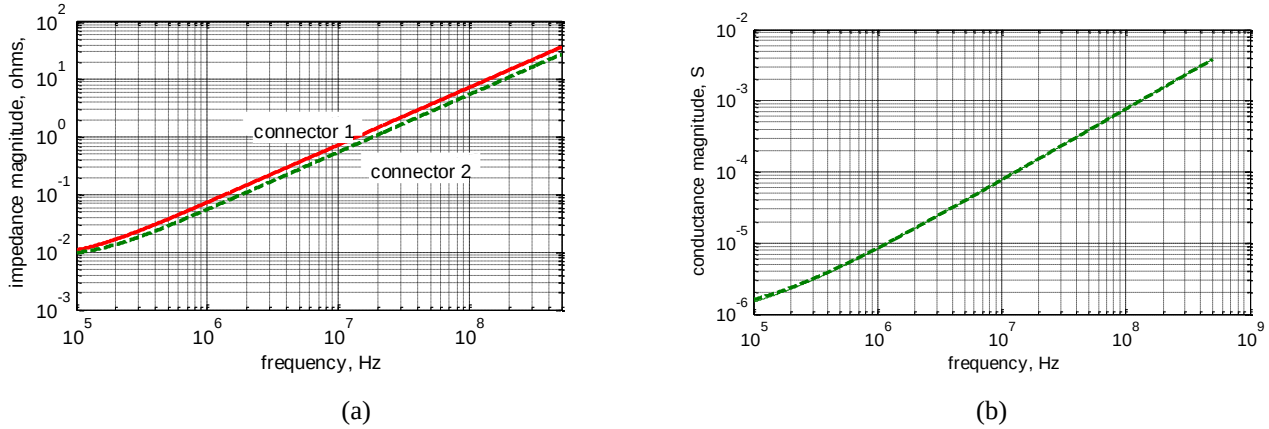


Fig. 7. Measured impedance magnitude of short connectors (a) and conductance magnitude of open connectors (b).

D. *Procedures for capacitor impedance measurement*

In this section the procedures developed for capacitor impedance measurements using all the three measurement

techniques are presented.

Procedure for the capacitor impedance measurement using the reflection technique

- make full two-port TOSM calibration;
- connect the type N RF connector with soldered capacitor (Fig.4(a)) to VNA cable;
- measure complex S_{11};
- desolder the capacitor from the connector and connect shorted connector without capacitor (Fig.4(b)) to the VNA cable and measure complex Z_{11s};
- connect open connector without capacitor to the VNA cable and measure complex Y_{11s};
- calculate complex impedance Z_{11} using (2);
- de-embed the connector parasitics using (1) to calculate the capacitor complex impedance Z_c;
- end.

Procedure for the capacitor impedance measurement using the shunt-through technique

- make full two-port TOSM calibration;
- connect both N type RF connectors in parallel and solder a capacitor under test to them (Fig.5(a));
- measure complex S_{21} using VNA;
- measure both connectors parasitcs Z_{11s1}, Z_{11s2}, Y_{11p1} and Y_{11p2};
- calculate the capacitor complex impedance Z_c using (3);
- end.

Procedure for the capacitor impedance measurement using the series-through technique

- make full two-port TOSM calibration;
- solder a capacitor under test to both connectors in series as shown in Fig.6(a);
- connect them to VNA cables;
- measure complex S_{21} (with the capacitor under test);
- measure both connectors parasitics Z_{11s1}, Z_{11s2}, Y_{11p1} and Y_{11p2};
- calculate the capacitor complex impedance Z_c using (5);
- end.

E. Extraction of capacitor parameters from complex impedance Z_c

- capacitance C can be calculated at the lowest frequency f_1 (at which parasitic inductive reactance is negligible):

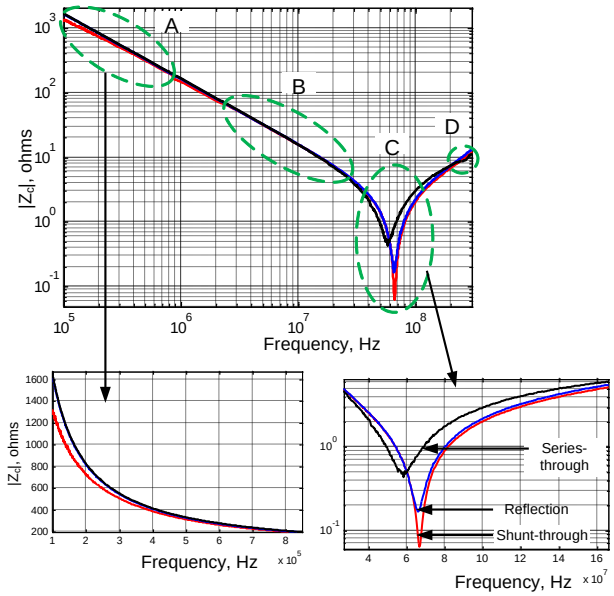
$$C = \frac{1}{2\pi f_1 \,|\, imag\{Z_c\}\,|} \qquad (7)$$

- capacitor ESR(f)=real{$Z_c(f)$};
- capacitor resonant frequency f_{res} is frequency at which capacitor complex impedance magnitude $|Z_c|$ is minimum;
- capacitor equivalent series inductance can be calculated as follows:

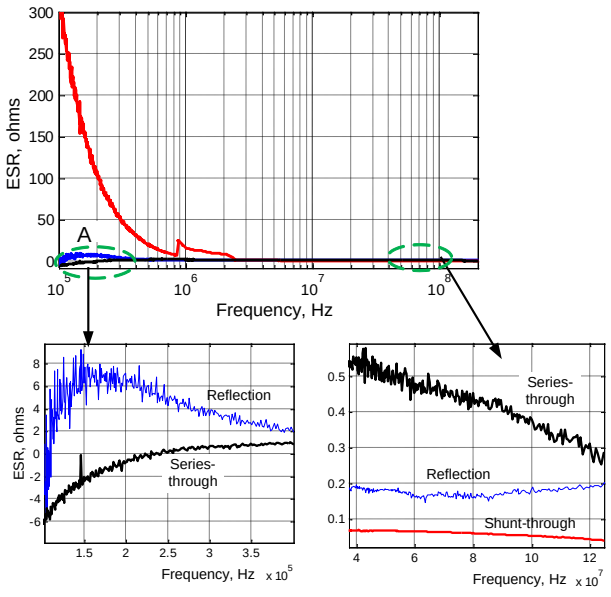
$$ESL = \frac{1}{4\pi^2 f_{res}^2 C} \,. \qquad (8)$$

F. Comparison of the measurement results

Using the procedures developed capacitors with different



(a)



(b)

Fig. 8. Measured 1nF capacitor impedance magnitude (a) and ESR (b) versus frequency.

nominal capacitance values (1nF, 10nF, 100nF and 1000nF) have been measured using the reflection, shunt-through and series-through techniques. The capacitors with nominal capacitances of 1nF, 100nF and 1000nF are metalized polypropylene film capacitors with capacitance tolerance of ±5%, but the capacitor with nominal capacitance of 10nF is

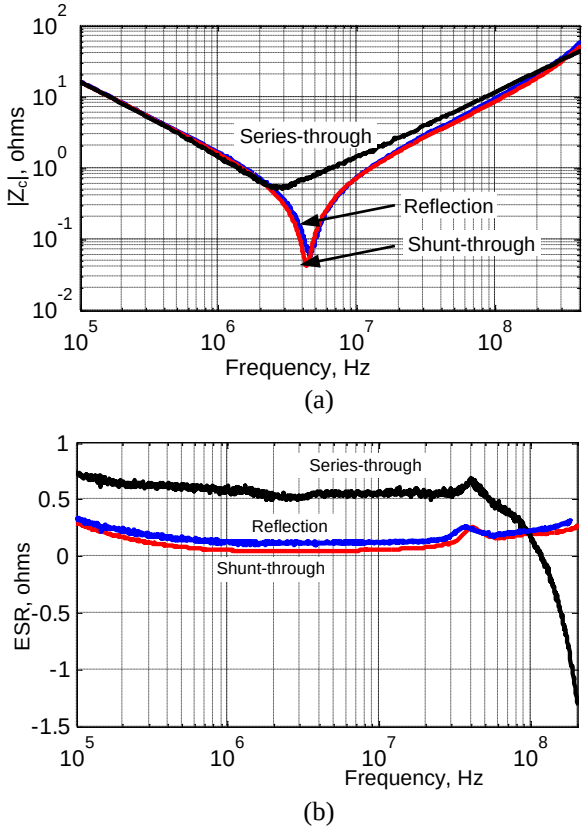


Fig. 9. Measured 100nF capacitor impedance magnitude (a) and ESR (b) versus frequency.

polyester film capacitor with tolerance of ±20%.

Measured $|Z_c|$ and ESR versus frequency for several capacitors are comparatively depicted in Fig.8 and Fig.9. But in Table 1 measured capacitor C and ESR (at f=100kHz) using VNA and high-precision LCR meter Instek LCR 819 are tabulated. Analysis of the measurement results and measurement errors will be presented in section IV.

TABLE 1
COMPARISON OF MEASURED CAPACITANCE AND ESR (AT 100KHZ) USING HIGH-PRECISION LCR METER AND VNA FOR DIFFERENT CAPACITORS

Nominal capacitance, nF			**1**	**10**	**100**	**1000**
LCR meter		**C, nF**	1	8.48	100.2	998.7
		ESR, Ω	2.5	3.6	0.29	0.04
VNA	**reflection**	**C, nF**	0.96	8.22	98.5	984.8
		ESR, Ω	*-2.61*	2.86	0.19	0.041
	shunt-through	**C, nF**	1.25	8.74	99.3	991
		ESR, Ω	288	7.8	0.28	0.04
	series-through	**C, nF**	0.98	8.29	99	1200
		ESR, Ω	*-5.9*	2.86	0.59	0.39

G. Effect of connectors parasitics on capacitor parameters measurements

It is interesting to observe how uncalibrated connectors parasitics affect measurement results. Measured $|Z_c|$ and ESR versus frequency before (using (2), (4) and (6)) and after (using (1), (3) and (5)) de-embedding is depicted in Fig.10. As it can be seen the reflection and the series-through techniques are very sensitive to the connectors parasitics in MHz range for both ESR and $|Z_c|$ (Fig.10(a,b,e,f)). If the connector parasitics are not removed from the measurement results using (1) and (5), then measured f_{res} is noticeably lower, but $|Z_c|$ is noticeably higher (for f above f_{res}) than it is. This is mainly due to parasitic inductance of the connectors used.

$|Z_c|$ and f_{res} measured using the shunt-through technique is much less affected due to the connector parasitics. For f up to several hundreds of MHz effect of the connector's parasitics is negligible (Fig.10(c)). However ESR measured using the shunt-through technique is affected much higher due to connectors parasitics for f higher than several tens of MHz (Fig.10(d)).

III. CALCULATION OF THE CAPACITOR PARAMETERS MEASUREMENT ERRORS

Since capacitor Z_c, ESR and other parameters are indirect measurement results then for measurement error calculation partial differentiation method [12] can be used. When partial errors are known, indirect measurement absolute error (ΔY) can be calculated

$$\Delta Y = \sqrt{\sum_{i=1}^{K}\left(\frac{\partial Y}{\partial x_i}\Delta x_i\right)^2}, \qquad (9)$$

and relative error

$$\varepsilon = 100\frac{\Delta Y}{Y}, \ \%, \qquad (10)$$

where $\partial Y / \partial x_i$ is partial derivative with respect to argument x_i; Δx_i is x_i argument absolute error.

- For the reflection technique capacitor impedance is the function of 6 arguments (according to (1) and (2)): $x_1=|S_{11}|$, $x_2=\varphi$; $x_3=|Z_{11s}|$, $x_4=\varphi_s$, $x_5=|Y_{11p}|$, $x_4=\varphi_p$,

where $S_{11}= |S_{11}|e^{j\varphi}$, $Z_{11s}=|Z_{11s}|e^{j\varphi_s}$, $Y_{11p}==|Y_{11p}|e^{j\varphi_p}$.

- For the shunt-through technique capacitor impedance is the function of 10 arguments (according to (3)): $x_1=|S_{21}|$, $x_2=\varphi$; $x_3=|Z_{11s1}|$, $x_4=\varphi_{s1}$, $x_5=|Y_{11p1}|$, $x_6=\varphi_{p1}$, $x_7=|Z_{11s2}|$, $x_8=\varphi_{s2}$, $x_9=|Y_{11p2}|$, $x_{10}=\varphi_{p2}$.
- For the series-through technique capacitor impedance is the function of 10 arguments (according to (5)): $x_1=|S_{21}|$, $x_2=\varphi$; $x_3=|Z_{11s1}|$, $x_4=\varphi_{s1}$, $x_5=|Y_{11p1}|$, $x_6=\varphi_{p1}$, $x_7=|Z_{11s2}|$, $x_8=\varphi_{s2}$, $x_9=|Y_{11p2}|$, $x_{10}=\varphi_{p2}$.

Absolute measurement errors of the main capacitor parameters can be calculated using the following expressions:

$$\Delta ESR = \Delta \operatorname{Re}\{Z_c\} = \sqrt{\sum_{i=1}^{K}\left(\frac{\partial \operatorname{Re}\{Z_c\}}{\partial x_i}\Delta x_i\right)^2} \qquad (11)$$

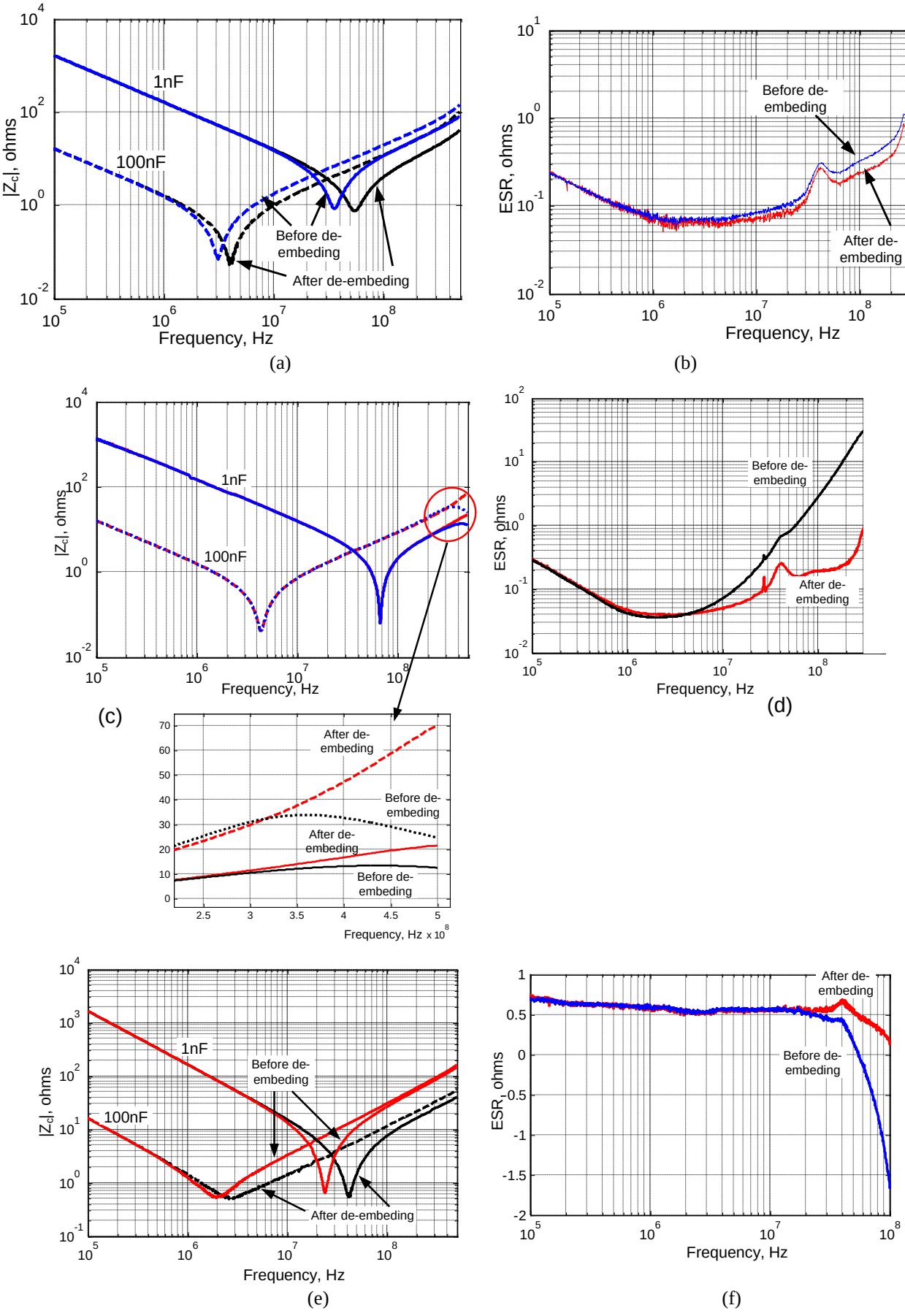


Fig. 10. Measured $|Z_c|$ and ESR versus frequency before and after de-embedding: (a), (b) for the reflection technique; (c), (d) for the shunt-through technique; (e), (f) for the series-through technique. Note: ESR versus frequency before and after de-embedding (b,d,f) is shown only for 100nF capacitor.

$$\Delta \operatorname{Im}\{Z_c\} = \sqrt{\sum_{i=1}^{K}\left(\frac{\partial \operatorname{Im}\{Z_c\}}{\partial x_i}\Delta x_i\right)^2} \quad (12)$$

$$\Delta |Z_c| = \sqrt{\left(\frac{\operatorname{Re}\{Z_c\}}{|Z_c|}\Delta \operatorname{Re}\{Z_c\}\right)^2 + \left(\frac{\operatorname{Im}\{Z_c\}}{|Z_c|}\Delta \operatorname{Im}\{Z_c\}\right)^2} \quad (13)$$

$$\Delta C = \frac{\Delta \operatorname{Im}\{Z_c\}}{2\pi f_1 (\operatorname{Im}\{Z_c\})^2} \quad . \quad (14)$$

If capacitor impedance is calculated using approximate expressions (2), (4) and (6), then more simple expressions for measurement error can be used. For the reflection technique the approximate expressions are

$$\Delta ESR = \Delta \operatorname{Re}\{Z_c\} \approx \sqrt{\left(\frac{\partial \operatorname{Re}\{Z_c\}}{\partial |S_{11}|}\Delta |S_{11}|\right)^2 + \left(\frac{\partial \operatorname{Re}\{Z_c\}}{\partial |\varphi|}\Delta\varphi\right)^2} \quad (15)$$

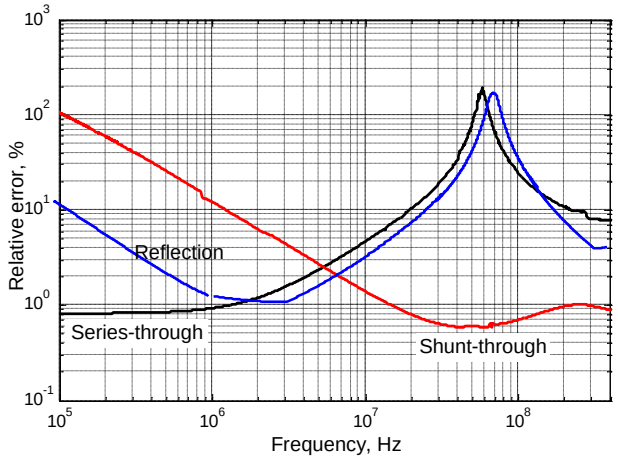
$$\Delta \operatorname{Im}\{Z_c\} \approx \sqrt{\left(\frac{\partial \operatorname{Im}\{Z_c\}}{\partial |S_{11}|}\Delta |S_{11}|\right)^2 + \left(\frac{\partial \operatorname{Im}\{Z_c\}}{\partial |\varphi|}\Delta\varphi\right)^2} \quad (16)$$

For shunt-through and series-through technique the approximate expressions are

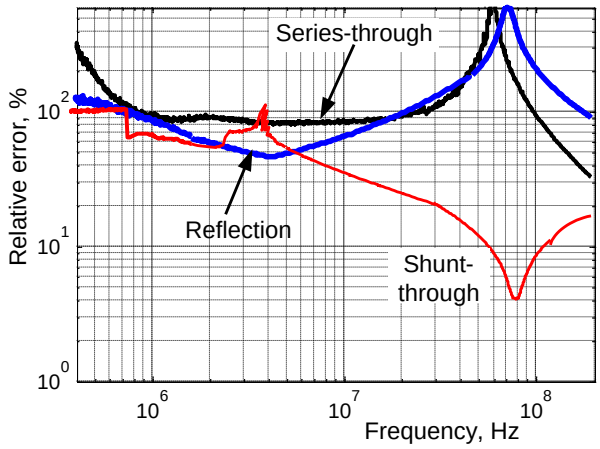
$$\Delta ESR = \Delta \operatorname{Re}\{Z_c\} \approx \sqrt{\left(\frac{\partial \operatorname{Re}\{Z_c\}}{\partial |S_{21}|}\Delta |S_{21}|\right)^2 + \left(\frac{\partial \operatorname{Re}\{Z_c\}}{\partial |\varphi|}\Delta\varphi\right)^2} \quad (17)$$

$$\Delta \operatorname{Im}\{Z_c\} \approx \sqrt{\left(\frac{\partial \operatorname{Im}\{Z_c\}}{\partial |S_{21}|}\Delta |S_{21}|\right)^2 + \left(\frac{\partial \operatorname{Im}\{Z_c\}}{\partial |\varphi|}\Delta\varphi\right)^2} \quad . \quad (18)$$

It should be noted that in VNA technical data-sheet manufacturer shows transmission and reflection coefficients magnitude errors (in *dB*) and phase errors (in degrees) [11]. But in the expressions described above, $\Delta|S_{11}|$ and $\Delta|S_{21}|$ are not in *dB*, and φ is in radians. For example, if it is needed to calculate absolute error $\Delta|S_{11}|$, then the following expression

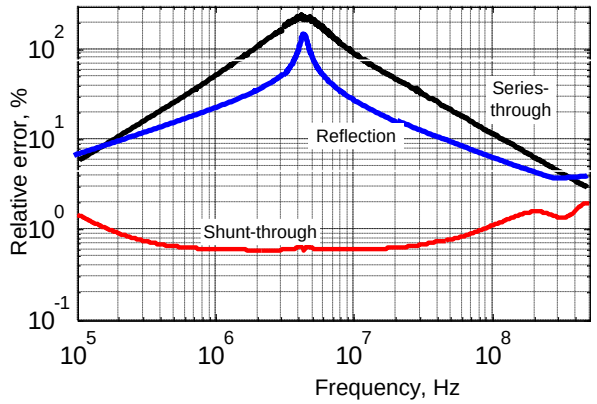


(a)

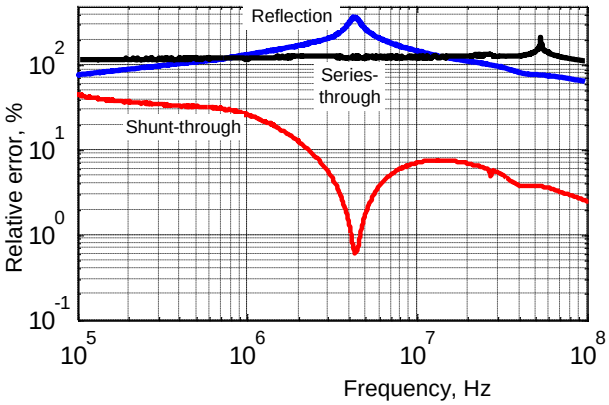


(b)

Fig. 11. Calculated relative measurement errors of 1nF capacitor |Zc| (a) and ESR (b) versus frequency for all the three measurement techniques.



