4th

PROCEEDINGS
of the

Small Systems Simulation Symposium 2012

February 12-14
Faculty of Electronic Engineering
Niš, Serbia

organized by the
Yugoslav Simulation Society
and the
Faculty of Electronic Engineering Niš

SSSS 2012

Proceedings

CONTENTS

Design of ultra-low-energy wireless sensor nodes powered by kinetic harvesters
Invited Paper

Tom J Kaźmierski

Abstract—**In an energy harvester powered wireless sensor node system, as the energy harvester is the only energy source, it is crucial to configure the microcontroller and the sensor node so that the harvested energy is used efficiently. This paper outlines modelling, performance optimisation and design exploration of the complete, complex system which includes the analog mechanical model of a tunable kinetic microgenerator, its magnetic coupling with the electrical blocks, electrical power storage and processing parts, the digital control of the microgenerator tuning system, as well as the power consumption models of sensor node. Therefore not only the energy harvester design parameters but also the sensor node operation parameters can be optimized in order to achieve the best system performance. The power consumption models of the microcontroller and the sensor node are built based on their operation scenarios so that the parameters of the digital algorithms can be optimised to achieve the best energy efficiency. In the proposed approach, two Hardware Description Languages, VHDL-AMS and SystemC-A is used to model the system's analogue components as well as the digital control algorithms which are implemented in the microcontroller and the sensor node. Simulation and performance optimisation results are verified experimentally. In the development of the fast design exploration tool based on the response surface technique, the response surface model (RSM) is constructed by carrying out a series of simulations. The RSM is then optimised using MATLAB's optimisation toolbox and the optimisation results are presented.**

I. INTRODUCTION

Wireless sensor networks (WSNs) have attracted a great research interest in recent years. Since wireless sensor nodes can provide information from previously inaccessible locations and from previously unachievable number of locations, many new application areas are emerging, such as environmental sensing [1], structural monitoring [2] and human body monitoring [3]. Although wireless sensor nodes are easy to deploy, the lack of physical connection means they must have their own energy supply. Because batteries have limited lifetime and are environmentally hazardous, it has become widely agreed that energy harvesters are needed for long-lasting sensor nodes [4–6]. The idea is to use energy harvester to capture small amounts of energy from the environment and use the generated energy to power the nodes in wireless sensor networks.

Vibration-based energy harvesters are used in many commercial applications since mechanical vibrations are widely

Tom J Kaźmierski is with the Faculty of Physical and Applied Sciences, University of Southampton, Southampton SO17 1BJ, UK, email:tjk@ecs.soton.ac.uk

present. Most of the reported vibration energy harvester designs are based on a spring-mass-damper system with a characteristic resonant frequency. These devices normally have a high Q-factor and generate maximum power when their resonant frequency matches the dominant frequency of the input ambient vibration [7]. Consequently, the output power generated by the microgenerator drops dramatically when there is a difference between the dominant ambient frequency and the microgenerator's resonant frequency. Tunable microgenerators, which can adjust their own resonant frequency through mechanical or electrical methods to match the input frequency, are therefore more desirable than the fixed frequency microgenerators [8]. A wireless sensor node powered by tunable energy harvester typically has the following key components (Fig. 1) [9]: a microgenerator which converts ambient environment vibration into electrical energy, a power processing circuit which regulates and stores the generated energy, an actuator used for the frequency tuning mechanism, a digital controller that monitors and retunes the tunable energy harvesting system based on vibration measurements from an accelerometer, and the wireless transceiver or transmitter.
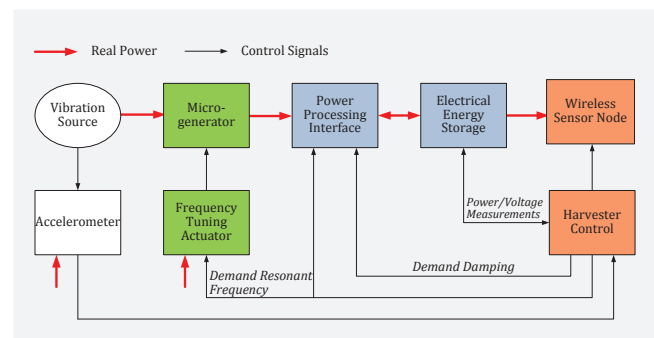
Fig. 1. Components of a energy harvester powered sensor node system [9].

Hardware description languages, such as VHDL-AMS and SystemC-A, have been used to model energy harvesters in recent years [10, 11]. HDLs with mixed signal and multi-domain capabilities are suitable for energy harvester modelling because an energy harvester is naturally a mixed-physical-domain system.

The technique outlined below models the complete system including the analogue mechanical, magnetic and electrical power storage and processing parts, the digital control of the

microgenerator tuning system, as well as the power consumption models of sensor node. Additionally, the paper proposes a response surface based design space exploration and optimisation technique so that not only the energy harvester design parameters but also the sensor node operation parameters can be optimised in order to achieve the best system performance.

II. PERFORMANCE OPTIMISATION

An automated energy harvester design flow must be implemented holistically and based on a single software platform that can be used to model, simulate, configure and optimise an entire energy harvester systems. Such a design flow is outlined in the pseudo-code of Algorithm 1 and also shown in Figure 2. Naturally, the process starts with initial design specification, such as available energy source (light, heat, vibration, etc), environmental energy density, device size, minimum voltage level/power output. According to these specifications, HDL models are constructed from component cells available in the component library. The component library contains parameterised models of different kind of micro-generator structures (solar cell, electromagnetic, piezoelectric, etc), various booster circuit topologies and storage elements. The outer loop in the algorithm represents this structure configuration process, which involves examining and comparing those HDL models from the library with the aim of identifying a set of components that meet specific user requirements. The inner design flow loop will then find the best performance of each candidate design by adjusting electrical and non-electrical parameters of the design's mixed-technology HDL model. The parametric optimisation of the generated structure will further improve the energy harvester efficiency by employing suitable optimisation algorithms. The design flow ends at the best performing design for fabrication subject to the user-defined performance characteristics.