(a)



(b)

Fig. 12. Calculated relative measurement errors of 100nF capacitor |Zc| (a) and ESR (b) versus frequency for all the three measurement techniques.

should be used

$$\Delta|S_{11}| = |S_{11}|\left(10^{(\Delta|S_{11}|_{dB})/20} - 1\right), \quad (19)$$

where $\Delta|S_{11}|_{dB}$ is reflection coefficient measurement uncertainty in *dB* from VNA data-sheet. In its turn, phase absolute error (in radians) from VNA data-sheet can be calculated as follows

$$\Delta\varphi = \pi\frac{\varphi_{\mathbf{deg}}}{180}, \quad (20)$$

From VNA ZVRE data-sheet [11] it can be deduced that in frequency range 20kHz – 4GHz: $\Delta|S_{11}|_{dB}$<0.4dB and $\Delta\varphi_{deg}$<3° (when $|S_{11}|$ is in range -15dB – +3dB); $\Delta|S_{11}|_{dB}$<1dB and $\Delta\varphi_{deg}$<6° (when $|S_{11}|$ is in range -25dB – -15dB). Accuracy of transmission measurements: in the frequency range 20kHz – 300kHz: $\Delta|S_{21}|_{dB}$<0.2dB and $\Delta\varphi_{deg}$<2° (when $|S_{21}|$ is in range -20dB – +3dB); $\Delta|S_{21}|_{dB}$<0.5dB and $\Delta\varphi_{deg}$<4° (when $|S_{21}|$ is in range -20dB – -30dB); in the frequency range 300 kHz to 1 MHz: $\Delta|S_{21}|_{dB}$<0.1 dB or $\Delta\varphi_{deg}$<1°; in the frequency range 1MHz to 4 GHz: $\Delta|S_{21}|_{dB}$<0.05 dB or $\Delta\varphi_{deg}$<0.4°. This data will be used for capacitors measurement error calculation. Calculated relative measurement errors of ESR and $|Z_c|$ versus frequency for all the three measurement techniques are shown in Fig.11 and Fig.12. In the next section the measurement error analysis will be presented.

IV. Analysis of the Measurement Results and Measurement Errors

Capacitors impedance magnitude $|Z_c|$ measurement analysis

Results presented in Fig.11 and Fig.12 clearly show that measurement error of capacitors $|Z_c|$ depends not only on capacitor impedance magnitude (and frequency) but also on the measurement technique used.

For the shunt-through technique, the lower $|Z_c|$ is, the lower measurement error is. The best measurement accuracy (below 1%) can be achieved at f_{res} and in the vicinity of it using the shunt-through technique. When using the technique for $|Z_c|$>30Ω, the measurement error exceeds several %. In kΩ range the shunt-through technique exhibits very poor accuracy (region A in Fig, 8(a)), because $|S_{21}|$ is very close to 0dB.

As opposed to the shunt-through technique, series-through technique exhibits lower measurement error for higher $|Z_c|$. Measurement accuracy at f_{res} and in the vicinity of it using the series-through technique is very poor, because $|S_{21}|$ is very close to 0dB. However when $|Z_c|$>30Ω, the measurement error is lower than several %, and in kΩ range the series-through technique exhibits very good accuracy.

The reflection technique should not be used for capacitor parameters measurements because it can give accurate results only when capacitor $|Z_c|$ slightly differs from 50Ω (Fig.11(a)).

Capacitors resonant frequency f_{res} measurement analysis

The best suited technique for capacitor f_{res} measurement is the shunt-through technique, because measurement accuracy of capacitor impedance at f_{res} and in the vicinity of it is very good (Fig.11(a) and Fig.12(a)) and this technique is less sensitive to the connectors parasitics.

Capacitors capacitance C measurement analysis

C measurement accuracy highly depends on the measurement technique used. Since *C* is calculated using (7) and capacitor $|\mathrm{Im}\{Z_c\}|$ at f_1 can vary from hundreds of mΩ to kΩ (depends on capacitor *C*), then for *C* measurements it is better to use the shunt-through technique when *C*>50nF. However, when *C*<50nF, then the series-through technique is the best choice for *C* measurements. This can also be deduced from the Table 1. For capacitors with nominal capacitance values of 100nF and 1000nF, the lowest difference between *C* values measured using VNA and high-precision LCR meter is achieved using the shunt-through technique (the difference below 1%). However, for capacitors with nominal capacitance of 1nF and 10nF, the lowest difference between *C* values measured using VNA and high-precision LCR meter is achieved using the series-through technique (the difference below 3%). The reflection technique should not be used for capacitance measurements.

Capacitors ESR measurement analysis

According to the results presented in Table 1 and Fig.8(b) - Fig.12(b) the most problematic parameter to measure accurately with VNA is capacitor ESR. Only the shunt-through technique should be used for accurate ESR measurements. However there are some limitations. ESR can accurately be measured only at f_{res} and in vicinity of it (ESR measurement error below 3% can be achieved). For frequencies which are much higher or lower than f_{res} the measurement accuracy can be very poor, even using the shunt-through technique, because when capacitor reactance is much higher than its ESR, then the measurement error can increase enormously. Moreover accurate measurements of ESR of capacitors with *C* lower than several tens of nF are impossible: series-through and reflection techniques for capacitors with *C* below 1nF, can give even negative ESR which is completely unrealistic (Table 1). Even the shunt-through technique gives very erroneous ESR measurement results for small-capacitance capacitors. So ESR measurements of small-capacitance capacitors (below 10nF) using VNA is completely useless, because of very poor measurement accuracy.

V. Conclusion

Vector network analyzers can be very useful for accurate capacitor impedance, resonant frequency, capacitance and ESL measurements in broad frequency range if proper measurement technique and de-embedding is used. Measurement accuracy of several % and even lower can be achieved using proper measurement technique. Measurement error of capacitors

impedance and ESR depends not only on capacitor impedance magnitude (and frequency) but also on the measurement technique used.

The best suited technique for capacitor f_{res} measurement is the shunt-through technique, because measurement accuracy of capacitor impedance at f_{res} and in the vicinity of it is very good (even below 1%) and this technique is less sensitive to the experimental setup parasitics. For capacitor impedance measurements in broad frequency range (when capacitor $|Z_c|$ can change from mΩ to kΩ) it is better to use combination of the shunt-through and series-through techniques: when $|Z_c|$ is below several tens of ohms, then it is better to use shunt-through technique; when $|Z_c|$ is above several tens of ohms, then it is better to use series-through technique. The reflection technique should not be used for capacitor parameters measurements because, firstly, it can give accurate results only when capacitor $|Z_c|$ slightly differs from 50Ω and secondly, it is very sensitive to uncalibrated connectors parasitics.

The most problematic parameter to measure accurately with VNA is capacitor ESR. Only shunt-through technique should be used for accurate ESR measurements. However there are some limitations. ESR can accurately be measured only at f_{res} and in vicinity of it (ESR measurement error below 3% can be achieved). For frequencies which are much higher or lower than f_{res} the measurement accuracy can be very poor, even using shunt-through technique, because when capacitor reactance is much higher than its ESR, then the measurement error can increase enormously. Accurate measurements of ESR of capacitors with C lower than several tens of nF are impossible. ESR measurements of small-capacitance capacitors using VNA is completely useless.

Overall it can be concluded that VNA can substitute more-expensive impedance analyzers for accurate capacitor parameters measurements in broad frequency range with some limitations in terms of ESR measurements.

REFERENCES

[1] Kazunari Okada, Toshimasa Sekino. Impedance Measurement Handbook. Agilent Technologies, 2006.

[2] A.M.R. Amarel, A.J.M. Cordoso, "An economic offline technique for estimating equivalent circuit of aluminum electrolytic capacitors," *IEEE Transactions on Instrumentation and Measurement*, vol.57, No.12, pp. 2697 – 2710, Dec. 2008.

[3] R. Dosoudil, „Determination of permeability from impedance measurement using vector network analyzer," *Journal of electrical engineering*, vol. 63, No.7, pp. 97-101, Oct. 2012.

[4] Satish Prabhakaran, Charles Sullivan, "Impedance-analyzer measurements of high-frequency power passives: Techniques for high power and low impedance," *In proc. IEEE Industry Applications Conference*, Pittsburgh, PA, USA, 2002, pp. 1360 – 1367.

[5] William B. Kuhn, Adam P. Boutz, "Measuring and reporting high quality factors of inductors using vector network analyzers," *IEEE Transactions on Microwave Theory and Techniques*, vol.58, No.4, pp. 1046-1055, Apr. 2010.

[6] Advanced Impedance Measurement Capability of the RF I-V Method Compared to the Network Analysis Method. Agilent Application Note No. 1369-2.

[7] Ultra Low Impedance Measurements Using 2-Port Measurements. Agilent Application Note, 2004.

[8] L.D. Smith, D. Hockanson, "Distributed SPICE circuit model for ceramic capacitors," *In. proc. Electronic Components and Technology Conference*, 2001, pp. 523-528.

[9] Rosa M. García Salvador, José A. Gázquez Parra, Nuria Novas Castellano, "Characterization and Modeling of High-Value Inductors in ELF Band Using a Vector Network Analyzer," *IEEE Transactions on Instrumentation and Measurement*, vol.62, No.4, pp. 415-423, Feb. 2013.

[10] Gregory L. Amorese, "Reducing ESR measurement errors," *Microwaves & RF*, vol. 42, No.4, 2003.

[11] Vector Network Analyzer Family ZVR datasheet. Rohde and Schwarz.

[12] John R. Taylor, *An Introduction To Error Analysis: The Study Of Uncertainties In Physical Measurements*. Sausalito, University Science Books, 1996.

Data-Driven Gradient Descent Direct Adaptive Control for Discrete-Time Nonlinear SISO Systems

Igor R. Krcmar, Milorad M. Bozic, and Petar S. Maric

***Abstract*—A novel data-driven gradient descent (GD) adaptive controller, for discrete-time single-input and single output (SISO) systems, is presented. The controller operates as the least mean squares (LMS) algorithm, applied to a nonlinear system with feedback. Normalization of the learning rate parameter provides robustness of the overall system to modeling errors and environment nonstationarity. Convergence analysis reveals that the controller forces tracking error to zero, with bounded control signal and the controller weights. The experiments on the benchmark systems support the analysis.**

***Index Terms*—Data-Driven Controller, Gradient Descent Algorithms, Direct Adaptive Control, Nonlinear SISO Systems, Discrete Time.**

Original Research Paper
DOI: 10.7251/ELS1418039K

I. Introduction

Control means decision making. In order to decide well, one has to have a goal, as well as, information on a controlled process and process environment. Within automatic control setup, the goal specifies desired behavior of the overall system. Information, regarding the process and influence of the environment, is given by the process model and/or process measurements. In many automatic control applications process model is not available. Instead, process measurements contain information on the process behavior. Therefore, many control algorithms were developed to cope with this situation. Usually, they are referred to as data-driven or model-free. This emphasizes situation where control is based, mainly, on the available measurement data. In these situations, concepts like feedback linearization [1], regardless their firm theoretical foundation and good performance, are not applicable. If we look back to the past, probably, the first data-driven control procedure is one proposed by Ziegler and Nichols for tuning proportional-integral-derivative (PID) controllers. However, many data-driven control algorithms were reported in the literature, till now. These are virtual reference feedback tuning (VRFT), iterative feedback tuning (IFT) and data-driven model-free adaptive control (MFAC), to mention a few. The VRFT is reference model adaptive control technique. Within VRFT framework, it is assumed that controller structure is known a-priori and controller parameters are determined by system identification procedure, using virtual reference signal. It is designed for discrete-time single-input and single output (SISO) linear time invariant (LTI) systems. The VRFT procedure may produce controller that results in a non stable closed loop system [2,3]. The IFT, also proposed for an unknown discrete-time SISO LTI system, determines controller parameters through iterative gradient-based local search procedure. Computation of the gradient of criterion function, with respect to the controller parameters, is based on available input/output (I/O) measurement data of the controlled plant [4]. The MFAC approach, in opposition to the VRFT and IFT, was developed for a class of discrete-time nonlinear systems. The main characteristic of the MFAC is utilization of local dynamic linearization data models. The models are computed along the dynamic operation points of the closed-loop system. Computation is performed by the dynamic linearization techniques, with a pseudo-partial derivative concept, based on the real-time I/O measurements of the controlled plant [3]. The above mentioned algorithms, i.e. VRFT, IFT and data-driven MFAC, are direct adaptive control algorithms, as they tune controller without prior plant identification procedure. Further, in applications rich with measurement data, regarding processes with high complexity, neural networks (NNs), with their ability to approximate functions through training procedure [5, 1], might be a good solution. Specially, in nonlinear control tasks, NNs are employed as nonlinear process models and/or nonlinear controllers [6, 7, 1]. However, design of an appropriate NN, meaning choice of an NN structure, input and output vectors, training set and training algorithm, might be a difficult task. Within signal prediction applications, we design a model that should produce, on its output, the desired signal. Usually, the model is parameterized and model parameters are tuned in the real-time based on the incoming process measurements, i.e.

Manuscript received 9 May 2014. Accepted for publication 11 June 2014.

I. R. Krcmar is with the Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, Bosnia and Herzegovina (phone: +387-65-927-212; fax: +387-51-211-408; e-mail: ikrcmar@etfbl.net).

M. M. Bozic is with the Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, Bosnia and Herzegovina (e-mail: mbozic@etfbl.net).

P. S. Maric is with the Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, Bosnia and Herzegovina (e-mail: pmaric@etfbl.net).

adaptively, in order to minimize a certain criterion. The criterion is a function of the error signal, the signal that represents the difference between some desired signal and the model output [8]. Similar setup one can find in the automatic control reference tracking tasks. In the reference tracking tasks, the controller is tuned, so the overall system output tracks the reference signal, as close as possible. Thus, adaptive signal prediction and reference tracking control are similar problems, as they can share the same criterion to be minimized through the parameters correction procedure. For the given parameterization of the process model, different procedures can be applied to achieve optimal value of the model parameters, regarding the imposed criterion. The gradient descent (GD) approach performs search for the optimal value of the parameters, in the parameter space, as it corrects the parameter values in direction opposite to the gradient of the criterion. A common choice for the criterion is a function of the squared error. Models, applied in the signal prediction tasks, with the feedback from the output to the input, introduce nonlinearity into the signal prediction algorithm. Algorithms, based on such models, are referred to as recurrent [8]. The least mean squares (LMS) algorithm, as a member of the GD class, is a simple, yet most frequently used algorithm for signal prediction and system identification. However, it might have a poor performance, when applied to a nonlinear signal and in varying environment [10]. To that cause, a class of normalized algorithms has been developed [12, 8, 9, 10]. They successfully cope with process nonlinearity and nonstationarity, thus provides faster convergence and higher accuracy. The proposed data-driven GD adaptive controller relies on the ideas established within signal prediction framework. It is direct adaptive controller, designed for discrete-time SISO nonlinear system to perform reference tracking. The controller is parameterized, thus provides linear combination of the reference signal and the output signal discrete-time samples. The parameters are tuned, in the real-time, through the GD procedure, to minimize a certain criterion. The criterion is square of the instantaneous output error, therefore the algorithm operates as the LMS algorithm applied to a nonlinear system with feedback. In order to cope with modeling errors and environment nonstationarity, the learning rate is modified as in normalized algorithms [9, 10, 12]. The proposed algorithm has ability of a constant disturbance rejection. Convergence analysis of the proposed algorithm can be performed, following the approach given in [11, 10]. The rest of the paper is organized as follows. In the Section 2, problem formulation and the controller design procedure are given. The Section 3 gives convergence analysis, as well as, analysis of the disturbance rejection ability of the proposed algorithm. The Section 4 contains experimental verification of the algorithm, while the Section 5 concludes the paper.

II. Problem formulation and controller design procedure

It is assumed that the controlled plant is discrete-time nonlinear SISO system, described by the equation

$$\begin{aligned} y(k+1) &= f(y(k), y(k-1), \ldots, y(k-n_y); \\ & u(k), u(k-1), \ldots, u(k-n_u)) = f(k), \end{aligned} \tag{1}$$

where $y(k) \in R$ and $u(k) \in R$ are the system output and input, respectively, n_y and n_u are the unknown orders, and k denotes discrete-time instant. Further, $f(\cdot) \in R$ is some nonlinear function with continuous partial derivatives with respect to its arguments. The controller parameterization is given by the equation

$$\begin{aligned} u(k) &= w_1(k) r(k) + w_2(k) r(k-1) + \ldots \\ &+ w_{n_c}(k) r(k-n_c+1) + w_{n_c+1}(k) y(k) \\ &+ w_{n_c+2}(k) y(k-1) + \ldots w_{2n_c}(k) y(k-n_c+1), \end{aligned} \tag{2}$$

where $r(k) \in R$ denotes the reference signal, $w_i(k); i = 1, 2, \ldots, 2n_c$ are controller parameters, in what follows called weights, and n_c is the controller order. If we introduce the following vectors $\mathbf{w}(k) = [w_1(k), w_2(k), \ldots, w_{2n_c}(k)]^T$, $\boldsymbol{\varphi}_r(k) = [r(k), \ldots, r(k-n_c+1)]^T$, $\boldsymbol{\varphi}_y(k) = [y(k), \ldots, y(k-n_c+1)]^T$, and $\boldsymbol{\varphi}(k) = [\boldsymbol{\varphi}_r^T(k), \boldsymbol{\varphi}_y^T(k)]^T$, where $(\cdot)^T$ denotes vector transpose, the equation (2) can be written in compact form as

$$u(k) = \boldsymbol{\varphi}^T(k) \mathbf{w}(k). \tag{3}$$

The cost function, as the system output should track the reference signal, is defined by

$$J(k) = (1/2) e^2(k), \tag{4}$$

where $e(k)$ denotes instantaneous error at the system output, and it is given as follows

$$e(k) = r(k) - y(k). \tag{5}$$

The weight vector is updated in the real-time according to the equation

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta \mathbf{w}(k). \tag{6}$$

The weights are corrected, through the GD procedure, thus minimizing the cost function (4). Therefore, the weights correction vector is given by

$$\Delta \mathbf{w}(k) = -\eta(k) \nabla_{\mathbf{w}(k)} J(k), \tag{7}$$

where $\eta(k)$ denotes the learning rate and $\nabla_{\mathbf{w}(k)}(\cdot)$ denotes gradient of a scalar function with respect to the weight vector $\mathbf{w}(k)$. In order to compute the weight correction vector (7),

first we have to compute the gradient of the cost function

$$\nabla_{\mathbf{w}(k)} J(k) = \nabla_{\mathbf{w}(k)} \left(\frac{1}{2} e^2(k) \right) = e(k) \nabla_{\mathbf{w}(k)} e(k). \tag{8}$$

To determine the gradient on the right hand side of the equations (8), we have to compute partial derivatives of the error $e(k)$ with respect to the weight $w_i(k)$, where $1 \le i \le 2n_c$

$$\frac{\partial e(k)}{\partial w_i(k)} = -\frac{\partial y(k)}{\partial w_i(k)} = -\sum_{j=1}^{n_c} \frac{\partial f(k)}{\partial y(k-j)} \frac{\partial y(k-j)}{\partial w_i(k)} - \sum_{j=1}^{n_c} \frac{\partial f(k)}{\partial u(k-j)} \frac{\partial u(k-j)}{\partial w_i(k)}. \tag{9}$$

Now we introduce

$$\alpha_j(k) = \frac{\partial f(k)}{\partial y(k-j)}; \beta_j(k) = \frac{\partial f(k)}{\partial u(k-j)}. \tag{10}$$

Combining (9) and (10) we have

$$\frac{\partial e(k)}{\partial w_i(k)} = -\frac{\partial y(k)}{\partial w_i(k)} = -\sum_{j=1}^{n_c} \alpha_j(k) \frac{\partial y(k-j)}{\partial w_i(k)} - \sum_{j=1}^{n_c} \beta_j \frac{\partial u(k-j)}{\partial w_i(k)}. \tag{11}$$