Algorithm 1 Automated energy harvester design flow.
  Initial design structure and specification
  Structure configuration loop:
  for all design structures **do**
    Build HDL model of design
    Optimisation loop:
    repeat
      Simulate and evaluate performance
      if best performance not achieved **then**
        Update design parameters
      end if
    until best performance achieved
    if there are more structures to try **then**
      Select new structure
    end if
  end for

The requirements for energy harvester component models are: 1) models need to be computationally efficient for fast performance optimisation when used in complete energy-harvester systems and yet accurate; these are conflicting
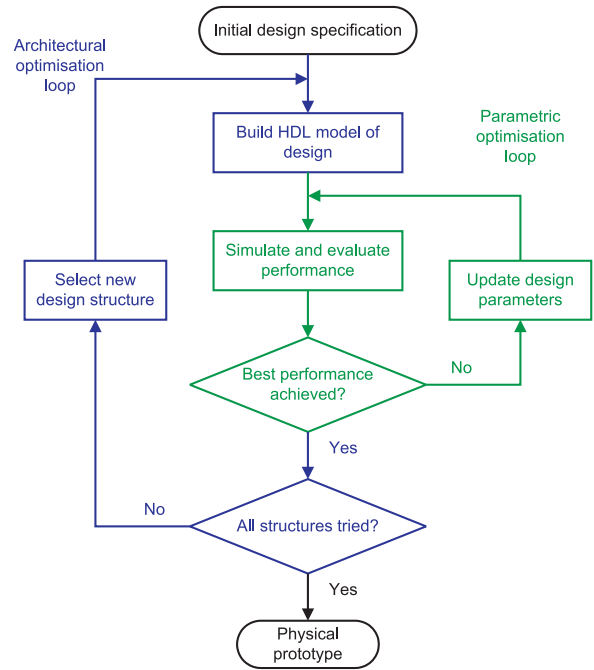
Fig. 2.   Energy harvester design flow.

requirements, 2) models need to capture both theoretical equations and practical non-idealities required for accurate performance estimation. The models should support different mechanical-electrical structures and will be expressed in terms of HDL descriptions. They will be able to predict the behaviour of the actual device accurately while remaining reconfigurable.

A small HDL model library of energy harvester components has been built. It contains two types of micro-generator, each of which can be configured with different coils (wire diameter of 12/16/25 μm), and two types of voltage multipliers that have three to six stages. The voltage transformer has not been included because it cannot be made and tested with available resources. But the simulation based optimisation of energy harvester with voltage transformer has been performed and will be discussed in Section II-A2. The configuration target has been set to find the set of components that can charge the 0.047F super capacitor to 2V in shortest time. These values were chosen because there has been reported energy harvester systems that use 0.047F storage capacitor and 2V working voltage [12].

Simulations of every available energy harvester configuration were carried out simultaneously and a process has been developed to automatically track the best model. SystemVision VHDL-AMS simulator [13] has been used as the single software platform. The outcome design is listed in Table I.

The micro-generator II has been chosen because it's bigger and stores more kinetic energy. However, it's quite interesting that the coil with the largest wire diameter, which leads to fewest number of turns, and the VM with fewest stages have been chosen. To further investigate on this result, more simu-

TABLE I
PARAMETERS OF THE CONFIGURATION RESULT.

Micro-generator	Type II
Wire diameter	25 μm
Voltage booster	3-stage Dickson voltage multipler

lations have been done and an important trade-off between the electromagnetic micro-generator and the VM voltage booster has been found.

Figure 3 shows the charging waveforms of Type I micro-generator connected to the same 5-stage VM but configured with different coils. At the beginning, the energy harvester with 25 μm wire diameter charges the quickest and the 12μm configuration charges the slowest while the 16μm one is in between. But the 25μm configuration also saturates quickly and reaches the 2V mark slower than the 16 μm energy harvester. Due to simulation time limitation, the figure does not show how the other two waveforms end. But it could be foreseen that the 16μm configuration will also saturate at some point while the 12μm one reaches highest voltage.
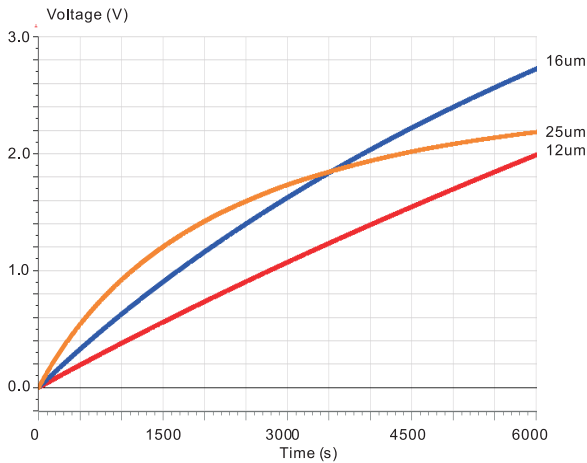
Fig. 3.   Simulation of Type I micro-generator with different coils.

Similar results have been obtained from the voltage booster end. Figure 4 shows the charging waveforms of Type II micro-generator with 25μm coil connecting with 3, 4 and 5 stages Dickson VMs. It can be seen that the energy harvester with 3-stage VM charges the super capacitor to 2V first and the one with 5-stage VM can reach the highest voltage.

From the simulation results it can be concluded that in an energy harvester design that combines electromagnetic micro-generator and voltage multiplier, the fewer number of turns in the coil and the fewer VM stages, the higher initial charging rate the energy harvester can get but the lower voltage it can finally reach. Therefore, although the micro-generator with more coil turns can generate more power and VMs with more stages can boost the voltage higher, under certain circumstances the optimisation of subsystems in isolation does not lead to a globally optimised design. It proves that when combining different components of an
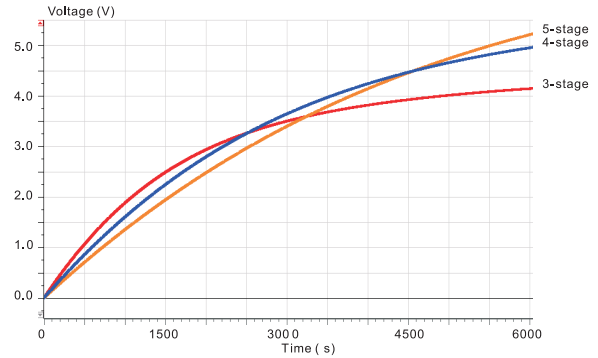
Fig. 4.   Simulation of Type II micro-generator with different VMs.

energy harvester, the gain at one part may come at the price of efficiency loss elsewhere, rending the energy harvester much less efficient than expected. This information is very useful for the development of future, more complicated systems and model libraries.

A. Performance optimisation

The close mechanical-electrical interaction (micro-generator and voltage booster) that takes place in energy harvesters, often lead to significant performance loss when the various parts of the energy harvesters are combined. Here the loss expressed in terms of energy harvesting efficiency:

$$\eta_{Loss} = \frac{E_{Harvested} - E_{Delivered}}{E_{Harvested}} \qquad (1)$$

In the proposed design flow, the generated energy harvester design should be parameterised such that automated performance optimisation will be able to further improve the energy harvester efficiency by employing suitable optimisation algorithms. The parameters used for the optimisation are from both the micro generator and the voltage booster. The optimisation object is to increase the charging rate of the super capacitor.

1) Exhaustive search: The micro-generator parameters that can be optimised are related to the coil size, i.e the thickness (t) and the outer radius (R). Because other components such as the magnets and cantilever determine the resonant frequency of the micro-generator and thus should be based on application requirements. The circuit parameters of voltage booster are the capacitor values of each VM stage. The entire energy harvester is optimised as an integrated model. The searching space of parameters has been given in Table II

TABLE II
OPTIMISATION SEARCHING SPACE.