Computation of the partial derivatives, on the right hand side of (11), will be carried out under the assumption of slowly changing weights [13, 8]. Therefore, the following holds

$$\frac{\partial y(k-j)}{\partial w_i(k)} \approx \frac{\partial y(k-j)}{\partial w_i(k-j)}; \frac{\partial u(k-j)}{\partial w_i(k)} \approx \frac{\partial u(k-j)}{\partial w_i(k-j)}. \tag{12}$$

At this point, it is worth noting that the coefficients $\alpha_j(k)$ and $\beta_j(k)$ are time varying and they define the linearized model of the nonlinear system (1), computed along the trajectory that system (1) follows as it advances through time. As we assume the system (1) is unknown, the coefficients $\alpha_j(k)$ and $\beta_j(k)$ can not be determined. Even if we can determine them, from the available measurement data, to represent dominant dynamics of the system (1), the modeling error is still present. To remedy the situation, normalization of the learning rate parameter, as proposed in [8, 9, 10, 11], will be applied. Therefore, we can continue the algorithm derivation by taking into account the assumptions

$$\alpha_1(k) = \alpha, \alpha_j(k) \equiv 0; j = 2, 3, \ldots, n_c, \tag{13}$$

and

$$\beta_1(k) = \beta, \beta_j(k) \equiv 0; j = 2, 3, \ldots, n_c. \tag{14}$$

Now, (11) can be rewritten as follows

$$\frac{\partial e(k)}{\partial w_i(k)} = -\frac{\partial y(k)}{\partial w_i(k)} = -\alpha \frac{\partial y(k-1)}{\partial w_i(k-1)} - \beta \frac{\partial u(k-1)}{\partial w_i(k-1)}. \tag{15}$$

In order to compute partial derivatives $\partial u(k-1)/\partial w_i(k-1)$ we have to refer to the equation (2). Now, we can distinguish two cases. In the first one, the weight $w_i(k-1)$ is connected to the reference signal. Thus, we have

$$\frac{\partial u(k-1)}{\partial w_i(k-1)} = r(k-i) + \sum_{j=1}^{n_c} w_{j+n_c}(k-1) \frac{\partial y(k-j)}{\partial w_i(k-1)} \approx r(k-i) + \sum_{j=1}^{n_c} w_{j+n_c}(k-1) \frac{\partial y(k-j)}{\partial w_i(k-j)}, \tag{16}$$

and in the case the weight $w_i(k-1)$ is connected to the output signal, we have

$$\frac{\partial u(k-1)}{\partial w_i(k-1)} = y(k-i) + \sum_{j=1}^{n_c} w_{j+n_c}(k-1) \frac{\partial y(k-j)}{\partial w_i(k-1)} \approx y(k-i) + \sum_{j=1}^{n_c} w_{j+n_c}(k-1) \frac{\partial y(k-j)}{\partial w_i(k-j)}. \tag{17}$$

Let us introduce the following

$$\pi_i(k) = \frac{\partial y(k)}{\partial w_i(k)}; i = 1, 2, \ldots, 2n_c, \tag{18}$$

and

$$\boldsymbol{\pi}(k) = [\pi_1(k), \pi_2(k), \ldots, \pi_{n_c}(k)]^T, \tag{19}$$

where $\boldsymbol{\pi}(k)$ denotes gradient, with respect to the weight vector, at the system output. Now, the weight update equation becomes

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta(k) e(k) \boldsymbol{\pi}(k), \tag{20}$$

with

$$\boldsymbol{\pi}(k) = \alpha \boldsymbol{\pi}(k-1) + \beta \left(\boldsymbol{\varphi}(k-1) + \sum_{j=1}^{n_c} w_{j+n_c} \boldsymbol{\pi}(k-j) \right). \tag{21}$$

To achieve normalization of the learning rate parameter, first we shall expand the error term (5) into a Taylor series [9]

$$e(k+1) = e(k) + \sum_{i=1}^{2n_c} \frac{\partial e(k)}{\partial w_i(k)} \Delta w_i(k) + \frac{1}{2!} \times \times \sum_{i=1}^{2n_c} \sum_{j=1}^{2n_c} \frac{\partial^2 e(k)}{\partial w_i(k) \partial w_j(k)} \Delta w_i(k) \Delta w_j(k) + \ldots. \tag{22}$$

From (5) and (18) we have

$$\frac{\partial e(k)}{\partial w_i(k)} = -\frac{\partial y(k)}{\partial w_i(k)} = \pi_i(k), \tag{23}$$

while from the weight update we have

$$\Delta w_i(k) = \eta(k) e(k) \pi_i(k). \tag{24}$$

If, for the time being, we neglect higher order terms in the Taylor series expansion (22), as well as, influence of the modeling errors, which comes through the local

gradients $\pi_i(k)$, then the error term becomes

$$e(k+1) = e(k)\left[1-\eta(k)\left\|\boldsymbol{\pi}(k)\right\|_2^2\right], \tag{25}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. The error term in (25) equals zero for

$$\eta(k) = \frac{1}{\left\|\boldsymbol{\pi}(k)\right\|_2^2}. \tag{26}$$

The expression (26) has to be corrected to compensate for the higher order terms in (22), as well as, for modeling errors, thus we have

$$\eta(k) = \frac{1}{C+\left\|\boldsymbol{\pi}(k)\right\|_2^2}. \tag{27}$$

Even though the learning rate (27) provides an optimal solution, it might yield a large bandwidth of the overall system, which in turn may result with the system instability. Therefore, the following learning rate is adjusted according to

$$\eta(k) = \frac{\mu}{C+\left\|\boldsymbol{\pi}(k)\right\|_2^2}, \tag{28}$$

where $0<\mu<1$. Thus, equations (2), (20), (21), and (28) define the proposed algorithm. From the derivation of the algorithm, it is clear that the coefficients α and β, given by (13) and (14), the controller order n_c, the constant C, introduced in (28), as well as, the constant μ are the designed parameters of the algorithm.

III. Convergence analysis and disturbance rejection

The goal of the analysis is to examine whether, the overall system can track the reference signal $r(k)=r^*$, so both signals $u(k)$ and $y(k)$ are bounded. Therefore, the following should hold

$$\lim_{k\to\infty} e(k) = 0 \Rightarrow \lim_{k\to\infty} y(k) = r^*, \tag{29}$$

with $|u(k)|<L_u$ and $|y(k)|<L_y$, where $|\cdot|$ denotes an absolute value, while L_u and L_y are some positive constants. In order to (29) holds, the mapping (25) has to be a contraction, i.e.

$$\left|1-\eta(k)\left\|\boldsymbol{\pi}(k)\right\|_2^2\right| < 1 \tag{30}$$

If inequality (30) holds the equation (25) defines fixed point iteration, thus (29) holds. The inequality (30) can be written as

$$-1 < 1-\eta(k)\left\|\boldsymbol{\pi}(k)\right\|_2^2 < 1. \tag{31}$$

Now, we substitute the learning rate $\eta(k)$, defined by the equation (28), into the inequalities (31), which yields the lower bound on the constant C

$$C > -\frac{2-\mu}{2}\left\|\boldsymbol{\pi}(k)\right\|_2^2. \tag{32}$$

Further, if we refer to the equation (3), we can write

$$\Delta u(k) = \Delta\boldsymbol{\varphi}^T(k)\mathbf{w}(k) + \boldsymbol{\varphi}^T(k)\Delta\mathbf{w}(k). \tag{33}$$

Now, from the weight update equation (20), we have

$$\lim_{k\to\infty} e(k) = 0 \Rightarrow \lim_{k\to\infty} \left\|\mathbf{w}(k)\right\| = 0. \tag{34}$$

To investigate behavior of $\Delta\boldsymbol{\varphi}(k)$ we proceed as follows

$$\begin{aligned} \left|y(k+1)-y(k)\right| &= \left|y(k+1)-r^*+r^*-y(k)\right| \\ &< \left|e(k+1)\right|+\left|e(k)\right| = \left|e(k)\right|\left(1+\varepsilon\right), \end{aligned} \tag{35}$$

where $\varepsilon = \left|1-\eta(k)\left\|\boldsymbol{\pi}(k)\right\|_2^2\right| < 1$. Therefore, the following holds

$$\lim_{k\to\infty} e(k) = 0 \Rightarrow \lim_{k\to\infty} \left|y(k+1)-y(k)\right| = 0, \tag{36}$$

and, as $r(k)=r^*$, we have

$$\lim_{k\to\infty} e(k) = 0 \Rightarrow \lim_{k\to\infty} \left\|\Delta\boldsymbol{\varphi}(k)\right\| = 0. \tag{37}$$

From (33), (34), and (37) we conclude

$$\lim_{k\to\infty} e(k) = 0 \Rightarrow \lim_{k\to\infty} \Delta u(k) = 0, \tag{38}$$

which completes the convergence analysis. If the system is under influence of constant disturbances at its output, then equation (1) can be rewritten as

$$y(k+1) = f(k) + d, \tag{39}$$

where d denotes the disturbance. It is obvious that computation of the partial derivatives (15) is not affected by the disturbance, and the algorithm derivation holds in this case. However, the disturbance affects the error signal, thus it will appear in the Taylor series (22). Thus, if adjust the constant C to compensate the disturbance, the analysis holds.

IV. Simulations

In order to investigate performance of the proposed algorithm, the controller was applied, in a reference tracking task, on three benchmark processes. In all experiments the parameters in the equation (15) were set to $\alpha=0$ and $\beta=1$, while the learning rate parameter in (28) was set to $\mu=0.1$. Further, the order of the controller, n_c, was varied from 1 to 15, and values of the constant C were from the set $\{0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.5, 3, 4, 5, 10\}$. The averaged squared error was adopted as the performance measure.

A. Experiment 1

In the first experiment, the controller was applied to the process [7] given by the equation

$$y(k+1)=\sin(y(k))+u(k)(5+\cos(y(k)u(k))) \quad (40)$$

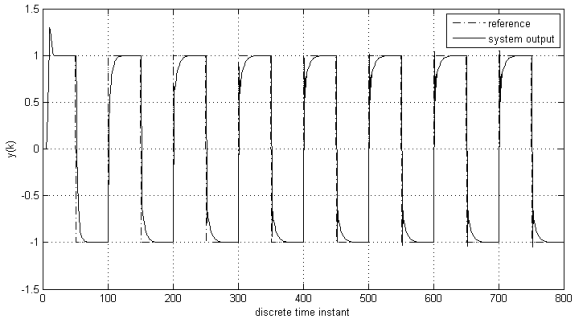


Fig. 1.a Tracking performance of the proposed algorithm in the Experiment 1.

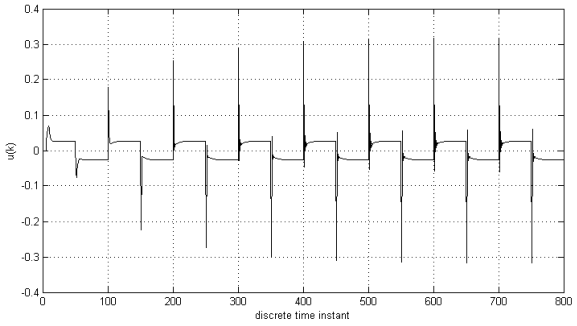


Fig. 1.b Control signal in the Experiment 1.

Figures 1.a to 1.c gives tracking performance, control signal, and error signal of the proposed algorithm, respectively, applied to the system (40), with parameters $C=0.5$ and $n_c=3$. Value of the performance measure was 0.1059.

Figure 1.d gives performance surface in the experiment 1. Minimum of the performance measure was 0.09166, and it was attained for $n_c=1$ and $C=3$.

B. Experiment 2

In the second experiment, the controller was applied to the process [6] given by the equation

$$y(k+1)=\frac{y(k)}{1+y^2(k)}+u^3(k) \quad (41)$$

Figures 2.a to 2.c shows tracking performance, control signal, and error signal of the proposed algorithm, respectively, applied to the system (41), with parameters $C=0.5$ and $n_c=3$. The performance measure was 0.13196.

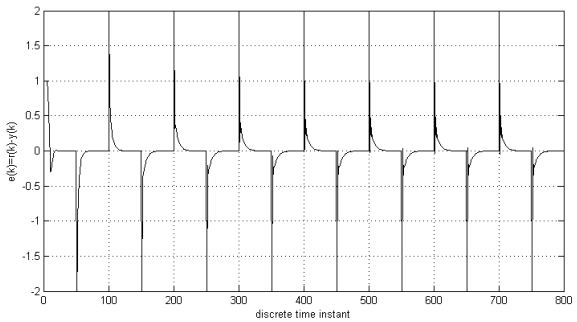


Fig. 1.c The error signal in the Experiment 1.

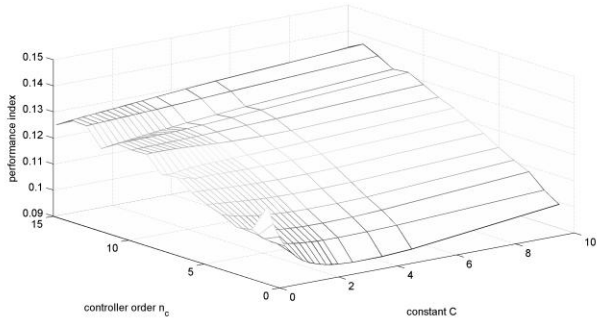


Fig. 1.d Performance index in the Experiment 1.

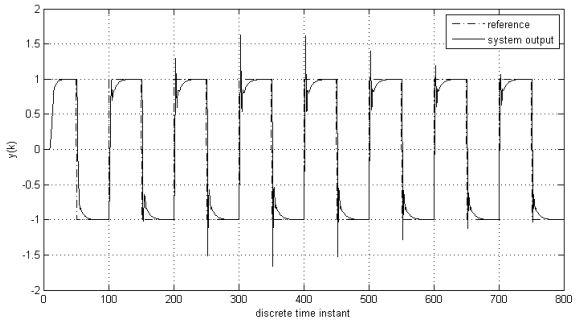


Fig. 2.a Tracking performance of the proposed algorithm in the Experiment 2.

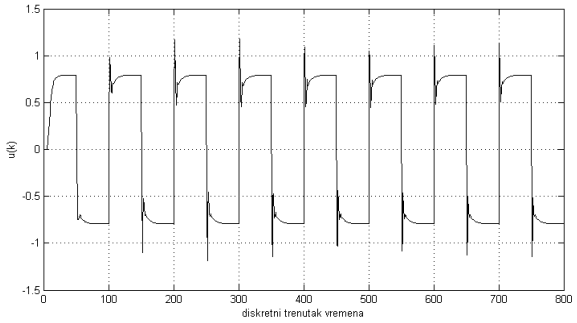


Fig. 2.b Control signal in the Experiment 2.

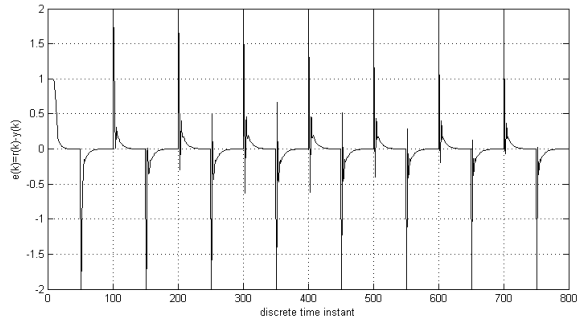


Fig. 2.c The error signal in the Experiment 2.

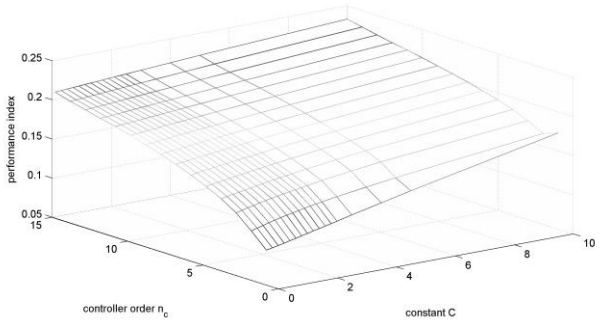


Fig. 2.d Performance index in the Experiment 2.

Figure 2.d gives performance surface in the experiment 2. A minimal value of the performance measure was 0.09296, and it was attained for $n_c = 1$ and $C = 0.1$.

C. Experiment 3

The third experiment was designed to investigate disturbance rejection property of the proposed algorithm. Due to that cause, the controller was applied to the process [6] given by the equation

$$y(k+1) = \frac{y(k)y(k-1)(y(k)+2.5)}{1+y^2(k)+y^2(k-1)} + u(k) \tag{42}$$

Also, at the discrete time instant $k = 370$, the step disturbance of the amplitude $d = 0.2$ was applied to the system output. In this case, a minimal value of the performance measure was 0.02527, and it was attained for $n_c = 1$ and $C = 0.1$. Figures 3.a to 3.c gives tracking performance, control signal, and error signal of the proposed algorithm, respectively, with parameters $C = 0.1$ and $n_c = 1$.

Figure 3.d shows a performance surface in the experiment 3. The minimal value of the performance measure was attained for $n_c = 1$ and $C = 0.1$.

V. Conclusions

The new data-driven GD adaptive controller, for discrete-time nonlinear SISO systems, has been proposed. It was designed for the reference tracking tasks. The controller has simple structure, i.e. it is linear combination of the reference signal and the output signal discrete-time samples. Parameters

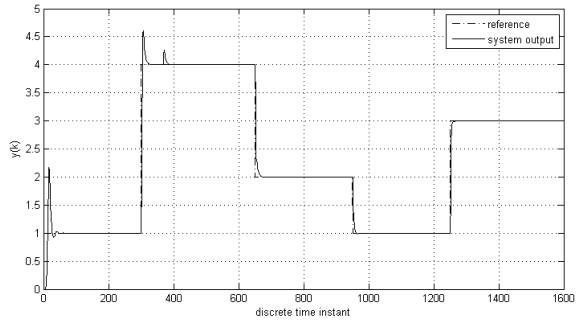


Fig. 3.a Tracking performance of the proposed algorithm in the Experiment 3.

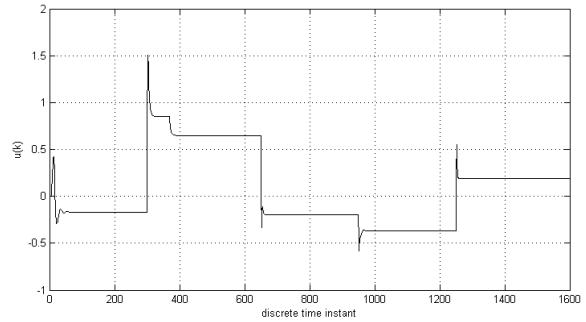


Fig. 3.b Control signal in the Experiment 3.

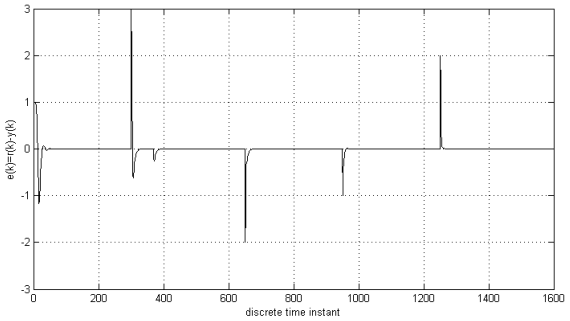


Fig. 3.c The error signal in the Experiment 3.

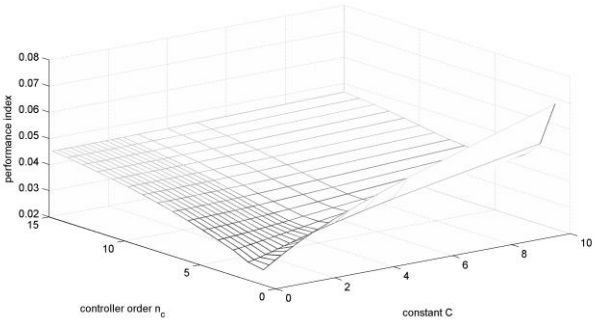


Fig. 3.d Performance index in the Experiment 3.

of the controller are updated in the real-time, as in the case of the LMS algorithm applied to a nonlinear system with feedback. The algorithm is robust to the modeling errors and environment nonstationarity, due to the learning rate parameter normalization. The convergence analysis has been provided. The controller forces tracking error to zero with bounded control signal and controller weights. The experiments, on the

benchmark systems, demonstrate performance of the proposed controller and support the analysis.

REFERENCES

[1] Jeffrey T. Spooner, Manfredi Maggiore, Raul Ordonez, and Kevin M. Passino, *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*, New York: John Wiley & Sons, Inc., 2002.

[2] M. C. Campi, A. Lecchini, and S. M. Savaresi, "Virtual reference feedback tuning: a direct method for the design of feedback controllers," *Automatica*, vol. 38, pp. 1337-1346, 2002.

[3] Zhongsheng Hou and Shangtai Jin, "Data-Driven Model-Free Adaptive Control for a Class of MIMO Nonlinear Discrete-Time Systems," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 2173-2188, Dec. 2011.

[4] H. Hjalmarsson, S. Gunnarsson, and M. Gevers, "Iterative feedback tuning: Theory and applications," *IEEE Control System Magazine*, vol. 18, no. 4, pp. 26-41, Aug. 1998.

[5] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Englewood Cliffs, NJ: Prentice-Hall, 1999.

[6] K. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 1, pp. 4-27, Jan. 1990.

[7] Hua Deng, Han-Xiong Li, and Yi-Hu Wu, "Feedback-Linearization-Based Neural Adaptive Control for Unknown Nonaffine Nonlinear Discrete-Time Systems," *IEEE Transactions on Neural Networks*, vol. 19, no. 9, Sep. 2008.

[8] D. Mandic and J. Chambers, *Recurrent Neural Networks for Prediction*, New York: Wiley, 2001.

[9] D. P. Mandic, "NNGD algorithm for neural adaptive filters," *Electronics Letters*, vol. 39, no. 6, pp. 845-846, 2000.