Coil thickness(mm)	1.0-1.3
Coil radius(mm)	2.0-2.45
Capacitor values(uF)	47/100/150

The optimisation is based on the concurrent simulations of design instances from uniform sampling the search space and track the best result (Figure 5). This is relative simple and

straightforward because after the automatic structure configuration the search space is quite small and the VM capacitors can only have discrete values. However, other optimisation algorithms may also be employed and in Section II-A2 a VHDL-AMS based genetic optimisation has been successfully applied to the integrated optimisation of energy harvester systems.
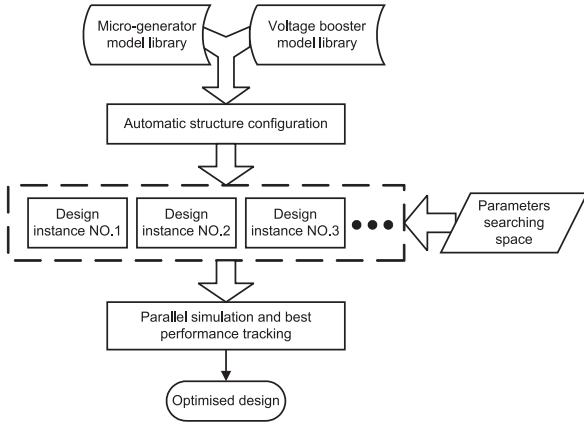
Fig. 5. Implementation of the proposed energy harvester design flow in VHDL-AMS.

To validate the effectiveness of the proposed approach to improve energy harvesting efficiency, the following simulations and experimental measurements have been carried out.

Original design: combines Type II micro-generator with a 5 stage Dickson VM. The used VM has been reported in literature as optimal configuration [14]. However, in the original design these two parts are optimised separately, which is quite common in existing energy harvester design approach. Parameters of original design are listed in Table III.

TABLE III
PARAMETERS OF ORIGINAL ENERGY HARVESTER.

Micro-generator	
Wire diameter(μm)	16
Coil thickness(mm)	1.3
Coil radius(mm)	2.45
Voltage booster	
VM configuration	5-stage Dickson
Capacitor values(C1-C5,uF)	47,150,150,47,150

Optimised design: has been obtained using the proposed design flow (Figure 5). Table IV gives the new micro-generator and voltage booster parameters.

TABLE IV
PARAMETERS OF OPTIMISED ENERGY HARVESTER.

Micro-generator	
Wire diameter(μm)	25
Coil thickness(mm)	1.3
Coil radius(mm)	2.0
Voltage booster	
VM configuration	3-stage Dickson
Capacitor values(C1-C3,uF)	100,100,47

The impact of these values on improving the energy harvester performance has been validated in both simulation and experimental measurements. According to the optimisation result, a new coil has been ordered from Recoil Ltd, UK [15] and replaced the original one for testing (see Figure 6).
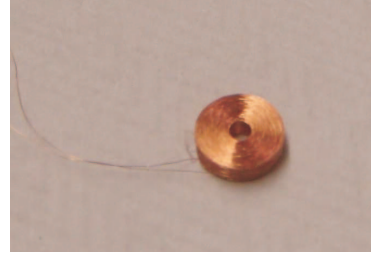
Fig. 6. New coil according to optimisation result (R=2.0mm, r=0.5mm, t=1.3mm, d=25μm).

Simulation and experimental waveforms of the original and optimised design are shown in Figure 7. As can be seen from the figure, there is good a correlation between the simulation and experimental waveforms in both of the energy harvester designs, which validates the effectiveness and accuracy of the proposed design flow. The energy harvester from original design can charge the super capacitor to 2V in 6000 seconds while the optimised design only uses 1500 seconds, which represents a 75% improvement.
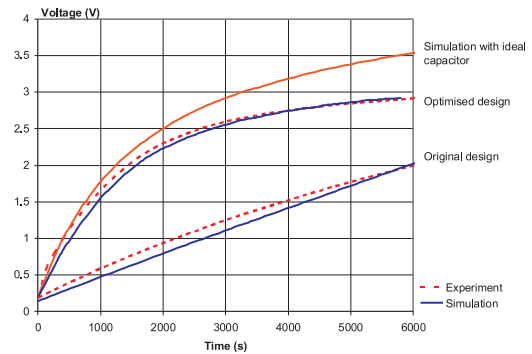
Fig. 7. Simulation and experimental waveforms of original and optimised energy harvesters.

2) Genetic optimisation: This section demonstrates another possible optimisation method to improve the energy harvester efficiency. Figure 8 shows that in the proposed approach, not only the energy harvester model but also the optimisation algorithm is implemented in a single VHDL-AMS testbench. The parameters used for the optimisation are from both the micro generator and the voltage booster. The optimisation object is to increase the charging rate of the super capacitor. The optimisation algorithm generates design parameters to the model and obtains the charging rate through simulation. The optimisation loop runs continuously until the design parameters reach an optimum.

A super capacitor of 0.22F has been used in the performance optimisation experiment. The micro-generator parameters that

4
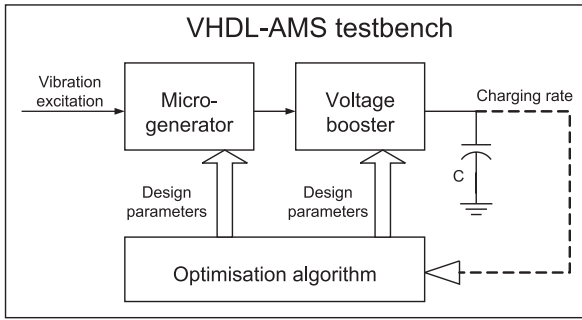
Fig. 8.   Integrated performance optimisation in VHDL-AMS testbench.

can be optimised are the number of coil turns (N), the internal resistance (R_c) and the outer radius (R). The voltage booster circuit here is a voltage transformer. The optimisation parameters are the number of turns and the resistance of the transformer's primary and secondary windings. For proof of concept, a genetic algorithm (GA) [16] has been employed to optimise the energy harvester with a voltage transformer booster. The implemented GA has a population size of 100 chromosomes. Each chromosome has 7 parameters (3 from the micro-generator and 4 from the voltage booster). The crossover and mutation rate are 0.8 and 0.02 respectively. Other optimisation algorithms may also be applied based on the proposed integrated model. The "un-optimised" model parameters are given in Table V.

TABLE V
PARAMETERS OF UN-OPTIMISED ENERGY HARVESTER.

Micro-generator		
Outer radius of coil (R)	1.2 mm	
Coil turns (N)	2300	
Internal resistance (R_c)	1600 Ω	
Voltage transformer		
	Resistance(Ω)	No. of turns
Primary winding	400	2000
Secondary winding	1000	5000

Applying the proposed modelling and performance optimisation, Table XIV gives the new micro-generator and voltage booster parameters which are referred to as "optimised" design. The impact of these values on improving the charging of the super is shown in Figure 9. As can be seen from the simulation results, in 150 minutes the un-optimised energy harvester charges the super capacitor to 1.5V and the optimised energy harvester reaches 1.95V, which represents a 30% improvement.

Performance of the developed GA has been further investigated by comparing the power transfer efficiency before and after optimisation. The maximum average power that can be delivered to the electrical domain is about 144μW. Table VII lists the average electrical power output from the micro generator and the voltage transformer. It can be seen that the optimisation improves the efficiency of both the micro generator and voltage booster, which validates the effectiveness of the developed genetic optimisation.

TABLE VI
PARAMETERS OF GA OPTIMISED ENERGY HARVESTER.

Micro-generator		
Outer radius of coil (R)	1.1 mm	
Coil turns (N)	2100	
Internal resistance (R_c)	1400 Ω	
Voltage transformer		
	Resistance(Ω)	No. of turns
Primary winding	340	1900
Secondary winding	690	3800
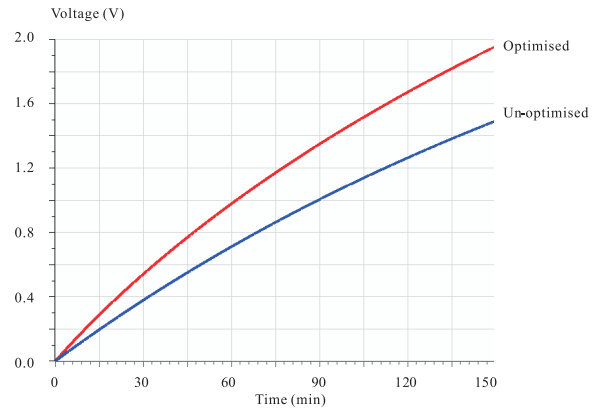
Fig. 9.   Simulation waveforms of super capacitor charging by different energy harvester models.

TABLE VII
ENERGY HARVESTER POWER EFFICIENCY.

	Generated power(μW)	Delivered power(μW)	Overall efficiency
Pre-optimisation	26.875	15.750	10.94%
Post-optimisation	29.250	19.625	13.63%

III. COMPLETE WIRELESS SENSOR NODE

Fig. 10 shows the diagram of the wireless sensor node system powered by tunable energy harvester. The wireless sensor node has a temperature sensor and a 2.4GHz radio transceiver. Once activated, the measured data is transmitted to another transceiver which is connected to a PC's USB port. The microgenerator converts the input vibration into electrical energy. The generated AC voltage is rectified by a diode bridge and stored in a 0.55F supercapacitor. The supercapacitor acts as the energy source for the microcontroller that controls the frequency tuning of the microgenerator and for the sensor node. In order to tune the resonant frequency of the microgenerator to match the frequency of the vibration source, the microcontroller uses two input signals, one from the microgenerator and one from the accelerometer. The operational amplifier acts as a comparator to generate square waves from the microgenerator output so that it is easy for the microcontroller to calculate the frequency. The detailed tuning algorithms are presented in Section III-A3. The microcontroller also provides energy for the accelerometer, the operational amplifier and the actuator so that these devices can be turned off when not in use. Table VIII lists the type and

make of the system components.
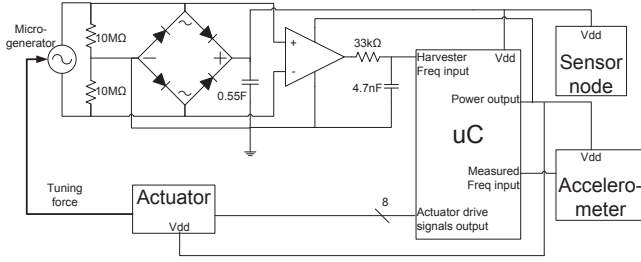
Fig. 10. System diagram of a tunable energy harvester powered wireless sensor node

TABLE VIII
SYSTEM COMPONENTS POWERED BY THE ENERGY HARVESTER

Component	Type	Make
Microcontroller	PIC16F884	Microchip
Accelerometer	LIS3L06AL	STMicroelectronics
Linear actuator	21000 Series Size 8 stepper motor	Haydon
Sensor node	eZ430-RF2500	Texas Instruments

A. System component models

1) Tunable microgenerator: Fig. 11(a) shows a diagram of the electromagnetic microgenerator together with its tuning mechanism. The microgenerator is based on a cantilever structure. The coil is fixed to the base, and four magnets (which are located on both sides of the coil) form the proof mass. The tuning mechanism uses magnetic force to change the effective stiffness of the cantilever which leads to a change of resonant frequency. One tuning magnet is attached to the end of the cantilever beam and the other tuning magnet is connected to a linear actuator. The linear actuator moves the magnet to the calculated desired position so that the resonant frequency of the microgenerator matches the frequency of the ambient vibration. The control algorithm is modelled as a SystemC digital process described in Section III-A3. Fig. 11(b) shows a photo of the microgenerator which is used to validate the proposed technique [17].

The dynamic model of the microgenerator is [18]:

$$m\frac{\mathrm{d}^2z(t)}{\mathrm{d}t^2} + c_p\frac{\mathrm{d}z(t)}{\mathrm{d}t} + k_s z(t) + F_{em} + F_{t\_z} = F_a \quad (2)$$

where m is the proof mass, $z(t)$ is the relative displacement between the mass and the base, c_p is the parasitic damping factor, k_s is the effective spring stiffness, F_{em} is the electromagnetic force, F_{t_z} is the z component of tuning force F_t and F_a is the input acceleration force. The z component of tuning force is:

$$F_{t\_z} = F_t\frac{z(t)}{l_c} \quad (3)$$

where l_c is the length of the cantilever.
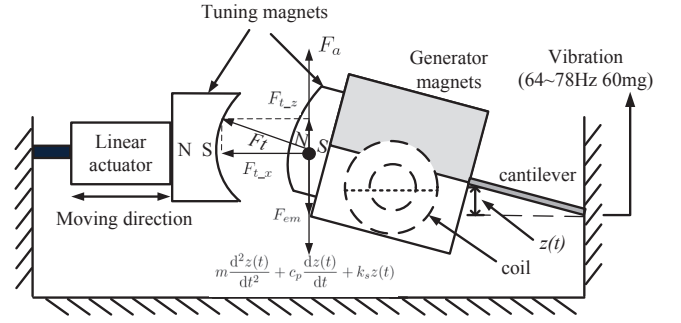
(a) Mechanical part
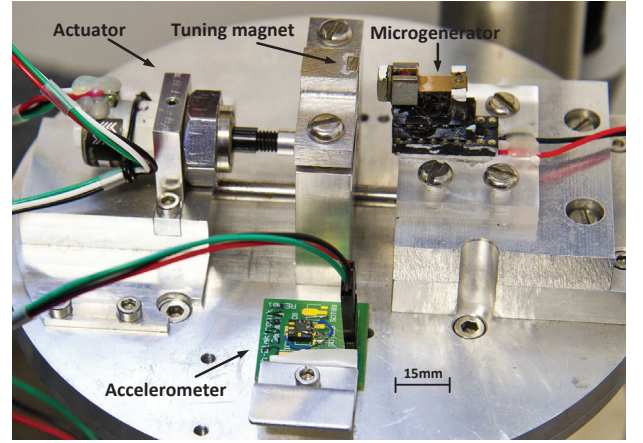
(b) Photo of tunable microgenerator

Fig. 11. Tunable electromagnetic microgenerator

The resonant frequency ω_0 and damping coefficient ζ are:

$$\omega_0 = \sqrt{\frac{k_s}{m}} \quad (4)$$

$$\zeta = \frac{c_p}{2\sqrt{mk_s}} \quad (5)$$

The resonant frequency of the tuned microgenerator (f'_r) is:

$$f'_r = f_r\sqrt{1 + \frac{F_t}{F_b}} \quad (6)$$

where f_r is the un-tuned resonant frequency, F_t is the tuning force between two magnets and F_b is the buckling load of the cantilever.