[10] Danilo P. Mandic, "A Generalized Normalized Gradient Descent Algorithm," *IEEE Signal Processing Letters*, vol.11, no. 2, pp. 115-118, Feb. 2004.

[11] D. P. Mandic and I. R. Krcmar, "Stability of NNGD algorithm for nonlinear system identification," *Electronics Letters*, vol. 37, no. 3, pp. 200-202, 2001.

[12] Danilo P. Mandic and Jonathon A. Chambers, "A normalised real time recurrent learning algorithm," *Signal Processing*, vol. 80, pp. 1909-1916, 2000.

[13] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," Neural Computations, vol. 1, no. 3, pp. 270-280, 1989.

A Modified Approximate Internal Model-based Neural Control for the Typical Industrial Processes

Jasmin Igic and Milorad Bozic

***Abstract*—A modification of the Approximate Internal Model-based Neural Control (AIMNC), using Multi Layer Neural Network (MLNN) is introduced. A necessary condition that the system provides zero steady-state error in case of the constant reference and constant disturbances is derived. In the considered control strategy only one neural network (NN), which is the neural model of the plant, should be trained off-line. An inverse neural controller can be directly obtained from the neural model without need for a further training. Simulations demonstrate performance improvement of the modified AIMNC strategy. An extension of the modified AIMNC controller for the typical industrial processes is proposed.**

***Index Terms*—Industrial processes, Neural networks, Nonlinear internal model control, Zero steady-state error.**

Original Research Paper
DOI: 10.7251/ELS1418046I

I. Introduction

We may imagine an Internal Model Control (IMC) architecture as a system with two degrees of freedom consisting of plant model dynamics and some approximation of the inverse dynamics of the plant model. In this respect the process as well as its approximate inverse dynamics must be stable. If the dynamics of the process is not stable, then the objective of the system design using an IMC approach is, primarily, stabilization of the process. Such a stabilized system is then viewed as an equivalent process that must satisfy forward-mentioned characteristics in terms of stability.

In addition to stability not less important aspect is the system accuracy. An architecture of the IMC has many positive characteristics in terms of stability, robustness and accuracy in a steady-state with respect to one degree of freedom approaches developed in the area of control engineering [1],[2].

Manuscript received 13 May 2014. Received in revised form 7 June 2014. Accepted for publication 11 June 2014.

J. Igic is with the mtel a.d. Banja Luka, Kralja Petra I Karadjordjevica 93, 78.000 Banja Luka, Bosnia and Herzegovina (e-mail: jasmin.igic@mtel.ba).

M. Bozic is with the University Banja Luka, Faculty of Electrical Engeneering, Banja Luka, Bosnia and Herzegovina. (e-mail: mbozic@etfbl.net).

The IMC design algorithms based on linear models have attracted much attention of control theorists, as well as of practitioners, especially within industrial applications in 1980s. Almost in parallel with such developments many successful trials were conducted using the IMC architecture combined with some kind of adaptation mechanism applied to control slowly varying processes [3],[4].

Further more, as the MLNN have been proved as universal approximators they were used in IMC architectures to control nonlinear processes [2],[5]-[10]. In that sense an Approximate Internal Model-based Neural Control (AIMNC) and its modification were proposed for unknown nonlinear discrete time processes [11]-[13]. High performance of the mentioned algorithms have been demonstrated at applications to some benchmark examples of industrial processes. However, we have noticed unacceptable behavior in applications of such designs to control slow industrial processes. The typical industrial processes show some accumulation effects plus dead time, i.e. they usually have a very slow dynamics.

The well-known examples of such processes are tanks, where the liquid level is controlled by using the difference between the input and output flow rates to form a manipulated variable [14], so in this paper we consider a double tank system as a plant.

To overcome the above mentioned problem in applications of the modified AIMNC algorithm, in this paper we present an extension of the modified AIMNC controller design for slow nonlinear discrete time processes. This extension is based on a heuristic consideration of the needed control changes which provide requested tracking performance of the control system. However, simulation results confirm the validity of the proposed control design in cases of slow nonlinear discrete time processes.

The rest of the paper is organized as follows. In the Section II, a plant modeling, structure and the modified AIMNC controller design procedure are given. The Section III gives an extension of the modified AIMNC algorithm applied to double tank system together with simulation results demonstrating performance of the proposed approach in controlling slow processes. In the Section IV we give conclusions of the work.

II. THE AIMNC CONTROLLER DESIGN

A. The plant modeling

A general input–output representation for an *n*-dimensional unknown nonlinear discrete time system with relative degree *d* is as follows [15]

$$y(k+d) = f[w_k, u(k)], \tag{1}$$

where a vector w_k is composed of current $y(k)$ and past values of the output $y(k-1), i=1,\ldots,n-1$, and past values of the input $u(k-1), i=1,\ldots,n-1$, at the system, and nonlinear mapping is defined by $f: R^n \times R^n \to R$ with $f \in C^\infty$.

Modeling is the first step in designing a controller for unknown nonlinear system. Different types of neural networks have been considered for modeling and control of nonlinear dynamical systems. In this paper, a MLNN has been used for modeling of nonlinear discrete time dynamical systems due to its general approximation abilities. If there is an appropriate number of the neurons in the hidden layers and adequately determined free parameters, the MLNN can approximate arbitrary continuous nonlinear function on a compact subset C^∞ of $R^n \times R^n$ to the desired accuracy [16].

A Neural Network *Nonlinear Autoregressive Moving Average* (NARMA) model, i.e. NN NARMA model is defined as follows [11],[12]

$$y(k+d) = N[w_k, u(k)] + \xi_k, \tag{2}$$

where $N[\bullet]$ is a neural model of the nonlinear dynamical system and the weight vector of the NN is omitted for simplicity, ξ_k is an output model error.

Taking into account the disturbances acting on the plant, (2) can be written as

$$y(k+d) = N[w_k, u(k)] + v_k, \tag{3}$$

where v_k represents the effect of uncertainties (model error ξ_k and disturbances).

B. An approximation of the NN model

The NN controller in the IMC control structure represemts an inversion of the NN model of the plant [7]. Hence, it is necessary to find the inversion of nonlinear mapping, represented by the NN, that models the nonlinear plant [2], [3]. Thus [11]-[13], an approximate model for the system (3) using Taylor series expansion of $N[w_k, u(k)]$ with respect to $u(k)$ around $u(k-1)$ is as

$$\begin{aligned} y(k+d) &= N[w_k, u(k)] + v_k = \\ &= N[w_k, u(k-1)] + N_1[w_k, u(k-1)]\Delta u(k) + R_k + v_k \end{aligned}, \tag{4}$$

where $N_1[w_k, u(k-1)] = \left(\partial N[w_k, u(k-1)]\right)/\left(\partial u(k)\right)$, $\Delta u(k) = u(k) - u(k-1)$, and remainder R_k is given by

$$R_k = N_2[w_k, \zeta_k](\Delta u(k))^2 / 2, \tag{5}$$

where $N_2[w_k, \zeta_k] = \left(\partial^2 N[w_k, \zeta_k)]\right)/\left(\partial u^2(k)\right)$ with ζ_k as a point between $u(k)$ and $u(k-1)$.

Based on the assumptions made in [11], after neglecting the reminder R_k and the uncertainty v_k in (4), the NN approximate model output $\hat{y}_m(k+d)$ is derived as follows

$$\hat{y}_m(k+d) = N[w_k, u(k-1)] + N_1[w_k, u(k-1)]\Delta u(k). \tag{6}$$

Since in (6), a control increment $\Delta u(k)$ appears linearly in the output $\hat{y}_m(k+d)$ of the NN approximate model, thus the design of the inverse NN controller is facilitated. To obtain the NN approximate model (6) and NN controller, the calculation of $N_1[w_k, u(k-1)]$ is needed. As was shown in [12], when using the NN with two hidden layers and neurons with hyperbolic activation functions within it, both the NN model and a calculation of $N_1[w_k, u(k-1)]$ can be obtained by one neural network, whose parameters are trained off-line.

C. The controller design

From (6) the control increment $\Delta u(k)$ is as follows

$$\Delta u(k) = (\hat{y}_m(k+d) - N[w_k, u(k-1)]) / N_1[w_k, u(k-1)]. \tag{7}$$

The control increment in (6) can be divided into two parts [12], as shown below

$$\Delta u(k) = \Delta u_n(k) + \Delta u_c(k), \tag{8}$$

where $\Delta u_n(k)$ refers to the nominal control increment and $\Delta u_c(k)$ is used to compensate the model error and disturbances. Using (8), (6) becomes

$$\begin{aligned} \hat{y}_m(k+d) &= N[w_k, u(k-1)] \\ &+ N_1[w_k, u(k-1)](\Delta u_n(k) + \Delta u_c(k)). \end{aligned} \tag{9}$$

When the model is exact and there is no disturbances, i.e. in the nominal case, the NN approximate model is given as follows

$$\hat{y}_m(k+d) = N[w_k, u(k-1)] + N_1[w_k, u(k-1)]\Delta u_n(k). \tag{10}$$

If the NN approximate model given by (10) has a stable inverse, then the nominal control increment $\Delta u_n(k)$ can be obtained directly by

$$\begin{aligned} \Delta u_n(k) &= \\ &= (r(k+d) - N[w_k, u(k-1)]) / N_1[w_k, u(k-1)] \end{aligned}, \tag{11}$$

where $r(k+d)$ is the reference signal at the time instant $(k+d)$. The nominal control law is determined as

$$u(k) = u(k-1) + \Delta u_n(k). \tag{12}$$

If the NN approximate model given by (10) has an unstable inversion, it is necessary to modify (11) by introducing a parameter α, as proposed in [12], according to

$$\begin{aligned}&\Delta u_n(k) = \\ &= \alpha(r(k+d) - N[w_k, u(k-1)]) / N_1[w_k, u(k-1)],\end{aligned} \tag{13}$$

where $0 < \alpha \le 1$. An introduction of the parameter α ensures that the control law given by (12) is bounded. The parameter α can be 1 in the case where the NN approximate model has the stable inverse. Conditions for the stability of the nominal NN controller given by (12) are detailed in [12].

In the presence of the plant model error and disturbances, substituting (8) in to (4), gives

$$\begin{aligned}&y(k+d) = N[w_k, u(k-1)] + N_1[w_k, u(k-1)]\Delta u_n(k) \\ &+ N_1[w_k, u(k-1)]\Delta u_c(k) + R_k + v_k.\end{aligned} \tag{14}$$

Define the NN approximate model error $\varepsilon(k)$ as

$$\varepsilon(k) = y(k) - \hat{y}_m(k). \tag{15}$$

The control increment $\Delta u_c(k)$ for compensation of the model error and disturbances is as follows [12]

$$\Delta u_c(k) = -\varepsilon(k) / N_1[w_k, u(k-1)]. \tag{16}$$

Based on (8), (13) and (16) the control law is expressed as

$$\begin{aligned}&u(k) = u(k-1) + \alpha(r(k+d) \\ &- N[w_k, u(k-1)]) / N_1[w_k, u(k-1)] - \varepsilon(k) / N_1[w_k, u(k-1)].\end{aligned} \tag{17}$$

The control law (17) consists of the nominal NN controller and uncertainties compensation. The analysis of robustness and stability of the control law (17) is given in [12].

The conceptual structure of the modified AIMNC with control law given by (17) and the NN approximate model given by (10) is shown in Fig. 1. With $S(z^{-1})$ is labeled a set point filter, with $F(z^{-1})$ robustness filter, where z^{-1} is backward shift operator [17]. The role of the blocks marked with "Scale" in Fig. 1. will be explained in the Section III.

D. The modified AIMNC controller design

A positive feature of the IMC control structure is that zero steady-state error in the system can be achieved if we ensure that the steady-state gain of the controller is the inverse value of the steady-state gain of the model [7]. On other hand, it has been shown in [18], that the controller design in the AIMNC structure can be highly facilitated when the reference signal and disturbances have constant values.

Here, we repeat conditions under which it is possible to achieve the satisfactory accuracy in the steady-state with AIMNC structure shown on the Fig. 1. in presence of the constant reference signal and constant disturbances [18]. If the system achieves the zero steady-state error then $y(k) = r(k)$ and $\Delta u(k) = u(k) - u(k-1) = 0$, then using (15) and (17), one has

$$\begin{aligned}&\Delta u(k) = \alpha(r(k+d) - N[w_k, u(k-1)]) / N_1[w_k, u(k-1)] \\ &- \varepsilon(k) / N_1[w_k, u(k-1)] = 0,\end{aligned}$$

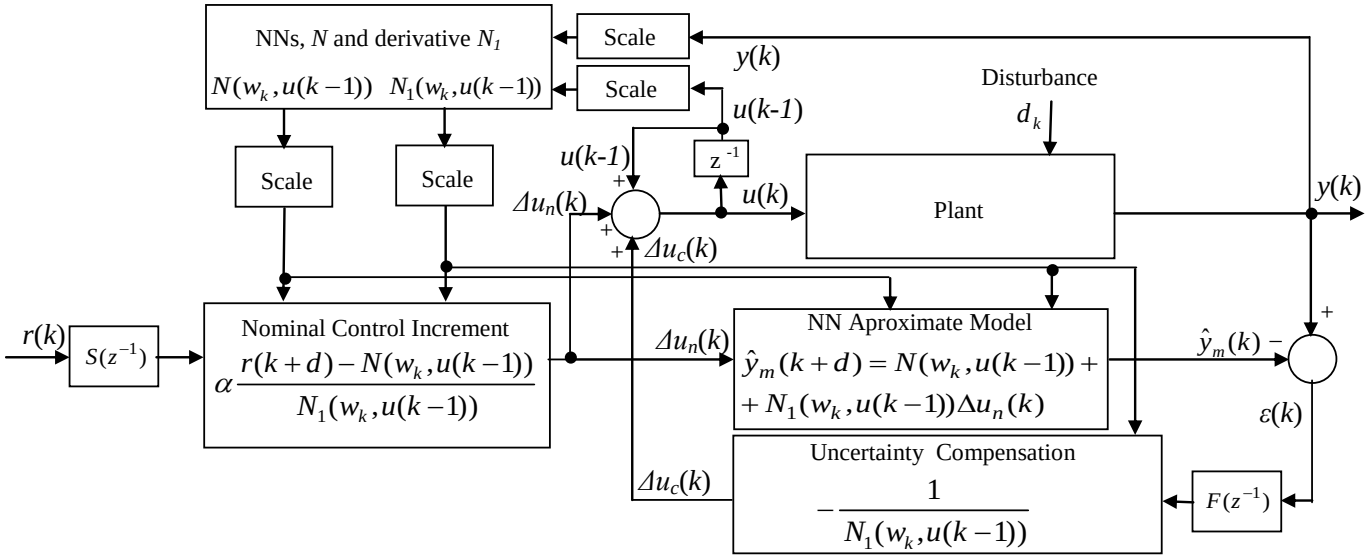


Fig. 1. The conceptual structure of the modified AIMNC

or

$$\alpha r(k+d) - \alpha N[w_k, u(k-1)] - \varepsilon(k) = 0,$$

and after substituting (15) in last equation it becomes

$$\alpha r(k+d) - \alpha N[w_k, u(k-1)] - y(k) + \hat{y}_m(k) = 0. \quad (18)$$

In the steady-state we have $\hat{y}_m(k) = \hat{y}_m(k+1) = ... = \hat{y}_m(k+d)$. Substituting (10) and (13) in to (18), we obtain

$$\alpha r(k+d) - \alpha N[w_k, u(k-1)] - y(k) + N[w_k, u(k-1)] + \alpha(r(k+d) - N[w_k, u(k-1)]) = 0,$$

or

$$2\alpha(r(k+d) - N[w_k, u(k-1)]) = y(k) - N[w_k, u(k-1)],$$

and consequently

$$\alpha = \frac{1}{2} \cdot \frac{y(k) - N[w_k, u(k-1)]}{r(k+d) - N[w_k, u(k-1)]}. \quad (19)$$

Also, when reference values are constant $r(k) = r(k+1) = ... = r(k+d)$ the system will have the zero steady-state error if $y(k) = r(k) = r(k+1) = ... = r(k+d)$. It follows from (19) that $\alpha = 0.5$ is the necessary condition that the system in Fig. 1. attains the zero steady-state error in the case of the constant reference signal and constant disturbances.

The comparison of the AIMNC strategy with respect to performance of the fixed and adaptive IMC has been presented in [18]. Through a sequence of simulations we have demonstrated performance of the AIMNC strategy and verified that the system with the parameter value of $\alpha = 0.5$ achieves the zero steady-state error in case of the constant reference signal and constant disturbances.

III. The Modified AIMNC for the Double Tank System

Since most industrial processes have slow dynamics we consider the control design procedure for the nonlinear plant in the cases of such kind of process dynamics. Also, in a large number of industrial processes inputs and outputs of the plant can not take negative values. In addition to that, it is necessary to take into account a downside of the possible actuator saturation.

In order to illustrate the validity of the proposed control design in cases of slow nonlinear discrete time processes we have chosen a double tank system as the plant. Its schematic representation is shown in Fig. 2. It is a nonlinear plant with slow dynamics that is often used for a verification of various control strategies: adaptive IMC [4], neuro-adaptive IMC [9], self-tuning regulator [19], Generalized Predictive Control

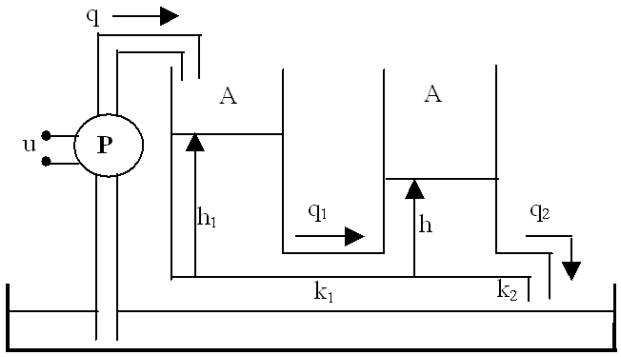


Fig. 2. The schematic representation of the double tank system

(GPC) [20], self-tuning IMC-PID regulator [21], direct control with neural networks [22], swarm adaptive tuning of hybrid PI-NN controller [23], Robust $\mathcal{H}2/\mathcal{H}\infty$/reference model dynamic output-feedback control [24].

The input flow q into the first tank is proportional to the voltage u of the pump P, which is the input to the plant. The fluid flows from the first tank to the second and flow rate q_1 is a function of the difference of liquid level h_1 and h in the first and second tank, respectively. From the second tank the fluid flows freely and a flow q_2 is a function of liquid level h which is the output of the plant. The task of the control system is control of the liquid level $y = h[\mathrm{m}]$ in the second tank by changing the voltage $u[\mathrm{V}]$ at the pump input.

The parameters of the plant, shown in Fig. 2., are: a cross-sectional area of the tank $A = 0.0154[\mathrm{m}^2]$, the discharge coefficient of the first tank k_1 and discharge coefficient of the second tank k_2. Based on the balance between changes in the amount of fluids in the tanks and flows we have a system of nonlinear equations [21],[22]:

$$\frac{dh_1}{dt} = \frac{1}{A}\left(ku - k_1\sqrt{2g(h_1 - y)}\right), \quad (20)$$

$$\frac{dy}{dt} = \frac{1}{A}\left(k_1\sqrt{2g(h_1 - y)} - k_2\sqrt{2gy}\right), \quad (21)$$

where $k = q/u = 1.174 \cdot 10^{-5}[\mathrm{m}^3/\mathrm{Vs}]$ is the coefficient of the pump and $g = 9.81[\mathrm{m/s}^2]$ is the acceleration of gravitational force. Since the dynamics of the pump is negligible compared to the dynamics of the two connected tanks, it is presented here as a static gain. The parameters k_1 and k_2 are determined experimentally in a steady state. Let in the steady state magnitudes h_1^0 and y^0 are achieved at a control signal u^0, then $dh_1/dt = dy/dt = 0$ and we have

$$u^0 = \frac{k_1\sqrt{2g}}{k}\sqrt{\left(h_1^0 - y^0\right)}, \quad (22)$$

$$y^0 = \frac{k_1^2}{k_1^2 + k_2^2} h_1^0 . \quad (23)$$

For the experimentally determined values of the parameters one obtains $k_1 = 2{,}46476 \cdot 10^{-5}[\mathrm{m}^2]$ and $k_2 = 1{,}816245 \cdot 10^{-5}[\mathrm{m}^2]$.

Simple forward-difference approximations of the plant model given by (20) and (21) are

$$h_1(k+1) = h_1(k) + \frac{1}{A}\left(ku(k) - k_1\sqrt{2g(h_1(k) - y(k))}\right), \quad (24)$$

$$y(k+1) = y(k) + \frac{1}{A}\left(k_1\sqrt{2g(h_1(k) - y(k))} - k_2\sqrt{2g \cdot y(k)}\right). \quad (25)$$

By entering the above mentioned parameter values of the plant into (24) and (25) we obtain a nonlinear discrete time model of the plant used for the design of the AIMNC controller:

$$h_1(k+1) = h_1(k) + 0.0007625u(k) - 0.0070893\sqrt{h_1(k) - y(k)}, \quad (26)$$

$$y(k+1) = y(k) + 0.0070893\sqrt{(h_1(k) - y(k))} - 0.0052243\sqrt{y(k)}. \quad (27)$$

In this model it is necessary to include the saturation of the actuator, i.e. the control signal u (the plant input) is in the range of $[0 \;\; 10]$ volts. Also, the output y of the plant can not take negative value. The fluid level in the first tank must not exceed the value $h_1 = 0.65[\mathrm{m}]$. From the (23), $h_1^0 / y^0 = 1.543$, the desired stationary value of the plant output must not exceed the value $y^0 = 0.4213[\mathrm{m}]$.