The electromagnetic voltage generated in the coil is:

$$V_{em} = -\Phi\frac{\mathrm{d}z(t)}{\mathrm{d}t} \quad (7)$$

where $\Phi = NBl$ is the transformation factor and N is the number of coil turns, B is the magnetic flux density and l is the effective length. The output voltage is:

$$V_m(t) = V_{em} - R_c i_c(t) - L_c\frac{\mathrm{d}i_L(t)}{\mathrm{d}t} \quad (8)$$

where R_c and L_c are the resistance and inductance of the coil respectively and $i_c(t)$ is the current through the coil. The

electromagnetic force is calculated as:

$$F_{em} = \Phi i_c(t) \tag{9}$$

2) Energy-aware sensor node behaviour and power consumption model: The eZ430-RF2500 wireless sensor node from Texas Instruments has been used in the system. The on-board controller is the MSP430F2274 and is paired with the CC2500 multi-channel RF transceiver, both of which are based on low-power design. The sensor node (Fig. 12) monitors the environment temperature as well as the supercapacitor voltage. Once activated, it transmits the temperature and voltage values through the radio link. Transmissions do not involve receiving acknowledgements. A program has been developed for the sensor control module to configure the sensor node in an energy-aware manner, namely that its transmission interval should depend on the available energy on the supercapacitor. The sensor node behaviour is summarised in Table IX. The transmission interval when the supercapacitor voltage is above 2.8V, i.e more energy stored, has been chosen as one parameter for optimisation. Although it is desirable to have as many transmissions as possible during a fixed time period, it may not always be the case that the transmission interval should be set as small as possible. This is because if the transmission is so frequency that the sensor node uses more energy than the harvester can generate, the supercapacitor voltage will drop below 2.8V and the transmission interval will increase in order for the energy storage to recover. Other factors such as frequency tuning also uses stored energy and therefore will affect how much energy is available for the sensor node.
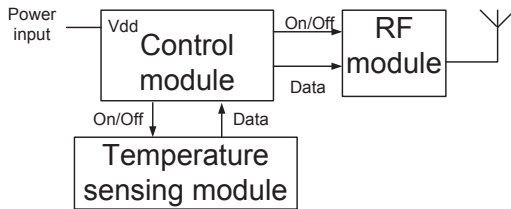
Fig. 12.   Block diagram of the sensor node

TABLE IX
SENSOR NODE BEHAVIOUR BASED ON SUPERCAPACITOR VOLTAGE

Supercapacitor voltage	Wireless transmission interval
Below 2.7V	No transmission
Between 2.7 and 2.8V	Every 1 minute
Above 2.8V	Every 5 seconds (parameter for optimisation)

In order to characterise the power consumption model of the sensor node, the current draw of the sensor node has been measured during each transmission. The results are listed in Table X.

The supply voltage was kept at 2.9V. So during each transmission lasting 4.5 ms, the sensor node consumes 227 μJ of energy and the equivalent resistance of its energy consumption model is:

$$R_{node} = \begin{cases} 167 \, \Omega & \text{when in transmission} \\ 5.8 \, \text{M}\Omega & \text{when in sleep} \end{cases} \tag{10}$$

TABLE X
CURRENT DRAW OF THE SENSOR NODE

Operation	Time	Current
Sleep mode	N/A	0.5μA
Wake-up	1 ms	4.5 mA
Sensing	1.5 ms	13.4 mA
Transmission	2 ms	26.8 mA

3) Tuning algorithms and power consumption models: In order for a energy harvester powered wireless sensor node (Fig. 1) to work autonomously, all the system components need to be powered by the harvested energy. The pseudo code of the tuning algorithm is shown in Algorithm 5. Standard SystemC modules were used to model the digital control process and in the experimental verification the control algorithm was implemented in a PIC16F884 microcontroller. As can be seen in Algorithm 5, a watchdog timer wakes the microcontroller periodically and the microcontroller first detects if there is enough energy stored in the supercapacitor. If there is not enough energy, the microcontroller goes back to sleep and waits for the watchdog timer again. If there is enough energy, the microcontroller will then compare the frequency of the microgenerator signal, which is close to the input vibration frequency, to the microgenerator's resonant frequency. When a difference is detected between the vibration frequency and the resonant frequency, the microcontroller retrieves the new desired position of the tuning magnet from a look-up table and begins a tuning process by controlling the actuator to move the tuning magnet to the new position (Fig. 11(a)). The watchdog timer and the microcontroller's clock frequency have been chosen as parameters for optimisation. Because these two parameters determine how much energy the microcontroller consumes and how quickly the system can response to the input vibration frequency change.

Algorithm 5 contains two subroutines: rough tuning (Algorithm 6) and fine tuning (Algorithm 7). The rough tuning measures the frequency of the microgenerator output and moves the actuator to the optimum position according to a predefined lookup table. However, the rough tuning alone cannot generate the best performance and a fine tuning algorithm is needed. This is because the measurement of the frequency of the microgenerator signal does not represent the input vibration frequency accurately enough and, in addition, there may also be a phase difference between the input vibration and the microgenerator motion that prevents the microgenerator from working at the resonance. The fine tuning takes another input, the raw vibration data from the accelerometer and moves the actuator to minimize the phase difference between microgenerator signal and the accelerometer signal so that the microgenerator is working as resonance. It can be seen that the fine tuning algorithm requires more calculation (thus more energy) than the rough tuning and additional energy is consumed by the accelerometer (see Table XI). Therefore it is not so energy efficient to use only the fine tuning algorithm as the proposed two-subroutine method. In the two-subroutine method, the rough tuning moves the actuator to

---

Algorithm 2 Harvester tuning control algorithm

1: **repeat**
2:     Energy generation while waiting for watchdog timer: 320 seconds (parameter for optimisation)
3:     **if** Enough energy stored in the supercapacitor ($V_s \geq 2.6$V, where 2.6V is the minimum voltage for the actuator to start) **then**
4:         Turn on Timer1 (clock frequency as parameter for optimisation)
5:         **repeat**
6:             Measure microgenerator period
7:         **until** 8 cycles have been measured
8:         Turn off Timer1
9:         Calculate input vibration frequency from 8 measurements
10:       Find optimum position (8-bit) of tuning magnet through look-up table which has been pre-obtained and stored in the microcontroller memory
11:       **if** Current position of tuning magnet matches optimum position (the accuracy is $1/2^8$) **then**
12:           Goto 2
13:         **else**
14:           Perform rough tuning (**Algorithm 2**)
15:         **end if**
16:       Measure the phase different between the accelerometer signal and the microgenerator signal
17:       **if** The phase difference is less than $100\mu s$ **then**
18:           Goto 2
19:         **else**
20:           Perform fine tuning (**Algorithm 3**)
21:         **end if**
22:     **end if**
23: **until** Forever

---

the approximate resonant position and the fine tuning finds the exact resonance.