The step response of the double tank system, represented by (26) and (27), for the pump voltage $u = 3.065[\mathrm{V}]$, takes a stationary value of $y = 0.2[\mathrm{m}]$ at the output, Fig. 3. The nonlinear discrete time plant model of double tank system can be approximated by a linear one with a single pole and gain as is given by

$$\tilde{P}_{DTS}(z^{-1}) = \frac{b_1 z^{-1}}{1 - a_1 z^{-1}} = \frac{0.0001893 z^{-1}}{1 - 0.9971 z^{-1}}, \quad (28)$$

This model has a pole at $a_1 = 0.9971$. It has a very slow dynamics, Fig. 3. The system needed over 2000 sample periods to reach a steady state from a zero initial level, by a constant excitation of the pump, whereby a sampling period for a real system is 1 second.

A. A Neural model of the double tank system

The MLNN structure with two hidden layers has been used for modeling of the nonlinear discrete time dynamical system. As inputs of the NN at time instant *k* were $w_k = [y(k)]$ and $u(k-1)$. The MLNN in the first and second hidden layer has had 10 neurons with hyperbolic tangent activation functions and bias inputs. The linear activation function and bias input has been used for the output neuron.

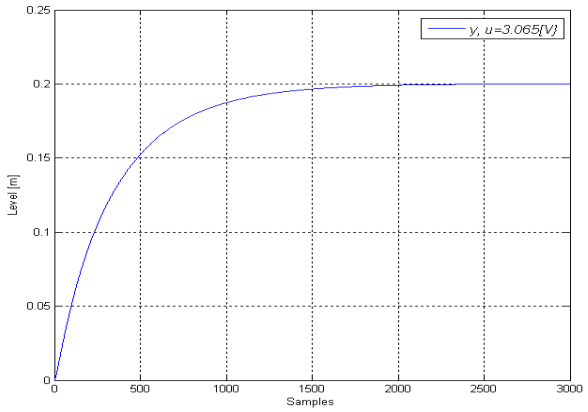


Fig. 3. The step response of the double tank system for pump voltage $u = 3.065[\mathrm{V}]$

Since the double tank system have slow dynamics it is necessary to pay special attention to the generation of adequate training set, as well as the scaling of inputs and outputs of a neural network in the training process, and then in the AIMNC structure.

The different types of training sets have been generated for different control vectors. A particular training set consisted of 12.000 pairs composed of $y(k)$ and $u(k-1)$, where each pair represents an input vector of the neural network in time instant *k*. The desired value $y(k+1)$ has been obtained based on (26) and (27).

A very good neural model for the AIMNC structure shown in Fig. 1. was obtained for the control input $u(k), k = 1,...,12000$ chosen as a random number in the range $[0 \;\; 8]$ with a mean value 4.0077, and the generated 12.001 values of $y(k)$ were in the range $[0 \;\; 0.373]$. Fig. 4. shows the double tank system output $y(k)$ generated by this random control input.

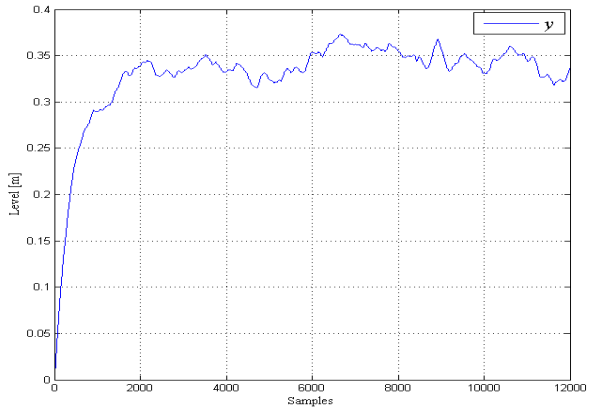


Fig. 4. The double tank system output y for the input vector $u = rand[0 \;\; 8][\mathrm{V}]$

A motivation for this choice of training pairs has been derived from a request to regulate the plant output in the range from 0 to 0.35 meters, and therefore it was chosen to change the control signal from 0 to 8 volts with a mean value near to 4. The steady state $y^0 = 0.35[\text{m}]$ corresponds to the voltage $u^0 = 4.0535[\text{V}]$ at the pump input. The initial value of the plant output has been set to $y(1) = 0$. With the proposed selection of training pairs and the initial conditions an opportunity was afforded to observe a plant dynamics contained in a transition process from the zero initial state to the end of the operating range, and in the vicinity of the steady-state.

For the NN training, a MATLAB function "newnarxsp" [25], has been used. The inputs to the MLNN were scaled to the range $[-1,\ 1]$, training algorithm used *Levenberg-Marquart* method [26]-[28], a number of training epochs was 1000 and the achieved mean square error was 4.45x10^{-5}. As a result of off-line training we obtained the matrix $W_{10} = [W_{11}\ \ W_{12}]$, $W_{10} \in R^{10\times 2}$, $W_{12} \in R^{10\times 1}$, $W_{21} \in R^{10\times 10}$ and $W_{32} \in R^{1\times 10}$ and vectors b_1, b_2 and b_3, with dimensions 10x1, 10x1 and 1x1, respectively.

The blocks marked with a "Scale" in Fig. 1. are introduced for the reason that the off-line training of MLNN performs scaling of inputs and the output to the desired range $[-1\ \ 1]$. So to calculate $N[w_k, u(k)]$ in the AIMNC control structure it is necessary to scale inputs within the range $[-1\ \ 1]$, and the output in the range $[0\ \ 0.373]$.

Fig. 5. shows the step responses of the neural network $N[y(k), u(k-1)]$ and of the double tank system given by (26) and (27) for the step change of the voltage $u = 3.065[\text{V}]$ at the pump input. The inputs in the neural network were scaled to the range $[-1,\ 1]$, and the output to the range $[0\ \ 0.373]$. From Fig. 5. is evident that the resulting neural model, with an order of the model error of $10^{-4}[\text{m}]$, is a good model of nonlinear plant at hand. The model error ε_m is shown on Fig. 5, too.

A derivative of the the neural network $N_1[y(k), u(k-1)]$ is shown in Fig. 6. In this case, the derivative of the NN for a modeling of the plant with slow dynamics has an order of 10^{-3}. It appears in the denominator of (13) and (16) for a calculation of the nominal control increment $\Delta u_n(k)$ and control increment for compensation of the model error and disturbances $\Delta u_c(k)$, respectively. For sake of comparison, Fig. 6b. also shows the derivative of the neural network for the modeling of the nonlinear plant with fast dynamics considered in [18].

B. The modified AIMNC controller for the double tank system

In the case of the nonlinear plant with slow dynamics the derivative of the neural network $N_1[y(k), u(k-1)]$ usually takes very small values. When using the AIMNC control structure this could represent a serious problem. The control increment in the control law given by (17) can take unacceptably high values. On the Fig. 7., the reference signal $r = 0.2[\text{m}]$, system output y, NN approximate model output $\hat{y}_m$ and model error ε are shown and on the Fig. 8. the corresponding control action u and control increment Δu in the case of reference signal $r = 0.2[\text{m}]$. From the Fig. 8. it is seen an unacceptable behavior of the AIMNC control structure. The system is practically stable only due to the saturation of the actuator.

The control increments, due to the very small values of derivative of the neural network, are too high and the voltage at the pump oscillates between the saturation limits 0 and 10.
It is therefore necessary to consider how it is possible to limit the values of the control increments in order to achieve acceptable system behavior.

An adequate control signal to control systems with slow dynamics, that most of the industrial processes posses, has three distinct segments. Typically, the first segment consists of a step change of the control signal, and the second of its exponential decrease or increase depending on the sign of the reference signal step change. In the case of the positive plant gain, if the step change of the reference signal is positive then the control signal in the second segment is decreasing and vice

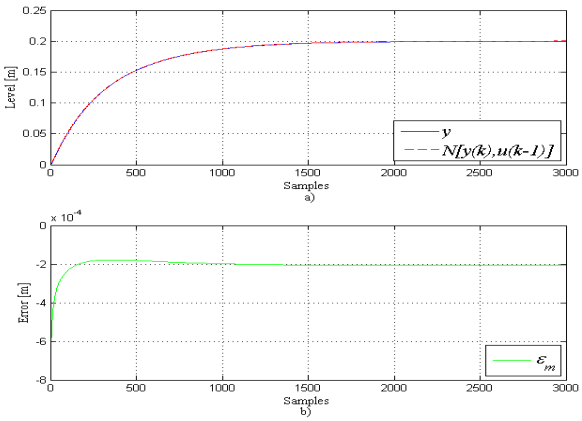


Fig. 5. The neural network $N[\cdot]$ and the double tank system y outputs, and model error ε_m for the step changes of voltage at the pump input

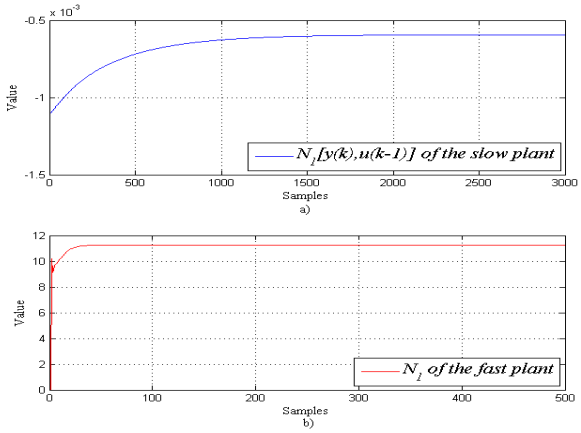


Fig. 6. The derivatives $N_1[\cdot]$ of the NNs of the slow plant and fast plant

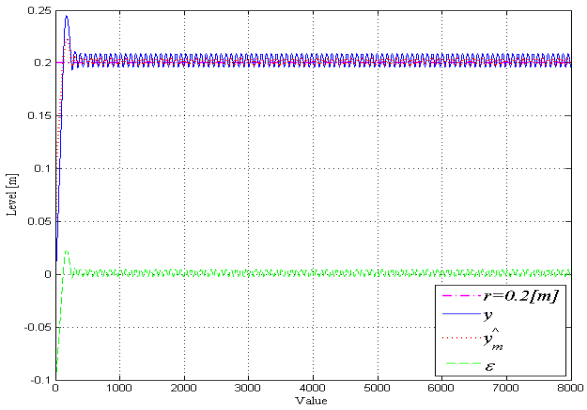


Fig. 7. Reference $r = 0.2[\mathrm{m}]$, system output y, NN approximate model output $\hat{y}_m$ and model error ε

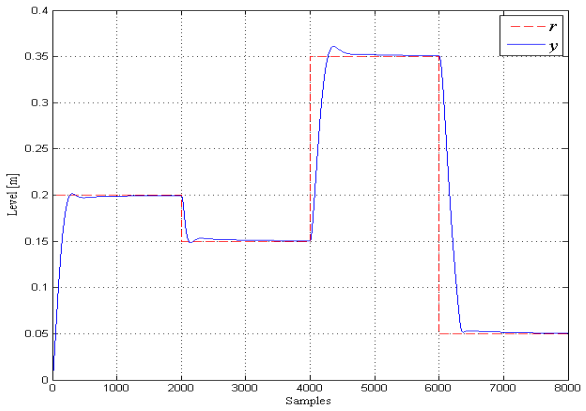


Fig. 9. System output y and reference r

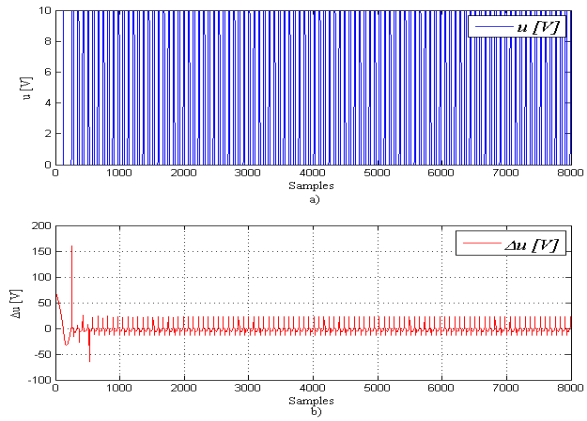


Fig. 8. Control u and control increment Δu for reference $r = 0.2[\mathrm{m}]$

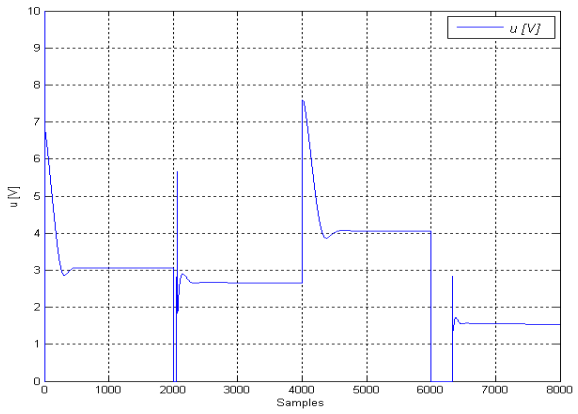


Fig. 10. Control action u

versa. The third segment is with a constant value of the control signal for a fixed reference signal that should provide the zero steady-state error.

It has already been shown that the satisfactory accuracy of the system with the AIMNC controller will be achieved if $\alpha = 0.5$, and we have therefore provided by the proper value of the control signal in the third segment.

A required form of the control signal can be achieved if we ensure that the step change of the reference signal create step change of the neural model output. This requires a different way to scale the control signal that is an input to the neural network, i.e. it is necessary to increase it. Therefore, in the considered case we have scaled the input of the neural model, to the range $[0 \quad 62]$.

We have performed the simulation of the AIMNC strategy shown in Fig. 1. with the obtained MLNN. The set point filter was chosen as $S(z^{-1}) = (1-r_1)/(1-r_1 z^{-1})$ with $r_1 = 0$, i.e. $S(z^{-1}) = 1$, and a robustness filter as $F(z^{-1}) = (1-0.99942)/(1-0.99942 z^{-1})$. Due to the proposed method of the control signal scaling, neural model error became bigger, and we have used the robustness filter $F(z^{-1})$.

The reference $r(k)$ was chosen to take the values of 0.2, 0.15, 0.35, and 0.05, successively for five periods of the 2000 samples.

On the Fig. 9. the response of the system with AIMNC controller for the double tank system with the proposed method of a control signal scaling is shown. The corresponding control signal is shown in Fig. 10.

Figs. 9.-10. depict satisfactory behavior of the proposed modified AIMNC strategy for the typical industrial processes and confirm that the choice of parameter $\alpha = 0.5$ provides the zero steady-state error in cases of the constant reference signals.

IV. Conclusion

In this paper we have presented an improvement of the AIMNC structure. The procedure of designing the MLNN model and controller is shown. The necessary condition for providing an appropriate accuracy of the system at steady state in the case of the constant reference signal and constant disturbances is derived.

Also, we have suggested an approach in which one can ensure satisfactory behavior of the AIMNC law for controling

slow industrial processes and provide the zero steady state error in the cases of constant reference signals and constant disturbances. Simulation results confirm performance improvements obtained by the proposed modification of the AIMNC algorithm.

REFERENCES

[1] T.Samad, and A.Annaswamy (Editors), "The Impact of Control Technology: *Overview, Success Stories, and Research Challenges",* IEEE Control Systems Society, February 2011. Available: http://ieeecss.org/main/IoCT-report.

[2] K.J.Hunt, and D.Sbarbaro, "Neural networks for nonlinear internal model control", *IEE Proceedings-D,* vol. 138, no. 5, pp. 431-438, September, 1991.

[3] J. M. Sanches and J. Rodellar, *Adaptive Predictive Control: From the concepts to plant optimization*, Prentice Hall, 1996.

[4] J.Igic, M.Bozic, P.Maric, and I.Krcmar, "Adaptive Control based on Internal Model", *Proc. XLV ETRAN, Conference* vol. 1, pp. 209-212, Bukovicka Banja, June, 2001.

[5] P.M.Mills, A.Y.Zomaya and M.O.Tade, *Neuro-Adaptive Process Control: A Practical Appoach*, John Willey&Sons, 1996.

[6] C.Kambhampati, R.J.Craddock, M.Tham, and K.Warwick, "Inverse model control using recurrent networks", *Mathematics and Computers in Simulation*, vol. 51, pp. 181-199, 2000.

[7] I.Rivals, and L.Personnaz, "Nonlinear internal model control using neural networks: Application to processes with delay and design issues," *IEEE Transactions Neural Networks,* vol. 11, no. 1, pp. 80–90, January, 2000.

[8] M.Mohammadzaheri, L.Chen, and S.Grainger, "A critical review of the most popular types of neuro control*", Asian Journal of Control,* vol. 14, no. 1, pp. 1-11, January 2012.

[9] J.Igic, M.Bozic, and P.Maric, "Neuro-Adaptive Control based on Internal Model", *Proc. XLVI ETRAN, Conference* vol. 1, pp. 161-164, Banja Vrucica, Teslic, June, 2002.

[10] V.G.Krishnapura, and A.Jutan, "A neural adaptive controller", *Chemical Engineering Science 55*, pp. 3803-3812, 2000.

[11] H.Deng, and H.-X.Li, "A Novel Neural Approximate Inverse Control for Unknown Nonlinear Discrete Dynamical Systems", *IEEE Transactions on Systems, Man, and Cybernetics*—Part B: Cybernetics, vol. 35, no. 1, pp. 553-567, February, 2005.

[12] H.-X.Li, and D.Hua, "An Approximate Internal Model-Based Neural Control for Unknown Nonlinear Discrete Processes," *IEEE Transactions on Neural Networks,* vol. 17, no. 3, pp. 659-670, May, 2006.

[13] Y-N.Wang, and X-F. Yuan, "SVM Approximate-based Internal Model Control Strategy", *Acta Automatica Sinica*, vol. 34, no. 2, pp. 172-179, February, 2008.

[14] A. Visioli, and Q.-C. Zhong, *Control of Integral Processes with Dead Time*. Springer Verlag, London, 2011.

[15] J.B.D. Cabrera, and K.S.Narendra, "Issues in the application of neural networks for tracking based on inverse control*", IEEE Transactions on. Automatic Control*, vol. 44, no. 11, pp. 2007–2027, November 1999.

[16] M.Sanguineti, "Universal Approximation by Ridge Computational Models and Neural Networks: A Survey", The Open Applied Mathematics Journal, pp. 31-58, 2008.

[17] L.Ljung, *System Identification-Theory for the User*, 2nd Edition., A Simon & Schuster Company, N.J., Prentice-Hall PTR, 1999.

[18] J.Igic, and M.Bozic, "An Modified Approximate Internal Model-based Neural Control", *Proc. XIII INFOTEH Conference*, vol.13, pp. 1026–1031, Jahorina, March, 2014.

[19] J.M.Fernandes, C.E.De Souza, and G.C.Goodwin, "Adaptive Control of Coupled Tank Apparatus", *International Journal of Adaptive Control and Signal Processing*, vol.3, pp.319-331, 1989.

[20] H.Demircioğlu, and E.Karasu, "Generalized Predictive Control", *IEEE Control Systems Magazine*, vol.20, no.5, pp.36-47, October 2000.

[21] W.K.Ho, T.H.Lee, H.P.Han, and Y.Hang, "Self-Tuning IMC-PID Control with Interval Gain and Phase Margins Assignment" *IEEE Transactions on Control Systems Technology*, vol.9, no.3, pp.535-541, May 2001.

[22] Y.Li, and A.Häußler, "Artificial Evolution of Neural Networks and its Application to Feedback Control", Centre for Systems and Control, University of Glasgow, Tech. Report: CSC-95011, Glasgow, UK, 1995

[23] M. S. Ramli, R. M. T. R. Ismail, M. A. Ahmad, S. M. Nawi, and M. V. Mat, "Improved Coupled Tank Liquid Levels System Based on Swarm Adaptive Tuning of Hybrid Proportional-Integral Neural Network Controller," *Am. J. Eng. Appl. Sci.*, vol. 2, no. 4, pp. 669–675, 2009.

[24] E. N. Gonçalves, W. E. G. Bachur, R. M. Palhares, and R. H. C. Takahashi, "Robust $\mathcal{H}2/\mathcal{H}\infty$/reference model dynamic output-feedback control synthesis," *Int. J. Control*, vol. 84, no. 12, pp. 2067–2080, November, 2011.

[25] H. Demuth, M. Beale, and M. Hagan, *Neural Network Toolbox 6, User's Guide*. The MathWorks, Inc., 2008.

[26] K. Levenberg, "A method for the solution of certain problems in least squares," *Quart Appl. Math*, vol. 2, pp. 164–168, 1944.

[27] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM J. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[28] J. Nocedal, and S. Wright, *Numerical Optimization*, 2nd Edition., Springer, 2006.

Integrated Cost-Benefit Assessment of Customer-Driven Distributed Generation

Čedomir Zeljković and Nikola Rajaković

***Abstract*—Distributed generation (DG) has the potential to bring respectable benefits to electricity customers, distribution utilities and community in general. Among the customer benefits, the most important are the electricity bill reduction, reliability improvement, use of recovered heat, and qualifying for financial incentives. In this paper, an integrated cost-benefit methodology for assessment of customer-driven DG is presented. Target customers are the industrial and commercial end-users that are critically dependent on electricity supply, due to high consumption, high power peak demand or high electricity supply reliability requirements. Stochastic inputs are represented by the appropriate probability models and then the Monte Carlo simulation is employed for each investment alternative. The obtained probability distributions for the prospective profit are used to assess the risk, compare the alternatives and make decisions.**

***Index Terms*—Distributed generation (DG), customer-perspective approach, integrated cost-benefit assessment, Monte Carlo simulation, long-term planning, uncertainty analysis.**

Original Research Paper
DOI: 10.7251/ELS1418054Z

I. Introduction

In a few recent decades, distributed generation (DG) draws an increasing attention due to its numerous techno-economic and environmental potentials [1]. Primarily, the community may be interested in introducing such a small-scale generation connected close to end-users in order to improve the overall efficiency and security of the nation's energy supply. Secondly, distributed generation may bring significant benefits to distribution power utilities through upgrade investment deferral, avoided electricity purchase, and loss reduction. Finally, the single electricity customers may also achieve remarkable benefits from the use of distribution generation especially through electricity bill reduction, improvement of supply reliability and combined production of heat and power (CHP).