Algorithm 3 Rough tuning algorithm

1: **repeat**
2:     Send the optimum position as 8-bit control signal to the actuator
3:     The actuator moves tuning magnet
4:     Wait 5 seconds for the microgenerator signal to settle down
5:     Compare the current position and optimum position
6: **until** Current position of tuning magnet matches optimum position

---

To tune the resonant frequency of the microgenerator effectively, the system incorporates a microcontroller, a linear actuator and an accelerometer. These three components need to be powered by the energy harvester in order to make an autonomous system. To characterise the power consumption models of these components, current measurements have been taken and power/energy consumptions have been calculated

---

Algorithm 4 Fine tuning algorithm

1: **repeat**
2:     Send the direction of movement that can reduce phase difference to the actuator
3:     The actuator moves tuning magnet by 1 step
4:     Wait 5 seconds for the microgenerator signal to settle down
5:     Measure the phase of the accelerometer signal
6:     Measure the phase of the microgenerator signal
7:     Calculate the phase difference
8: **until** The phase difference is less than $100\mu s$

(Table XI). According to the current and voltage values together with their operational times, the equivalent resistances for the power consumption models of these devices have been obtained.

TABLE XI
POWER CONSUMPTION MODELS OF THE SYSTEM COMPONENTS

Component (action)	Operation time(ms)	Current (mA)	Power (mW)	R_{eq} (Ω)
Accelerometer	153	5.1	13.2	509
Actuator				
(1 step)	5	312	811	8.33
(100 steps)	500	156	405	16.7
Microcontroller				
(Rough tuning)	149	1.9	5.0	1.38k
(Fine tuning)	325	5.1	6.5	250

IV. HDL IMPLEMENTATION

A. Analog part

The SystemC-A language [19] is used to build the system models. It is an extension to the SystemC language with analog and mixed-signal (AMS) capabilities. The digital part is modeled using standard SystemC modules. The analog part, consisting of non-linear differential and algebraic equations, is included using the extended syntax where the user defines the behavior of each analog component by specifying the *build* methods that contribute to the analog equation set of whole system. In Systemc-A, a *build* method is provided to support the automatic equation formulation of the user-defined system models. It is a virtual method in the abstract *component* base class and inherited by all derived components. It consists of two functions, $BuildM()$ and $BuildRhs()$. SystemC-A uses the $BuildM()$ method to add the Jacobian entries to the analog equation set and $BuildRhs()$ method to build the equations, i.e. the right hand side of the Newton-Raphson linearized equation set. The microgenerator equations and corresponding Jacobian matrix entries to be included in the SystemC-A model are listed in Table XII.

The SystemC-A code of the tunable microgenerator model, which is according to Table XII, is listed below:

```
generator::generator(){} //constructor
generator::generator(char nameC[5],TerminalVariable
*node_a,TerminalVariable *node_b,double value,double
```

TABLE XII
EQUATION FORMULATION OF THE MICROGENERATOR MODEL

	$z(t)$	$\frac{dz(t)}{dt}$	$i_L(t)$	Equation
$z(t)$	$-k_s$	$-c_p - mS$	$-\Phi$	$m\frac{d^2z(t)}{dt^2} + c_p\frac{dz(t)}{dt}$ $+k_sz(t) + \Phi i_L(t) + F_{t\_z} - F_a$
$\frac{dz(t)}{dt}$	S	-1	0	0
$i_L(t)$	0	$-\Phi$	R_c	$-R_c i_L(t) - L_c\frac{di_L(t)}{dt} + \Phi\frac{dz(t)}{dt}$

```
Freq): //node_a is Vm, node_b is Im, value is the
tuning force, Freq is the input frequency
component(nameC,node_a,node_b,value){
ztQ = new Quantity("ztQ");
//quantity zt is relative displacement
ytQ = new Quantity("ytQ");
//quantity yt is velocity
itQ = new Quantity("itQ");
//quantity it is inductor current
Fin=value; //tuning force
omega=Freq*2*3.14159;}

void generator::build(){ //model equations

t=TS->get_time(); //current time point
S=TS->get_S();
//time derivative, S=2/h for trapezoidal integration
mpytdotdot=-Mp*Yam*omega*omega*sin(omega*t);
//input acceleration force

zt=X(ztQ);
yt=X(ytQ); //X() return previous value
it=X(itQ);

ztdot=Xdot(ztQ); //Xdot() return previous time
derivative
ytdot=Xdot(ytQ);
itdot=Xdot(itQ);

BuildM(ztQ,ztQ,-Ks); //Jacobian of equation (2)
BuildM(ztQ,ytQ,-Cp-Mp*S);
BuildM(ztQ,itQ,-Phi);
BuildRhs(ztQ,mpytdotdot+Mp*ytdot+Cp*yt+Ks*zt+Phi*it);
//Right hand side of equation (2)

BuildM(ytQ,ztQ,S);
BuildM(ytQ,ytQ,-1);
BuildM(ytQ,itQ,0);
BuildRhs(ytQ,yt-ztdot);

BuildM(itQ,ztQ,0); //Jacobian of equation (8)
BuildM(itQ,ytQ,-Phi);
BuildM(itQ,itQ,Rc);
BuildRhs(itQ,-Rc*it-Lc*itdot-vt+Phi*yt);
//Right hand side of equation (8)
}
```

B. Digital part

The pseudo code of the tuning algorithm is shown in Algorithm 5. Standard SystemC modules were used to model the digital control process and in the experimental verification the control algorithm was implemented in a PIC16F884 microcontroller. As can be seen in Algorithm 5, a watchdog timer wakes the microcontroller periodically and the microcontroller first detects if there is enough energy stored in the supercapacitor. If there is not enough energy, the microcontroller goes back to sleep and waits for the watchdog timer again. If there is enough energy, the microcontroller will then compare the frequency of the microgenerator signal, which is close to the input vibration frequency, to the microgenerator's resonant frequency. When

a difference is detected between the vibration frequency and the resonant frequency, the microcontroller retrieves the new desired position of the tuning magnet from a look-up table and begins a tuning process by controlling the actuator to move the tuning magnet to the new position (Fig. 11(a)).