Manuscript received 31 May 2014. Accepted for publication 10 June 2014. This article partially presents results from the PhD thesis "Investment and Operation Aspects of Customer-Driven Distribution Generation under Uncertainty", defended at the Faculty of Electrical Engineering, University of Belgrade.

Čedomir Zeljković is with the Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, Bosnia and Herzegovina (phone: +387-51-221-878; e-mail: cedomir@etfbl.net).

Nikola Rajaković is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: rajakovic@etf.rs).

By the most authors, the concept of optimal investment in distributed generation is considered from the perspective of distribution power utility. This problem principally comes down to determination of rated powers and installation locations for prospective DG units and is usually referred to as optimal DG sizing and sitting. The optimality criteria are mainly loss minimization, improvement of reliability and power quality, network upgrade deferral or voltage profile improvement [2]-[4]. There are also multi-objective approaches with a combination of two or more optimization goals [5]-[7].

Several papers emphasize "a customer approach", although the term customers does not refer to electricity end-users, but to the independent DG developers which consider the investments in DG units that would be connected at the appropriate locations in the distribution network being optimal for investor's profit maximization. For example, Ref. [8] has presented a methodology based on the Monte Carlo simulation where DG units of random size were placed on random locations in the distribution network searching for the solution which would offer the maximum overall network benefits. In the paper is then highlighted that the total achieved benefits must be accordingly translated to the deserving parties, among them the DG owners.

On the contrary to the network-perspective approaches, this paper focuses on the significantly less studied problem of the benefits achievable by the single electricity customers. In particular, we deal with the industrial and commercial customers that are billed by the world's most common time-of-use (TOU) electricity tariff, consisted of both volumetric and demand cost elements. Our target customers observe the problem exclusively from their point of view and are not in possession of any data about the structure and operation of the distribution network they are connected to.

The problem of customer adoption of distributed generation and optimal operation of installed DG units has been illustrated in [9]. The optimization task, written in a form of mixed-integer linear program, has been solved by using commercial software and, in its original form, has taken into account only the benefits of electricity bill reduction. This approach is a pure customer-perspective, though through using the more general concept of a microgrid, where a group of related customers are observed together as a low voltage network. The original model has been extended in [10], where combined generation of heat as well as the use of local energy

storage devices have been considered, but still without taking into account the DG reliability benefits. It is not the only approach where the potentials of CHP systems to provide savings to the customers have been analyzed [11][12]. Finally, in this short literature review, the study of very important issue of stochastic change of fuel and electricity prices must necessarily be mentioned [13].

In our previous efforts we have also discussed about the customer electricity bill reduction by means of the optimal scheduling of distributed generation. We have introduced a method based on a search procedure capable of dealing with non-linear, non-convex and discontinuous DG cost functions [14]. We have also analyzed the impact of financial incentives on customer DG investments [15], as well as the potentials of DG to provide the reliability improvements [16].

In this paper, the results of recent studies have been utilized in order to establish a new integral DG installation assessment methodology that would include a majority of influential factors. The main goal of the methodology is to provide an answer to the customer is it profitable or not to invest in distributed generation and which is the best possible solution from the investment options available on the market. The approach considers the overall life time of the DG investment and, due to uncertainties in input parameters, involves not only deterministic but also stochastic elements. Since the life times of DG technologies are typically a couple of decades long, the change in worth of money over time must also be respected and thus all money amounts must be converted into their present values. Like in the other long-term planning methodologies, the employed models and procedures are kept as simple as possible, but tending to involve all the essential principles. The main benefits taken into account are the customer's bill reduction, reliability improvement, combined heat and power savings as well as financial incentives. On the opposite side of the scales lie the DG investment costs. The probability distributions of net profit calculated from the benefits and costs are used for declaring the best solution. They may also be used for the assessment of customer investment risk. It should be noted that the proposed methodology is modular and, in case of presence of additional facts not covered in this study, the new costs/benefits can be easily attached to the basic framework.

II. Problem Formulation and Solution Methodology

A. Overview of the System under Consideration

The target customer is a medium or large industrial or commercial company whose business is significantly dependent on electricity supply, due to high electricity consumption, high power peak demand or high electricity supply reliability requirements. The traditional way for customer energy supply is purchasing electricity from the distribution utility company. The alternative is buying fuel from the market and generating electricity locally, by using own generating units. The customer considers an investment in on-site dispatchable generating units such as reciprocating engines, gas turbines, microturbines, and fuel cells. The dispatchable units may also be combined with some renewable energy technologies like solar and wind power devices. Depending on their rated powers, DG units may partially or completely satisfy the customer load. A potential deficit can be compensated by purchasing electricity from the utility. The purchasing may also be preferred in the case of electricity prices being more attractive then fuel prices. On the other hand, it is assumed that the surplus electricity is not sold back to the grid. The opposite class of problems, assuming the two-way flow of energy and net-metering, deserves its own approach. Finally, since our research is focused on the application of distributed generation, the other customer cost-saving techniques like demand-side management or local energy storage are not considered in the paper. The concept of the system is illustrated in Fig. 1.

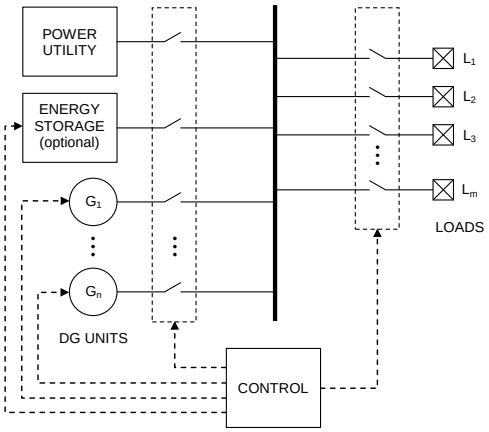


Fig. 1. The concept of the system under consideration

B. Stochastic vs. Deterministic Methodology

In case when all inputs are known with complete certainty, the economic cost-benefit assessment may be performed using the traditional discounted cash flow (DCF) method [17]. For each considered investment alternative, the net present value (NPV) is calculated using the following manner

$$NPV_j = \sum_{k \in L} \frac{B_{j,k} - C_{j,k}}{(1+d)^k} \quad (1)$$

where $B_{j,k}$ is the benefit obtained by the use of DG investment alternative j during the year k, $C_{j,k}$ is the corresponding annualized cost, d is the discount rate and L is the life time of the investment. The usual decision-making approach is choosing the alternative having the highest positive NPV.

On the contrary, in cases when several input variables exhibit stochastic changes, the conventional approach should be extended. One of the most comprehensive methods for analyzing problems that involve uncertainty is Monte Carlo simulation (MCS) [18]. In the MCS method, the main stochastic variables are assigned probabilistic models, a sufficient number of random scenarios is generated and, for each scenario, the behavior of the system is simulated. The output of MCS procedure is a probability distribution of NPV, which represents a full spectrum of possible outcomes of the

investment. By analyzing the obtained distributions for considered investment alternatives, risk is assessed and the best solution is selected.

The example of decision making is shown in Fig. 2. Scenario 1 is the situation where cumulative density functions (CDF) of the profit do not intersect. For the same probability, the profit of project B is always higher than the profit of project A. Or alternatively, given one particular profit, the probability that it will be achieved or exceeded is always higher by project B than it is by project A. Hence, the project B should be unambiguously selected for realization. In scenario 2 the decision may go into two opposite directions. Risk-loving investors will be attracted by the possibility of higher profit and therefore will choose project A. Risk-averse investors will be attracted by the possibility of low loss and will therefore choose project B.

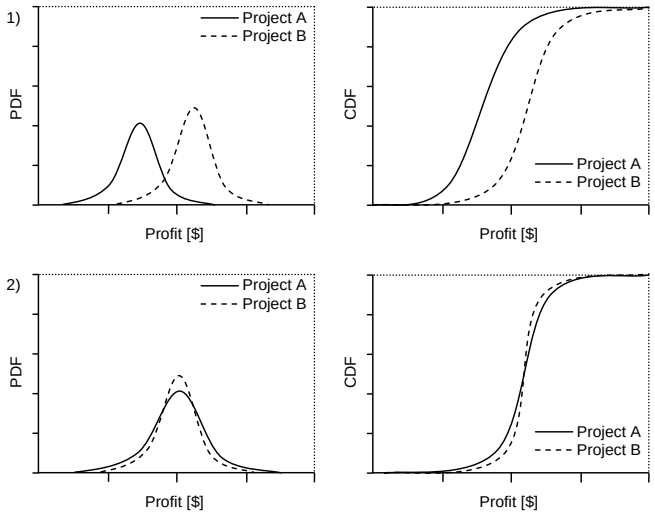


Fig. 2. Using the probability density functions (PDF) and cumulative density functions (CDF) for case comparisons: 1) unambiguous advantage of case B; 2) indeterminate situation - decision between low loss and high profit.

C. The Principle of Existence of Optimum Rated Power

The existence of the optimal investment solution may be demonstrated on a single DG unit example (Fig. 3). The investment costs grow linearly with the increase in rated power of the unit or, due to the economy of scale, their slope may have a slight decrease. The benefits are also getting higher when greater units are chosen, but the finite energy requirements of the customer are the reason for the growth saturation. Therefore, there exists the optimal rated power which guarantees the maximum difference between the benefits and costs.

III. Models and Methods

A. Modeling of Inputs

In procedures for quantification of achievable benefits, the following inputs are considered as crucial:

- Customer load diagram;
- Long term retail prices of electricity;

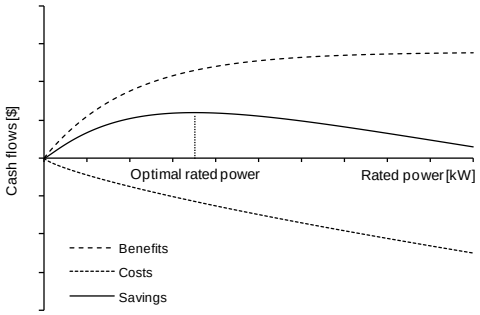


Fig. 3. The principle of finding the optimal rated power

- Long term retail prices of fuel for distributed generating units (such as natural gas, diesel, gasoline or hydrogen);
- Meteorological inputs such as wind speed and insolation (for renewable energy sources, if existent);
- Reliability of distributed generation;
- Reliability of the grid at the customer connection point;
- Start-up costs of distributed generation;
- Customer damage costs incurred by interruptions in power supply.

More details about the stochastic modeling of listed inputs can be found, for instance, in [16],[19]-[26],[28],[29].

B. Quantification of Benefits and Costs

1) Reduction in Customer Electricity Bill

The cost savings achievable by the customer are highly influenced by the way how the generating units are scheduled to run. The most simple but the less economical dispatch procedure is a continuous operation of DG. A quite better solution is dispatching the units at certain periods of day or year, during the hours of peak demand or high price of electricity [12]. A further improvement represents the threshold control, where DG is run whenever the customer load is greater than the predefined threshold value [12][30]. The best results are achievable by employing the heuristic dispatch methods which take into account the probabilistic nature of the input variables [19][31].

In this study, we employ our dispatch strategy which is presented and discussed in [19]. The so called *Enhanced threshold control algorithm* (ETC) is a quality solution which is capable of dispatching an arbitrary number of DG units. The algorithm is designed to successfully cope with realistic issues such as non-linearity in DG efficiency curves, stochastic nature of input variables and existence of peak demand charges in the customer electricity bill.

2) Reliability Benefits

Many industrial and commercial customers, that are sensitive on power interruptions, may also significantly benefit from DG in the role of a backup source. Evaluation of benefits of improved reliability are based on Monte Carlo simulation methodology that we described in [16]. Basically, we determine the customer damage costs incurred due to interruptions in power supply and comparing the cases without and with distributed generation. Avoided interruption costs are equivalent to achieved benefit of improved reliability (B_{REL}).

Since the crucial inputs such as number of failures per year or times of occurrences of particular failures are not deterministic but stochastic variables, the amount of reliability benefits is also obtained in a form of a probability distribution rather than as single value.

3) Benefits of Combined Generation of Heat and Power

For customers who have a substantial need for thermal energy, it could be thought about the investment in DG units that have the option of waste heat recovery (CHP). It is then necessary to assess whether the benefits of recovered heat from the CHP unit is greater than the cost of investment in such an upgrade.

The effectiveness of CHP system is primarily determined by a heat-to-power ratio of the customer [11]. The customers with high heat-to-power ratio, i.e. with heat load dominant over electrical load, utilize the CHP systems in heat-tracking (HT) mode of operation. HT means that the heat output of the CHP system follows the heat demand of the customer. Electricity generation has the second level of priority and the possible deficit of electrical energy is compensated by purchasing from the utility. Most of the savings is achieved on the basis of production of heat energy, while production of electricity, the peak shaving and reliability improvement are not of noticeable importance. Therefore, such kind of customers is not of interest for the methodology proposed in this paper.

Since the target for this study are the customers considerably dependent on supply of electric energy, the remainder of the analysis concentrates on the customers with low heat-to-power ratio. Considering this limitation, it only makes sense to analyze the electricity tracking mode for the operation of the CHP system. In electricity-tracking (ET) mode the primary goal is electricity production. The power output of the CHP system is adjusted to follow the electrical demand of the customer. Deficit of heat energy is compensated by the customer's auxiliary boiler or provided by an external heat supplier, while possible surplus of heat is dissipated into the atmosphere. By working in this regime, it is possible to operate DG units strictly according to previously calculated optimal schedules and thus to achieve maximum self-generation and peak-shaving benefits. The worth of CHP benefit is equivalent to the worth of heat energy which is avoided to be purchased or locally generated by customer boilers.

4) Incentives and Grants

Financial incentives are another important factor which can significantly improve the quality of the considered project and stimulate the customer to invest in distributed generation. DG technologies that are favored and predominantly financially supported by governments or other entities are renewable energy sources, environmentally improved fossil-fuel generation, as well as modern high-efficiency and cogeneration facilities. The diversity of possible types of incentives is practically unlimited and therefore it is hard to derive a universal approach. Instead, every available incentive program deserves its own separate analysis. However, despite of type diversity, every analysis should result in the stream of the present values of expected annual benefits which would the customer achieve from the incentive program.

5) Investment Costs

Analogously to the treatment of benefits, for the use in discounted cash flow, the investment costs are to be displayed in a form of stream of annual cash outflows. Due to a wide variety of options for financing the DG projects, the outflow stream formation procedure is not unambiguously defined. For example, for upfront paid investment, the cash flow comes down just to one outflow - the whole investment cost paid in the year zero. On the other hand, if the investment is loan based, the outflow stream contain values for the overall loan repayment period.

6) Other Benefits and Costs

Depending on regional, economical and regulatory occasions as well as the nature of the customer, it is possible to encounter other specific costs or earn some additional benefits related to the usage of distributed generation. New costs and benefits have also to be properly quantified in a form of annual cash flow streams and included in the evaluation model.

IV. Illustrative Example

A. Input Parameters

The algorithm is tested on a hypothetical customer which weekly diagram of expected load (L_m) and standard deviation (σ) are given in Fig. 4. It is assumed that precise load history is not available so as the random load time series is generated by using the following expression

$$L(h) = L_m(h) + \sigma(h) \cdot N_h(0,1) \tag{2}$$

where $N_h(0,1)$ are independent values drawn from the normal distribution.

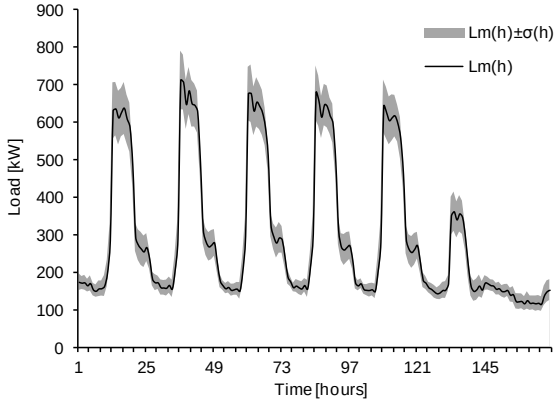


Fig. 4. Stochastic model of the customer load

The customer considers an investment in a natural gas fueled microturbine facility. The equipment manufacturer offers six alternatives (G1 to G6), with power outputs ranging from 65 to 1000 kW. The main characteristics are shown in Table I. The facilities G1 and G2 are made of a single turbine, while the others represent multiturbine configurations built by using 200kW units. The additional important parameters are listed in Table II. The example of extremely non-linear efficiency curve of a microturbine facility (alternative G4) is

shown in Fig. 5.

The life time is 15 years, for all six investment options. The investments costs are covered by a subsidized loan with 5,56% interest rate. The discount rate is set at 7%.

TABLE I
BASIC PARAMETERS FOR CONSIDERED INVESTMENT OPTIONS

Parameter	G1	G2	G3	G4	G5	G6
Technology			Microturbine			
Type	C65	C200	C400	C600	C800	C1000
Configuration	1×65	1×200	2×200	3×200	4×200	5×200
Max. output (kW)	65	200	400	600	800	1000
Min. output (kW)	20	50	50	50	50	50
Invest. cost (M$)	0,15	0,40	0,75	1	1,2	1,4
CHP module (M$)	0,026	0,06	0,12	0,18	0,24	0,3

TABLE II
ADDITIONAL PARAMETERS FOR CONSIDERED INVESTMENT OPTIONS

Parameter	G1-G6
Mean time to failure - MTTF (h)	14000
Mean time to repair - MTTR (h)	3,1
Start-up time (min)	3
Start-up cost ($/start-up)	Start-up time×Full load cost
Non-fuel variable cost ($/kWh)	0,005

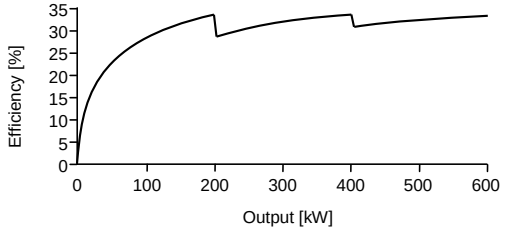


Fig. 5. Efficiency curve of a 3×200kW microturbine facility (alternative G4).

Wholesale prices of natural gas and electricity are simulated by using the Ornstein-Uhlenbeck (OU) correlated stochastic process with mean reverting drift, in accordance with the following equations:

$$dx = \kappa_x(\bar{x} - x)dt + \sigma_x dW_x \tag{2}$$

$$dy = \kappa_y(\bar{y} - y)dt + \rho\sigma_y dW_x + \sqrt{1-\rho^2}\,\sigma_y dW_y \tag{3}$$

where x and y are the logarithms of the electricity and natural gas prices, κ_x and κ_y are the corresponding mean-reversion rates, $\bar{x}$ and $\bar{y}$ are the mean reversion levels, σ_x and σ_y are the corresponding price volatility rates, W_x and W_y and are the standard Wiener processes.

The wholesale prices are transferred to the retail level by averaging on a monthly horizon and by adding the costs of distribution. The retail price of natural gas is assumed to be flat all the month long. The electricity prices are assumed different for both on-peak and off-peak volumetric consumption as well as the monthly peak demand, which is adjusted by appropriate factors. The most important parameters are summarized in Table III, while more detailed explanations can be found in [29].

The average number of grid outages is 1 failure/year, while the expected repair time is 2,3 hours. The outages are simulated using the failure pattern recorded at the customer connection point - the distributions of the moment of occurrence and repair time in terms of time of the day, day of the week, and season of the year. More details on reliability parameters are given in [16].

TABLE III
BASIC ENERGY PRICING PARAMETERS

Parameter	Value	Unit
Electricity mean level	3,61	ln($/MWh)
Electricity start price	3,73	ln($/MWh)
Electricity reversion rate	3,13	1/year
Electricity price volatility	0,41	1/year
Natural gas mean level	1,35	ln($/MWh)
Natural gas start price	1,31	ln($/MWh)
Natural gas reversion rate	1,69	1/year
Natural gas price volatility	0,39	1/year
Natural gas-electricity price correlation	0,82	-

B. Base Case

A Monte Carlo simulation is run for each investment alternative. The simulations contain 100 runs each, which is chosen as a compromise between computation time and the resolution of obtained results. The first conclusion for the base case can be drawn from distributions of profit NPV shown in Fig. 6. Since the cumulative probability functions do not intersect, it is possible to make unambiguous order of investments by the quality: (1) C800, (2) C600, (3) C1000, (4) C400, (5) C200 i (6) C65. Fig. 7 shows that the maximum expected value of the profit NPV corresponds to the rated power of 800kW. Expected profit for C65 is even negative, which makes the alternative G1 absolutely inappropriate for further analysis.

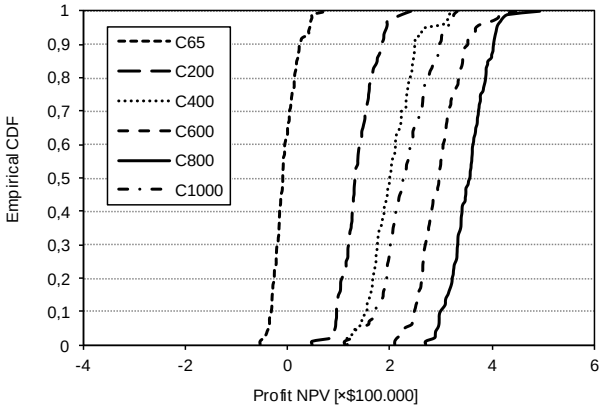


Fig. 6. Base case: Empirical CDFs for considered investment alternatives

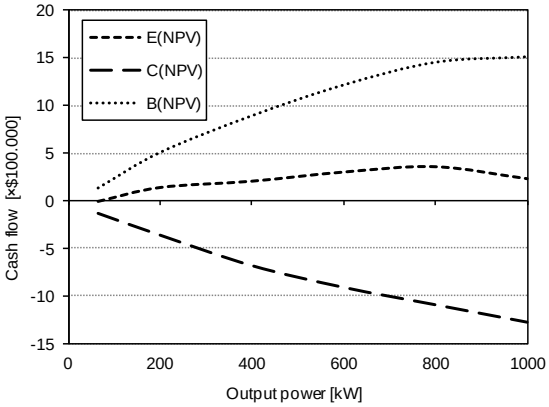


Fig. 7. Base case: Cash flows in terms of microturbine output power

C. Sensitivity Analysis

Errors in the input parameters may lead to erroneous results and conclusions. Therefore, the analysis is performed in order to determine how the results are influenced by the change in particular inputs. In this paper we show some representative results of our comprehensive sensitivity analysis. The investigation is limited to three most promising alternatives, namely G4-G6.