Algorithm 5 Harvester tuning control algorithm

1: **repeat**
2:    Energy generation while waiting for watchdog timer (320 seconds)
3:    **if** Enough energy stored in the supercapacitor ($V_s \geq 2.6$V, where 2.6V is the minimum voltage for the actuator to start) **then**
4:      Turn on Timer1
5:      **repeat**
6:        Measure microgenerator period
7:      **until** 8 cycles have been measured
8:      Turn off Timer1
9:      Calculate input vibration frequency from 8 measurements
10:      Find optimum position (8-bit) of tuning magnet through look-up table which has been pre-obtained and stored in the microcontroller memory
11:      **if** Current position of tuning magnet matches optimum position (the accuracy is $1/2^8$) **then**
12:        Goto 2
13:      **else**
14:        Perform rough tuning (**Algorithm 2**)
15:      **end if**
16:      Measure the phase different between the accelerometer signal and the microgenerator signal
17:      **if** The phase difference is less than 100μs **then**
18:        Goto 2
19:      **else**
20:        Perform fine tuning (**Algorithm 3**)
21:      **end if**
22:    **end if**
23: **until** Forever

Algorithm 5 contains two subroutines: rough tuning (Algorithm 6) and fine tuning (Algorithm 7). The rough tuning measures the frequency of the microgenerator output and moves the actuator to the optimum position according to a predefined lookup table. However, the rough tuning alone cannot generate the best performance and a fine tuning algorithm is needed. This is because the measurement of the frequency of the microgenerator signal does not represent the input vibration frequency accurately enough and, in addition, there may also be a phase difference between the input vibration and the microgenerator motion that prevents the microgenerator from working at the resonance. The fine tuning takes another input, the raw vibration data from the accelerometer and moves the actuator to minimize the phase difference between the microgenerator signal and the accelerometer signal so that the microgenerator is working as resonance. It can be seen that the fine tuning algorithm requires more calculation (thus

more energy) than the rough tuning and additional energy is consumed by the accelerometer (see Table XI). Therefore it is not so energy efficient to use only the fine tuning algorithm as the proposed two-subroutine method. In the two-subroutine method, the rough tuning moves the actuator to the approximate resonant position and the fine tuning finds the exact resonance.

Algorithm 6 Rough tuning algorithm

1: **repeat**
2: Send the optimum position as 8-bit control signal to the actuator
3: The actuator moves tuning magnet
4: Wait 5 seconds for the microgenerator signal to settle down
5: Compare the current position and optimum position
6: **until** Current position of tuning magnet matches optimum position

---

Algorithm 7 Fine tuning algorithm

1: **repeat**
2: Send the direction of movement that can reduce phase difference to the actuator
3: The actuator moves tuning magnet by 1 step
4: Wait 5 seconds for the microgenerator signal to settle down
5: Measure the phase of the accelerometer signal
6: Measure the phase of the microgenerator signal
7: Calculate the phase difference
8: **until** The phase difference is less than $100\mu s$

V. SIMULATION RESULTS AND EXPERIMENTAL VERIFICATION

A SystemC-A model of the complete system has been built and simulated. The SystemC-A code of the top-level testbench is listed below. The system components include the microgenerator, the diode bridge, the supercapacitor and the equivalent variable resistances of the actuator, the accelerometer, the microcontroller and the sensor node.

```
void testbench::system(){

ACT=new actuator;
ACM=new accelerometer;
uC=new control;
NODE=new sensor;
n0 = new Node("0");//don't write n0
n1 = new Node("n1");
n2 = new Node("n2");
n3 = new Node("n3");
n4 = new Node("n4");
n5 = new Node("n5");
n6 = new Node("n6");
//microgenerator generator *G1 =new
generator("G1",n1,n2,0.3192,64);
//diode bridge
diode *D1 =new diode("D1",n0,n1,2.117e-7,1.015);
diode *D2 =new diode("D2",n0,n2,2.117e-7,1.015);
diode *D3 =new diode("D3",n2,n3,2.117e-7,1.015);
```

```
diode *D4 =new diode("D4",n1,n3,2.117e-7,1.015);
resistor *R1 =new resistor("R1",n1,n0,10e6);
resistor *R2 =new resistor("R2",n2,n0,10e6);
//super capacitor model
resistor *Ri =new resistor("Ri",n3,n4,0.204);
resistor *Rd =new resistor("Rd",n3,n5,84.0);
resistor *Rl =new resistor("Rl",n3,n6,4375.0);
cap_ini *Ci0 =new cap_ini("Ci0",n4,n0,0.35,1.65);
cap_vary *Ci1 =new cap_vary("Ci1",n4,n0,0.21,1.65);
cap_ini *Cd =new cap_ini("Cd",n5,n0,0.21,1.65);
cap_ini *Cl =new cap_ini("Cl",n6,n0,0.06,1.65);
//power consumption models for actuator,
accelerometer, microcontroller and sensor node
res_vary *RAct =new res_vary("RAct",n3,n0,1.0e9);
res_vary *RAcc =new res_vary("RAcc",n3,n0,1.0e9);
res_vary *RuC =new res_vary("RuC",n3,n0,1.0e9);
res_vary *RNode =new res_vary("RNode",n3,n0,1.0e9);
}
```

The test scenario has been divided into two parts. During the first half of the test, the input vibration frequency changes by 5Hz every 25 minutes (Fig. 13(a)). The main objective of the this part of the test is to demonstrate the frequency tuning capability of the microgenerator. It can be seen that after the input frequency changes, the supercapacitor voltage drops because the generated voltage is not high enough to charge the supercapacitor. Then the microcontroller wakes up and tunes the resonant frequency of the microgenerator, which uses much of the energy stored on supercapacitor but the retuned microgenerator starts to charge the supercapacitor again. During the second half, the input frequency is fixed and the performance of the sensor node is being tested (Fig. 13(b)). The sensor node transmits at different time intervals according to the different voltage levels on the supercapacitor (Table IX). The transmission interval is reflected on the supercapacitor charging slope. The shorter transmission interval is, the more gradual charging slope gets. Experimental measurements have been carried out and the waveforms are also shown in Fig. 13. The comparison between the simulation and experimental waveforms of the supercapacitor voltage represents both the energy generation and consumption of the system. In both figures the simulation results correlate well with the experimental measurements which validates the presented technique.

VI. FAST DESIGN EXPLORATION USING A RESPONSE SURFACE MODEL

Response surface models are constructed from a data set extracted from either physical experiments or computer experiments (simulations) [20]. Due to space limitations, only two major steps of the methodology are given below, namely the formation of an approximated mathematical model by fitting the response under study in terms of design parameters using regression analysis (Section VI-A) and the design of a series of experiments or simulations based on design of experiments (DOE) methodology (Section VI-B). Discussions of the statistical assessment of the goodness of fit and the fitted model reliability are omitted in this paper.