1) Case A: Electricity Price Volatility

Case A covers the impact of electricity price volatility. Fig. 8 shows that the increase in volatility of approximately 50% leads to a just slight increase in profit NPV, for each investment alternative. The increase is logical since the greater deviation in electricity price provides room for distributed generation to participate more effectively both in peak shaving and energy generation. Fig. 9 shows that the increase in deviation of the electricity price also leads to the larger deviation of the profit NPV. Obviously, C800 remains the best alternative for the customer.

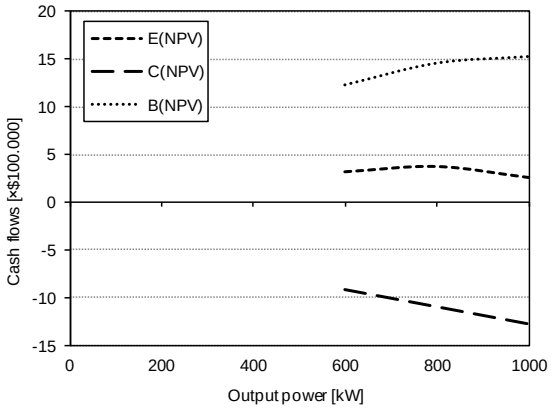


Fig. 8. Case A: Cash flows in terms of microturbine output power

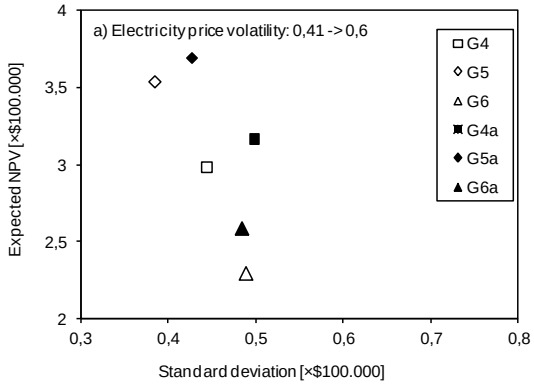


Fig. 9. Case A: Changes in expected value and standard deviation of profit NPV due to increase in electricity price volatility.

2) Case B: Natural Gas to Electricity Price Correlation

In the base case, the correlation between the prices of natural gas and electricity is assumed to be high, which is appropriate for the markets where a large portion of electricity is produced by gas power plants. The adopted value of the correlation coefficient ρ in the formula (3) is 0,82. In this analysis, the initial value of the correlation coefficient is decreased down to $\rho = 0{,}41$. Fig. 10 shows that halving the correlation factor significantly affects the standard deviation of NPV, while the expected values are left almost unchanged. The interpretation of this phenomenon is simple. When the prices of gas and electricity change with the correlated trajectories, the ratio of their prices are not changed within wide limits, so the customer savings are similar from month to month. When the correlation decreases, the ratio of the price of gas and electricity oscillates more. Therefore, in cases with expensive electricity the savings are getting greater and in cases with expensive gas the savings are getting lower. This increases the deviation of the results, but the average value of the profit does not experience a remarkable change.

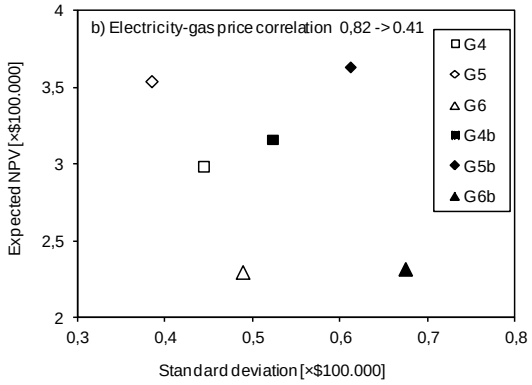


Fig. 10. Case B: Changes in expected value and standard deviation of profit NPV due to decrease in correlation between the prices of natural gas and electricity.

3) Case C: Natural Gas to Electricity Mean Level Ratio

The gas/electricity price ratio is one of the most important factors that dictate the amount of achievable savings for the customer. Case C will test the impact of this variable. The mean level of natural gas price logarithm is therefore changed from initial 1,35 to an arbitrary new value of 1,54. As might be expected, the impact on the achievable savings will be very strong. Because of the expensive gas, distributed generation will be used much rarely, mainly just for the peak shaving. Results in Fig. 11 show that the alternative with C1000 very likely finishes its life time with negative NPV of the profit. The alternatives C600 and C800 keep the profits positive, but the expected values of NPV are more than halved in comparison with the base case. Once again, C800 remains the best investment alternative.

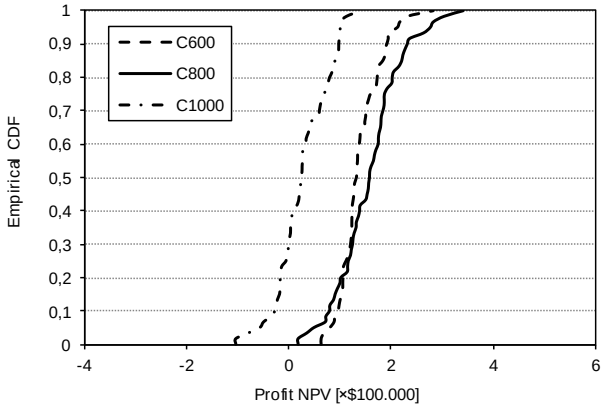


Fig. 11. Case C: Empirical CDFs for considered investment alternatives.

4) Case D: Load deviation

Case D covers the influence of uncertainty in customer load to the amount of savings achievable by DG. The standard deviation of the customer load from the base case is doubled here and the Monte Carlo simulations are performed again for the alternatives G4-G6. Increased deviation of the load logically leads to an increase in the monthly peak demand. Therefore, C1000 comes into play since it has the largest potential for the peak shaving. As per Fig. 12 the overall performance of C1000 almost reaches the results achievable by C800. C600 as the facility with the smallest output power is not significantly affected by the change in the load deviation.

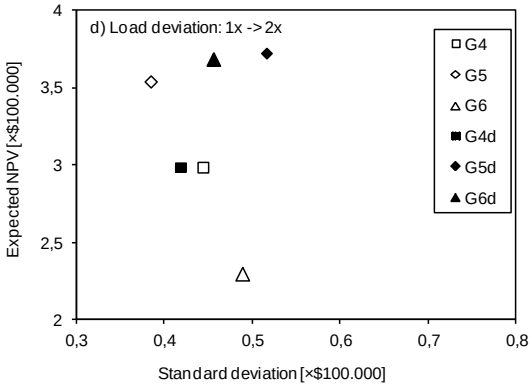


Fig. 12. Case D: Changes in expected value and standard deviation of profit NPV due to increase in deviation of the customer load.

D. Contribution in Reliability Improvement

It is interesting to determine which portion of the total benefits is related to improved reliability. According to principles described in [16], the customer loads are sorted in a priority order, so as in case of the grid outage, the most critical loads are served first by distributed generation. Although the less critical loads are left disconnected, the customer does not incur a significant additional damage. For the customer considered in this paper, microturbine facility C600 with 600kW rated power is sufficient to cover the most critical loads. Therefore, C800 and C1000 with their surplus of output power are not worth much in terms of reliability. Fig. 13 illustrates that the benefits of improved reliability are practically the same for each investment alternative. The expected present value of reliability benefits are about 16,5% of the overall customer benefits. With modifications of input parameters defined in cases A to D, the results obtained for the base case are not significantly changed.

E. Final Decision Based on Power Aspects

Using the results of presented comprehensive analysis, the customer would select the alternative based on C800 microturbine facility. This alternative shows the best results for the wide range of change in crucial input parameters. The risk of loss (i.e. negative profit NPV) is negligible.

The customer primarily benefits on cutting costs for energy supply. The benefits of improved reliability are just 1/6 of the total benefits. It should be noted that the ratio between particular benefits can be significantly changed under different conditions. For example, for the customers connected to unreliable distribution grid, the benefits of improved reliability can be even greater than the benefits of cheaper energy supply.

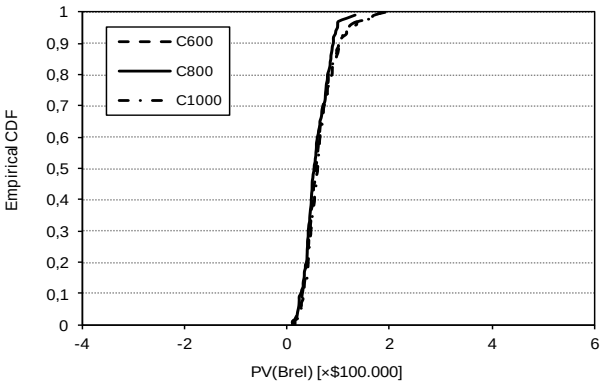


Fig. 13. Base case: Contribution of reliability improvement to the total profit NPV

F. Combined Heat and Power

Microturbine exhaust gases contain a respectable amount of heat energy. Heat to power ratio for microturbine technology is regularly greater than 1,5. By investing in a system of heat exchangers, available heat energy might be used for space heating, water heating, etc. It depends on the nature of the customer how much of the available heat can be effectively utilized.

We tested how much heat energy is available for six investment alternatives (G1–G6), while the microturbines are dispatched according to the proposed ETC algorithm. It is assumed that the customer used their own boiler to produce heat, prior to the investment in distributed generation. Natural gas is bough at the retail prices, as same as for microturbine purposes. The boiler efficiency is set to 80%. We calculated the amount of avoided purchase of natural gas thanks to the use of heat recovered from the microturbines. The heat exchanger investment costs are deducted from the value of benefit in order to compute the net profit of the investment.

Fig. 14 shows that the upgrade in CHP has a huge theoretical maximum potential. If total available heat energy would be utilized, the investment in CHP would be very profitable. On the other hand, it is not easy to achieve a perfect matching between the load profiles of electricity and heat. In a simple case of space heating, the available heat would be used only during the colder months.

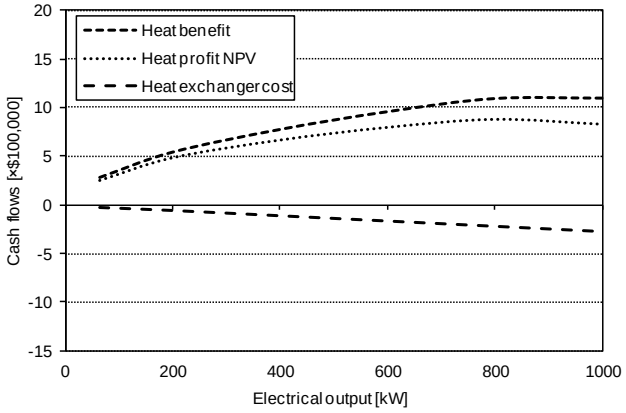


Fig. 14. Cost-effectiveness of investment in microturbine heat exchangers

In order to precisely calculate the benefits of recovered heat, a detailed model of customer heat load profile is necessary. In this paper, we determine just the profitability thresholds expressed as a percentage of the theoretical maximum, where the investment in CHP upgrade is still profitable for the customer. The results are summarized in Table IV.

It should be noted that the investments in CHP can be subsidized under certain circumstances, which would additionally lower the profitability thresholds.

TABLE IV
PROFITABILITY THRESHOLDS FOR INVESTMENT IN CHP UPGRADE

	Unit	C65	C200	C400
CHP investment cost NPV	$\times 10^5$\$	0,24	0,55	1,09
Benefit NPV	$\times 10^5$\$	2,79	5,44	7,76
Profitability threshold	%	8,5	10,1	14,1

	Unit	C600	C800	C1000
CHP investment cost NPV	$\times 10^5$\$	1,64	2,19	2,73
Benefit NPV	$\times 10^5$\$	9,61	10,98	11,02
Profitability threshold	%	17,1	19,9	24,8

V. CONCLUSION

In this paper the profitability of investing in customer-driven distributed generation has been considered. The target customers are industrial and commercial users of electricity, which are billed not only for the volumetric consumption but also for the monthly peak demand. Presented methodology is capable of providing the answer whether or not it is profitable to invest in distributed generation, and which is the most appropriate investment option for the customer. The proposed methodology has been comprehensively tested on one integrated illustrative example.

REFERENCES

[1] G. Pepermans, J. Driesen, D. Haeseldonckx, R. Belmans, and W. D'haeseleer, "Distributed generation: definition, benefits and issues", *Energy Policy*, vol. 33, pp. 787-798, 2005.

[2] A. A. Chowdhury, S. K. Agarwal, and D. O. Koval, "Reliability modeling of distributed generation in conventional distribution systems planning and analysis", *IEEE Transactions on Industry Applications*, vol. 39, pp. 1493-1498, 2003.

[3] A. Moreno-Munoz, J. J. G. de-la-Rosa, M. A. Lopez-Rodriguez, J. M. Flores-Arias, F. J. Bellido-Outerino, and M. Ruiz-de-Adana, "Improvement of power quality using distributed generation", *International Journal of Electrical Power Energy Systems*, vol. 32 pp. 1069-1076, 2010.

[4] H. A. Gil and G. Joos, "On the quantification of the network capacity deferral value of distributed generation", *IEEE Transactions on Power Systems*, vol. 21, pp. 1592-1599, 2006.

[5] H. A. Gil and G. Joos, "Models for quantifying the economic benefits of distributed generation", *IEEE Transactions on Power Systems*, vol. 23 pp. 327-335, 2008.

[6] I. -S. Bae, J. O. Kim, J. -C. Kim and C. Singh, "Optimal operating strategy for distributed generation considering hourly reliability worth", *IEEE Transactions on Power Systems*, vol. 19, pp. 287-292, 2004.

[7] A. Barina, L. F. Pozzatti, L. N. Canha, R. Q. Machado, A. R. Abaide, and G. Arend, "Multi-objective analysis of impacts of distributed generation placement on the operational characteristics of networks for distribution system planning", *International Journal of Electrical Power Energy Systems*, vol. 32, pp. 1157-1164, 2010.

[8] M. Zangiabadi, R. Feuillet, H. Lesani, N. Hadj-Said, and J. T. Kvaløy, "Assessing the performance and benefits of customer distributed generation developers under uncertainties", *Energy*, vol. 36, pp. 1703-1712, 2011.

[9] C. Marnay, J. Chard, K. Hamachi, T. Lipman, M. Moezzi, B. Ouaglal, and A.Siddiqui, "Modeling of customer adoption of distributed energy resources", Lawrence Berkeley National Laboratory Report LBNL-49582, Berkeley, CA, USA, 2001.

[10] A. S. Siddiqui, C. Marnay, R. M. Firestone, and N. Zhou, "Distributed generation with heat recovery and storage", Journal of Energy Engineering, vol. 133, pp. 181-210, 2007.

[11] Y. Ruan, Q. Liu, W. Zhou, R. Firestone, W. Gao, and T. Watanabe, "Optimal option of distributed generation technologies for various commercial buildings", Applied Energy, vol. 86, pp. 1641-1653, 2009.

[12] B. Coffey and E. Kutrowski, "Demand charge considerations in the optimization of cogeneration dispatch in a deregulated energy market", International Journal of Energy Research, vol. 30, pp. 535-551, 2006.

[13] A. S. Siddiqui and C. Marnay, "Distributed generation investment by a microgrid under uncertainty", *Energy*, vol. 33, pp. 1729-1737, 2008.

[14] Č. V. Zeljković, N. Lj. Rajaković, and S. J. Zubić, "A method for cost minimization applicable to load centers containing distributed generation", in *Proc. IEEE PowerTech Conf.*, Bucharest Romania, June 28 – July 2, 2009, pp. 1-6.

[15] Č. V. Zeljković, N. Lj. Rajaković, and S. J. Zubić, "An application of cost minimization algorithm to economic justification of installing distributed generation", in *Proc. IFAC CMTEE Conference* Vilamoura, Portugal, March 29-31, 2010, pp. 243-248.

[16] Č. V. Zeljković, N. Lj. Rajaković, and S. J. Zubić, "Customer-perspective approach to reliability evaluation of distributed generation", in *Proc. IEEE PowerTech Conf.*, Trondheim Norway, June 19-23, 2011.

[17] R. A. Brealey and S. C. Myers, *Principles of corporate finance*, 7th ed., Boston: McGraw-Hill, 2003.

[18] S. C. Savvides, "Risk analysis in investment appraisal", Project Appraisal, vol. 9, pp. 3-18, 1994.

[19] Č. Zeljković and N. Rajaković, "Cost-saving potential of customer-driven distributed generation," *Electric Power Systems Research*, vol. 92, pp. 87-95, 2012.

[20] J. Nowicka-Zagrajek and R. Weron, "Modeling electricity loads in California: ARMA models with hyperbolic noise", *Signal Processing*, vol. 82, pp. 1903-1915, 2002.

[21] B. Klöckl, G. Papaefthymiou, Multivariate time series models for studies on stochastic generators in power systems, Electr.Power Syst. Res. 80 (2010) 265-276.

[22] G. H. Kjølle, and A. T. Holen, Reliability and interruption cost prediction using time-dependent failure rates and interruption costs, Qual. Reliab. Eng. Int.14 (1998) 159-165.

[23] R. Billinton W. Li, Reliability Assessment of electric power systems using Monte Carlo methods, New York: Plenum, 1994.

[24] L. Goel, X. Liang, and Y. Ou, Monte Carlo simulation-based customer service reliability assessment, Electr. Power Syst. Res. 49 (1999) 185-194.

[25] R. Billinton, R. N. Allan, Reliability evaluation of power systems, 2nd ed., New York and London: Plenum Press, 1996.

[26] R. E. Brown, Electric power distribution reliability, 2nd ed., Boca Raton: CRC Press, 2008.

[27] H. Y. Yamin, Review on methods of generation scheduling in electric power systems, Electr. Power Syst. Res. 69 (2004) 227-248.

[28] K. H. LaCommare, J. H. Eto, Cost of power interruptions to electricity consumers in the United States (US), *Energy*, vol. 31, pp. 1845-1855, 2006.

[29] K. M. Maribu and S. -E. Fleten, "Combined heat and power in commercial buildings: Investment and risk analysis," *The Energy Journal*, vol. 29, pp. 123-150, 2008.

[30] P. S. Curtiss and J. F. Kreider, "Recent developments in the control of distributed electrical generation systems", *ASME Journal of Solar Energy Engineering*, vol. 125, pp. 352-358, 2003.

[31] R. Firestone, M. Stadler, C. Marnay, "Integrated energy system dispatch optimization", *in Proc. 4th Annual IEEE Conference on Industrial Informatics*, Singapore, August 16-18, 2006, pp. 357-362.

Instructions for Authors

First A. Author, Second B. Author, and Third C. Author

***Abstract*—These instructions give you guidelines for preparing papers for ELECTRONICS journal. Use this document as a template if you are using Microsoft *Word* 6.0 or later. Otherwise, use this document as an instruction set. The electronic file of your paper will be formatted further. Define all symbols used in the abstract. Do not cite references in the abstract. Do not delete the blank line immediately above the abstract; it sets the footnote at the bottom of this column.**

***Index Terms*—About four key words or phrases in alphabetical order, separated by commas.**

I. Introduction

THIS document is a template for Microsoft *Word* versions 6.0 or later.

When you open the file, select "Page Layout" from the "View" menu in the menu bar (View | Page Layout), which allows you to see the footnotes. Then, type over sections of file or cut and paste from another document and use markup styles. The pull-down style menu is at the left of the Formatting Toolbar at the top of your *Word* window (for example, the style at this point in the document is "Text"). Highlight a section that you want to designate with a certain style, then select the appropriate name on the style menu. The style will adjust your fonts and line spacing. **Do not change the font sizes or line spacing to squeeze more text into a limited number of pages.** Use italics for emphasis; do not underline. The length of the manuscript is limited to the maximum of 15 pages.

To insert images in *Word*, position the cursor at the insertion point and either use Insert | Picture | From File or copy the image to the Windows clipboard and then Edit | Paste Special | Picture (with "float over text" unchecked).

We will do the final formatting of your paper.

II. Procedure for Paper Submission

A. Review Stage

The manuscripts are to be delivered to the editor by e-mail: electronics@etfbl.net. Prepare it in two-column format as shown in this template. Place all figures and tables at the end of the paper (after the references) on separate page(s). Figures and tables must have the same caption names as referenced in the text. Only PDF format of the manuscript is allowed at the review stage. Please, check if all fonts are embedded and subset and that the quality of diagrams, illustrations, and graphics is satisfactory. Failing to provide above listed requirements is a valid reason for rejection.

B. Final Stage

When you submit your final version (after your paper has been accepted), prepare it in two-column format, including figures and tables in accordance with this template. Pack all of your files (manuscript source file in *Word*, figures, and manuscript PDF form) within one archive file (you may use any of the available file compression tools: *WinZip*, *WinRAR*, *7-Zip*, etc.). Do not forget to provide the manuscript converted in PDF format that will be used as a reference for final formatting of your paper. Figures should be named as referenced in the manuscript (e.g. *fig1.eps*, *fig2.tif*, etc.)

C. Figures and Tables

Format and save your graphic images using a suitable graphics processing program and adjusts the resolution settings. We accept images in the following formats: PS, EPS, TIFF, GIF, and PNG. Additionally, it is allowed to use images generated by using one of the following software tools: Microsoft Word, Microsoft PowerPoint, or Microsoft Excel. The resolution of a RGB color file should be 400 dpi. Please note that JPG and other lossy-compressed image formats are not allowed. Use available software tools to convert these images to appropriate format.

Image quality is very important to how yours graphics will reproduce. Even though we can accept graphics in many formats, we cannot improve your graphics if they are poor quality when we receive them. If your graphic looks low in quality on your printer or monitor, please keep in mind that cannot improve the quality after submission.

Manuscript received 15 September 2011 (write the date when you have first sent the manuscript for review). Received in revised form 20 October 2011 (write the date when you have sent the manuscript in its revised form if revisions required for your paper after review).

(Place here any sponsor and financial support acknowledgments).

F. A. Author is with the Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, Bosnia and Herzegovina (corresponding author to provide phone: +387-51-222-333; fax: +387-51-111-222; e-mail: author@etfbl.net).

S. B. Author was with Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia. He is now with the Institute "Mihailo Pupin", Belgrade, Serbia (e-mail: author@pupin.rs).

T. C. Author is with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, on leave from the Faculty of Electronic Engineering, University of Niš, Niš, Serbia (e-mail: author@elfak.ni.ac.rs).

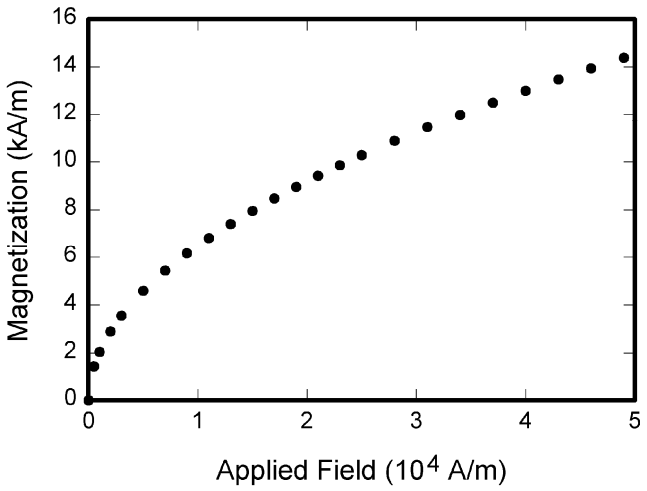


Fig. 1. Magnetization as a function of applied field. Note that "Fig." is abbreviated. There is a period after the figure number, followed by two spaces. It is good practice to explain the significance of the figure in the caption.

If you are importing your graphics into this Word template, please use the following steps:

Under the option EDIT select PASTE SPECIAL. A dialog box will open, select paste picture, then click OK. Your figure should now be in the Word Document.

If you are preparing images in TIFF, EPS, or PS format, note the following. High-contrast line figures and tables should be prepared with 600 dpi resolution and saved with no compression, 1 bit per pixel (monochrome).

Photographs and grayscale figures should be prepared with 300 dpi resolution and saved with no compression, 8 bits per pixel (grayscale).

Most charts graphs and tables are one column wide (3 1/2 inches or 21 picas) or two-column width (7 1/16 inches, 43 picas wide). We recommend that you avoid sizing figures less than one column wide, as extreme enlargements may distort your images and result in poor reproduction. Therefore, it is better if the image is slightly larger, as a minor reduction in size should not have an adverse affect the quality of the image.

III. Math

If you are using *Word*, use either the Microsoft Equation Editor or the *MathType* add-on (http://www.mathtype.com) for equations in your paper (Insert | Object | Create New | Microsoft Equation *or* MathType Equation). "Float over text" should *not* be selected.

IV. Units

Use either SI (MKS) or CGS as primary units. (SI units are strongly encouraged.) English units may be used as secondary units (in parentheses). **This applies to papers in data storage.** For example, write "15 Gb/cm^2 (100 Gb/in^2)." An exception is when English units are used as identifiers in trade, such as "3½-in disk drive." Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity in an equation.

TABLE I
Units for Magnetic Properties

Symbol	Quantity	Conversion from Gaussian and CGS EMU to SI [a]
Φ	magnetic flux	1 Mx $\rightarrow 10^{-8}$ Wb = 10^{-8} V·s
B	magnetic flux density, magnetic induction	1 G $\rightarrow 10^{-4}$ T = 10^{-4} Wb/m^2
H	magnetic field strength	1 Oe $\rightarrow 10^3/(4\pi)$ A/m
m	magnetic moment	1 erg/G = 1 emu $\rightarrow 10^{-3}$ A·m^2 = 10^{-3} J/T
M	magnetization	1 erg/(G·cm^3) = 1 emu/cm^3 $\rightarrow 10^3$ A/m
$4\pi M$	magnetization	1 G $\rightarrow 10^3/(4\pi)$ A/m
σ	specific magnetization	1 erg/(G·g) = 1 emu/g $\rightarrow$ 1 A·m^2/kg
j	magnetic dipole moment	1 erg/G = 1 emu $\rightarrow 4\pi \times 10^{-10}$ Wb·m
J	magnetic polarization	1 erg/(G·cm^3) = 1 emu/cm^3 $\rightarrow 4\pi \times 10^{-4}$ T
χ, κ	susceptibility	$1 \rightarrow 4\pi$
χ_ρ	mass susceptibility	1 cm^3/g $\rightarrow 4\pi \times 10^{-3}$ m^3/kg
μ	permeability	$1 \rightarrow 4\pi \times 10^{-7}$ H/m = $4\pi \times 10^{-7}$ Wb/(A·m)
μ_r	relative permeability	$\mu \rightarrow \mu_r$
w, W	energy density	1 erg/cm^3 $\rightarrow 10^{-1}$ J/m^3
N, D	demagnetizing factor	$1 \rightarrow 1/(4\pi)$

Vertical lines are optional in tables. Statements that serve as captions for the entire table do not need footnote letters.

[a] Gaussian units are the same as cgs emu for magnetostatics; Mx = maxwell, G = gauss, Oe = oersted; Wb = weber, V = volt, s = second, T = tesla, m = meter, A = ampere, J = joule, kg = kilogram, H = henry.

The SI unit for magnetic field strength H is A/m. However, if you wish to use units of T, either refer to magnetic flux density B or magnetic field strength symbolized as $\mu_0 H$. Use the center dot to separate compound units, e.g., "A·m^2."

V. Helpful Hints

A. Figures and Tables

Because we will do the final formatting of your paper, you do not need to position figures and tables at the top and bottom of each column. In fact, all figures, figure captions, and tables can be at the end of the paper. Large figures and tables may span both columns. Place figure captions below the figures; place table titles above the tables. If your figure has two parts, include the labels "(a)" and "(b)" as part of the artwork. Please verify that the figures and tables you mention in the text actually exist. **Please do not include captions as part of the figures. Do not put captions in "text boxes" linked to the figures. Do not put borders around the outside of your figures.** Use the abbreviation "Fig." even at the beginning of a sentence. Do not abbreviate "Table." Tables are numbered with Roman numerals.

Color printing of figures is not available Do not use color unless it is necessary for the proper interpretation of your figures.

Figure axis labels are often a source of confusion. Use words rather than symbols. As an example, write the quantity

"Magnetization," or "Magnetization *M*," not just "*M*." Put units in parentheses. Do not label axes only with units. As in Fig. 1, for example, write "Magnetization (A/m)" or "Magnetization ($A \cdot m^{-1}$)," not just "A/m." Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)," not "Temperature/K."

Multipliers can be especially confusing. Write "Magnetization (kA/m)" or "Magnetization (10^3 A/m)." Do not write "Magnetization (A/m) × 1000" because the reader would not know whether the top axis label in Fig. 1 meant 16000 A/m or 0.016 A/m. Figure labels should be legible, approximately 8 to 12 point type.

B. References

Number citations consecutively in square brackets [1]. The sentence punctuation follows the brackets [2]. Multiple references [2], [3] are each numbered with separate brackets [1]–[3]. When citing a section in a book, please give the relevant page numbers [2]. In sentences, refer simply to the reference number, as in [3]. Do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] shows ... ." Please do not use automatic endnotes in *Word*, rather, type the reference list at the end of the paper using the "References" style.

Number footnotes separately in superscripts (Insert | Footnote).[1] Place the actual footnote at the bottom of the column in which it is cited; do not put footnotes in the reference list (endnotes). Use letters for table footnotes (see Table I).

Please note that the references at the end of this document are in the preferred referencing style. Give all authors' names; do not use "*et al*." unless there are six authors or more. Use a space after authors' initials. Papers that have not been published should be cited as "unpublished" [4]. Papers that have been accepted for publication, but not yet specified for an issue should be cited as "to be published" [5]. Papers that have been submitted for publication should be cited as "submitted for publication" [6]. Please give affiliations and addresses for private communications [7].

Capitalize only the first word in a paper title, except for proper nouns and element symbols. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [8]. All references **must be** written in Roman alphabet.

C. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have already been defined in the abstract. Abbreviations such as IEEE, SI, ac, and dc do not have to be defined. Abbreviations that incorporate periods should not have spaces: write "C.N.R.S.," not "C. N. R. S." Do not use abbreviations in the title unless they are unavoidable (for example, "IEEE" in the title of this article).

[1]It is recommended that footnotes be avoided (except for the unnumbered footnote with the receipt date and authors' affiliations on the first page). Instead, try to integrate the footnote information into the text.

D. Equations

Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). First use the equation editor to create the equation. Then select the "Equation" markup style. Press the tab key and write the equation number in parentheses. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Use parentheses to avoid ambiguities in denominators. Punctuate equations when they are part of a sentence, as in

$$\int_0^{r_2} F(r,\varphi)\, dr\, d\varphi = [\sigma r_2 / (2\mu_0)] \cdot \int_0^{\infty} \exp(-\lambda\, |z_j - z_i|)\, \lambda^{-1} J_1(\lambda r_2)\, J_0(\lambda r_i)\, d\lambda \,. \tag{1}$$

Be sure that the symbols in your equation have been defined before the equation appears or immediately following. Italicize symbols (*T* might refer to temperature, but T is the unit tesla). Refer to "(1)," not "Eq. (1)" or "equation (1)," except at the beginning of a sentence: "Equation (1) is ... ."

E. Other Recommendations

Use one space after periods and colons. Hyphenate complex modifiers: "zero-field-cooled magnetization." Avoid dangling participles, such as, "Using (1), the potential was calculated." [It is not clear who or what used (1).] Write instead, "The potential was calculated by using (1)," or "Using (1), we calculated the potential."

Use a zero before decimal points: "0.25," not ".25." Use "cm^3," not "cc." Indicate sample dimensions as "0.1 cm × 0.2 cm," not "0.1 × 0.2 cm^2." The abbreviation for "seconds" is "s," not "sec." Do not mix complete spellings and abbreviations of units: use "Wb/m^2" or "webers per square meter," not "webers/m^2." When expressing a range of values, write "7 to 9" or "7-9," not "7~9."

A parenthetical statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.) In American English, periods and commas are within quotation marks, like "this period." Other punctuation is "outside"! Avoid contractions; for example, write "do not" instead of "don't." The serial comma is preferred: "A, B, and C" instead of "A, B and C."

If you wish, you may write in the first person singular or plural and use the active voice ("I observed that ..." or "We observed that ..." instead of "It was observed that ..."). Remember to check spelling. If your native language is not English, please get a native English-speaking colleague to carefully proofread your paper.

VI. Some Common Mistakes

The word "data" is plural, not singular. The subscript for the permeability of vacuum μ_0 is zero, not a lowercase letter "o." The term for residual magnetization is "remanence"; the adjective is "remanent"; do not write "remnance" or "remnant." Use the word "micrometer" instead of "micron." A

graph within a graph is an “inset,” not an “insert.” The word “alternatively” is preferred to the word “alternately” (unless you really mean something that alternates). Use the word “whereas” instead of “while” (unless you are referring to simultaneous events). Do not use the word “essentially” to mean “approximately” or “effectively.” Do not use the word “issue” as a euphemism for “problem.” When compositions are not specified, separate chemical symbols by en-dashes; for example, “NiMn” indicates the intermetallic compound $Ni_{0.5}Mn_{0.5}$ whereas “Ni–Mn” indicates an alloy of some composition Ni_xMn_{1-x}.

Be aware of the different meanings of the homophones “affect” (usually a verb) and “effect” (usually a noun), “complement” and “compliment,” “discreet” and “discrete,” “principal” (e.g., “principal investigator”) and “principle” (e.g., “principle of measurement”). Do not confuse “imply” and “infer.”

Prefixes such as “non,” “sub,” “micro,” “multi,” and “ultra” are not independent words; they should be joined to the words they modify, usually without a hyphen. There is no period after the “et” in the Latin abbreviation “*et al.*” (it is also italicized). The abbreviation “i.e.,” means “that is,” and the abbreviation “e.g.,” means “for example” (these abbreviations are not italicized).

An excellent style manual and source of information for science writers is [9].

VII. Editorial Policy

Each manuscript submitted is subjected to the following review procedure:

- It is reviewed by the editor for general suitability for this publication
- If it is judged suitable, two reviewers are selected and a single-blinded review process takes place
- Based on the recommendations of the reviewers, the editor then decides whether the particular paper should be accepted as is, revised or rejected.

Do not submit a paper you have submitted or published elsewhere. Do not publish “preliminary” data or results. The submitting author is responsible for obtaining agreement of all coauthors and any consent required from sponsors before submitting a paper. It is the obligation of the authors to cite relevant prior work.

Every paper submitted to “Electronics” journal are single-blind reviewed. For conference-related papers, the decision to accept or reject a paper is made by the conference editors and publications committee; the recommendations of the referees are advisory only. Undecipherable English is a valid reason for rejection.

VIII. Publication Principles

The contents of “Electronics” are peer-reviewed and archival. The “Electronics” publishes scholarly articles of archival value as well as tutorial expositions and critical reviews of classical subjects and topics of current interest.

Authors should consider the following points:

1) Technical papers submitted for publication must advance the state of knowledge and must cite relevant prior work.
2) The length of a submitted paper should be commensurate with the importance, or appropriate to the complexity, of the work. For example, an obvious extension of previously published work might not be appropriate for publication or might be adequately treated in just a few pages.
3) Authors must convince both peer reviewers and the editors of the scientific and technical merit of a paper; the standards of proof are higher when extraordinary or unexpected results are reported.
4) Because replication is required for scientific progress, papers submitted for publication must provide sufficient information to allow readers to perform similar experiments or calculations and use the reported results. Although not everything need be disclosed, a paper must contain new, useable, and fully described information. For example, a specimen’s chemical composition need not be reported if the main purpose of a paper is to introduce a new measurement technique. Authors should expect to be challenged by reviewers if the results are not supported by adequate data and critical details.
5) Papers that describe ongoing work or announce the latest technical achievement, which are suitable for presentation at a professional conference, may not be appropriate for publication in “Electronics”.

IX. Conclusion

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

Appendix

Appendixes, if needed, appear before the acknowledgment.

Acknowledgment

The preferred spelling of the word “acknowledgment” in American English is without an “e” after the “g.” Use the singular heading even if you have many acknowledgments. Avoid expressions such as “One of us (S.B.A.) would like to thank ... .” Instead, write “F. A. Author thanks ... .” **Sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page, not here.**

References

[1] G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

[2] W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[3] H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.

[4] B. Smith, “An approach to graphs of linear forms (Unpublished work style),” unpublished.
[5] E. H. Miller, “A note on reflector arrays (Periodical style—Accepted for publication),” *IEEE Trans. Antennas Propagat.*, to be published.
[6] J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.
[7] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
[8] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces (Translation Journals style),” *IEEE Transl. J. Magn.Jpn.*, vol. 2, Aug. 1987, pp. 740–741 [*Dig. 9th Annu. Conf. Magnetics* Japan, 1982, p. 301].
[9] M. Young, *The Techincal Writers Handbook.* Mill Valley, CA: University Science, 1989.
[10] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Periodical style),” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34–39, Jan. 1959.
[11] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. Neural Networks*, vol. 4, pp. 570–578, Jul. 1993.
[12] R. W. Lucky, “Automatic equalization for digital communication,” *Bell Syst. Tech. J.*, vol. 44, no. 4, pp. 547–588, Apr. 1965.
[13] S. P. Bingulac, “On the compatibility of adaptive controllers (Published Conference Proceedings style),” in *Proc. 4th Annu. Allerton Conf. Circuits and Systems Theory*, New York, 1994, pp. 8–16.
[14] G. R. Faulhaber, “Design of service systems with priority reservation,” in *Conf. Rec. 1995 IEEE Int. Conf. Communications*, pp. 3–8.
[15] W. D. Doyle, “Magnetization reversal in films with biaxial anisotropy,” in *1987 Proc. INTERMAG Conf.*, pp. 2.2-1–2.2-6.
[16] G. W. Juette and L. E. Zeffanella, “Radio noise currents n short sections on bundle conductors (Presented Conference Paper style),” presented at the IEEE Summer power Meeting, Dallas, TX, Jun. 22–27, 1990, Paper 90 SM 690-0 PWRS.
[17] J. G. Kreifeldt, “An analysis of surface-detected EMG as an amplitude-modulated noise,” presented at the 1989 Int. Conf. Medicine and Biological Engineering, Chicago, IL.
[18] J. Williams, “Narrow-band analyzer (Thesis or Dissertation style),” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993.
[19] N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.
[20] J. P. Wilkinson, “Nonlinear resonant circuit devices (Patent style),” U.S. Patent 3 624 12, July 16, 1990.
[21] *IEEE Criteria for Class IE Electric Systems* (Standards style), IEEE Standard 308, 1969.
[22] *Letter Symbols for Quantities*, ANSI Standard Y10.5-1968.
[23] R. E. Haskell and C. T. Case, “Transient signal propagation in lossless isotropic plasmas (Report style),” USAF Cambridge Res. Lab., Cambridge, MA Rep. ARCRL-66-234 (II), 1994, vol. 2.
[24] E. E. Reber, R. L. Michell, and C. J. Carter, “Oxygen absorption in the Earth’s atmosphere,” Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (420-46)-3, Nov. 1988.
[25] (Handbook style) *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, 1985, pp. 44–60.
[26] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.
[27] (Basic Book/Monograph Online Sources) J. K. Author. (year, month, day). *Title* (edition) [Type of medium]. Volume (issue). Available: http://www.(URL)
[28] J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: http://www.atm.com
[29] (Journal Online Sources style) K. Author. (year, month). Title. *Journal* [Type of medium]. Volume(issue), paging if given. Available: http://www.(URL)
[30] R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. *IEEE Trans. Plasma Sci.* [Online]. *21(3).* pp. 876–880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

Instruction for Authors

Editorial objectives

In the journal ”Electronics”, the scientific papers from different fields of electronics and electrical engineering in the broadest sense are published. Main topics are electronics, automatics, telecommunications, computer techniques, power engineering, nuclear and medical electronics, analysis and synthesis of electronic circuits and systems, new technologies and materials in electronics etc. The main emphasis of papers should be on methods and new techniques, or the application of existing techniques in a novel way.

The reviewing process

Each manuscript submitted is subjected to the following review procedures:

- It is reviewed by the editor for general suitability for this publication;
- If it is judged suitable, two reviewers are selected and a double-blind review process takes place;
- Based on the recommendations of the reviewers, the editor then decides whether the particular paper should be accepted as it is, revised or rejected.

Submissions Process

The manuscripts are to be delivered to the editor by e-mail: electronics@etfbl.net.

Manuscripts have to be prepared in accordance with the instructions given in the template for paper preparation that can be found on the journal’s web page (www.electronics.etfbl.net).

Authors should note that proofs are not supplied prior to publication and ensure that the paper submitted is complete and in its final form.

Copyright

Articles submitted to the journal should be original contributions and should not be under consideration for any other publication at the same time. Authors submitting articles for publication warrant that the work is not an infringement of any existing copyright and will indemnify the publisher against any breach of such warranty. For ease of dissemination and to ensure proper policing of use, papers and contributions become the legal copyright of the publisher unless otherwise agreed.

ELECTRONICS, VOL. 18, NO. 1, JUNE 2014

PAPERS

EDITORIAL **1**

ENERGY EFFICIENT MULTI-CORE PROCESSING **3**
Charles Leech and Tom J. Kazmierski

SPICE MODELING AND SIMULATION OF A MPPT ALGORITHM **11**
Miona Andrejević Stošović, Marko Dimitrijević, Duško Lukač, and Vančo Litovski

THE USE OF BAYESIAN NETWORKS IN DETECTING THE STATES OF VENTILATION MILLS IN POWER PLANTS **16**
Sanja Vujnović, Predrag Todorov, Željko Đurović, and Aleksandra Marjanović

USB HW/SW CO-SIMULATION ENVIRONMENT WITH CUSTOM TEST TOOL INTEGRATION **23**
Zargaryan Y. Grigor, Aharonyan K. Vahram, Melikyan V. Nazeli and Marko A. Dimitrijević

MEASURING CAPACITOR PARAMETERS USING VECTOR NETWORK ANALYZERS **29**
Deniss Stepins, Gundars Asmanis, and Aivis Asmanis

DATA-DRIVEN GRADIENT DESCENT DIRECT ADAPTIVE CONTROL FOR DISCRETE-TIME NONLINEAR SISO SYSTEMS **39**
Igor R. Krcmar, Milorad M. Bozic, and Petar S. Maric

A MODIFIED APPROXIMATE INTERNAL MODEL-BASED NEURAL CONTROL FOR THE TYPICAL INDUSTRIAL PROCESSES **46**
Jasmin Igic and Milorad Bozic

INTEGRATED COST-BENEFIT ASSESSMENT OF CUSTOMER-DRIVEN DISTRIBUTED GENERATION **54**
Čedomir Zeljković and Nikola Rajaković

ISSN 1450-5843

UDK 621.38