A. Response Surface Mathematical Model

Suppose there is a dependant variable(s) ($y \in R^n$) where n is the number of observations, believed to be affected
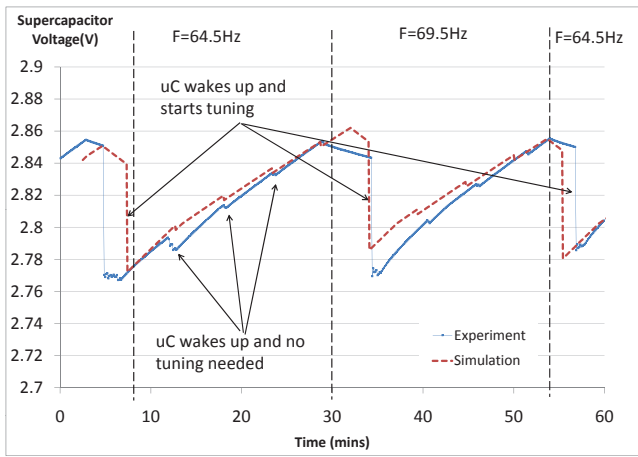
(a) Frequency tuning
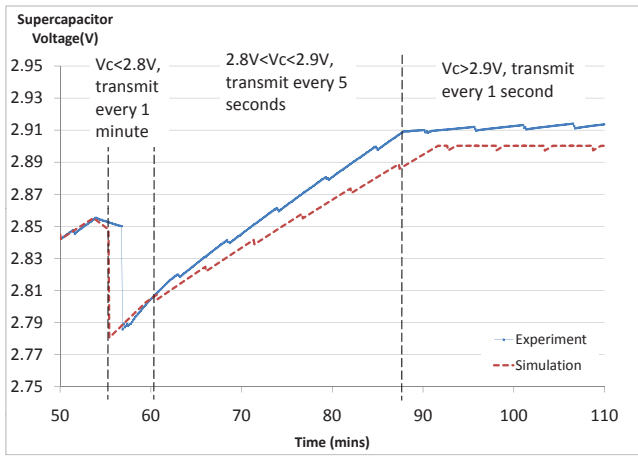
(b) Node behavior

Fig. 13. Simulations and experimental measurements of the supercapacitor voltages

by a vector of independent variables ($a \in R^k$) where k is the number of independent variables, then the relationship between the dependent variable(s) and independent variables can be expressed as:

$$y = f(a_1, a_2, ..., a_k) + \epsilon \qquad (11)$$

where ϵ represents the model errors, $a_1, a_2, ..., a_k$ are independent variables and $f()$ is called system function that relates dependant variable to independent variables. In most cases, the exact behaviour of the system function is unknown especially in engineering problems, so the system function $f()$ may be approximated by an empirical model as:

$$y = \hat{y}(a_1, a_2, ..., a_k) + \epsilon \qquad (12)$$

where $\hat{y}$ are a low order polynomials or a multi-dimensional splines, and this is called the response surface model. The independent variables or design parameters in equation (12) (i.e $a_1, a_2, ..., a_k$) are expressed in their corresponding physical units and must be converted to a dimensionless quantities with zero mean and the same standard deviation before proceeding

with further RSM analysis such as regression. These new quantities are called coded variables (i.e $x_1, x_2, ..., x_k$) of original design variables (parameters). The transformation process between natural representations and coded representations is achieved via equation (13):

$$x = \frac{a - [a_{max} + a_{min}/2}{[a_{max} + a_{min}/2} \qquad (13)$$

where a_{max} and a_{min} are the maximum and minimum value in the range of that specific design parameter. Now the approximated function $\hat{y}$ is expressed in term of coded variables ($x_1, x_2, ..., x_k$) and how to choose such a model $\hat{y}$ determines the success of applying RSM methodology. Typically, most engineering problems $\hat{y}$ can be approximated by a quadratic multi-variable polynomials as follows:

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j} \beta_{ij} x_i x_j \qquad (14)$$

where $\beta_0, \beta_i, \beta_i i, \beta_i j$ are the coefficients of the intercept, linear, quadratic and interaction in the regression model respectively, x_i, x_j are the design parameters in their coded format. The coefficients of the polynomial in equation (14) are determined through n simulation runs for the SystemC-A energy harvester model. The design points of the n runs are determined using DOE technique based on D-Optimal criteria. Using matrix notation, equation (14) can be written as:

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\beta \qquad (15)$$

where $\boldsymbol{X}_{n \times p}$ is $n \times p$ design matrix, p is the number of coefficients in the approximated polynomial, n is the number of simulation runs. $\beta_{p \times 1}$ are the unknowns parameters need to be solved. The difference between the observed values $\boldsymbol{y}$ and fitted values $\hat{\boldsymbol{y}}$ for the ith observation $\epsilon_i = y_i - \hat{y}_i$ is called the residual for that specific observation. The sum of the squares of the residuals (SSE) is defined as:

$$SSE = \sum_{i=1}^{n} \epsilon^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (16)$$

Combining equations (15 and 16) and differentiating with respect to β lead to:

$$\frac{\partial SSE}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{\partial}{\partial \beta_i} (y_i - \boldsymbol{X}\beta)^2 \right) \qquad (17)$$

Solving equation (17) for each β_i using least square method (LSM) will lead to an accurate model $\hat{y}$ that satisfy the condition of minimum residuals (i.e best fit).

B. D-Optimal Experimental Design

In the design matrix $\boldsymbol{X}_{n \times p}$, each specific run is represented by a single row and each column contains a specific design parameter that varies in each row based on predefined designed points. How to choose the predefined design points efficiently is called design of experiments (DOE) methodology. There are different types of design of experiments, such as full factorial, central composite design (CCD), Box Behnken

11

designs (BBD) and computer generated designs, such as D-optimal design [20]. Because D-optimal DOE explores design parameters space efficiently with minimum number of run that enable model construction with good accuracy [21], it has be used for the study in this paper. The algorithm of D-optimal criterion optimise the feasible potential design points to form a subset of D-optimal points that will be used in simulation runs. This optimisation is based on maximizing the determinant of XX', where XX' is called information matrix [21].

C. RSM optimisation results

As described in Section III-A, three parameters which affect the energy generation and consumption of the wireless sensor node system have been chosen for optimisation. Their value ranges and coded variable symbols are listed in Table XIII. Each of the three coded variables has three values [-1 0 1] which is the minimum number required to generate a quadratic approximation [20]. The full factorial design requires 27 (3^3) simulations while the D-optimal design only requires 10 simulations. As explained in Section VI-B, the D-optimal design points are obtained and 10 simulations have been carried out with the corresponding parameters. The acceleration level of the input vibration is fixed as 60mg and the input frequency changes by 5Hz every 25 minutes. The optimisation aim has been chosen as to maximise the number of transmissions during one hour. The MATLAB response surface toolbox has been used to generate the quadratic equation and the response surface model is:

$$\hat{y}(x_1, x_2, x_3) = 469.167 - 108.833x_1 - 18.833x_2 - 209.5x_3$$
$$+ 71.833x_1^2 + 90.5x_2^2 - 39.0x_3^2$$
$$- 32.333x_1x_2 - 71.333x_1x_3 + 43.333x_2x_3 \qquad (18)$$

TABLE XIII
SYSTEM PARAMETERS FOR OPTIMISATION

Description	Value range	Coded symbol
Microcontroller clock frequency(Hz)	125k - 8M	x_1
Watchdog timer wakeup time(sec)	160 - 480	x_2
Transmission time interval(sec)	1 - 10	x_3

The fitted model in equation (18) reflects the effects of each design parameters as well as the interactions effects between design parameters. Fig. 14 plots each single design parameter against the total number of transmissions while holding other design parameters constant.

Two algorithms from the MATLAB optimisation toolbox have been used to maximise the number of transmission, i.e maximise equation (18). The chosen algorithms are Simulated Annealing and Genetic Algorithm, both of which are capable of global searching. The optimisation results, together with the original design, are listed in Table XIV. It can be seen that both of the optimised design improved the system performance massively. The total number of transmissions doubled with the optimised design, which validates our proposed technique.
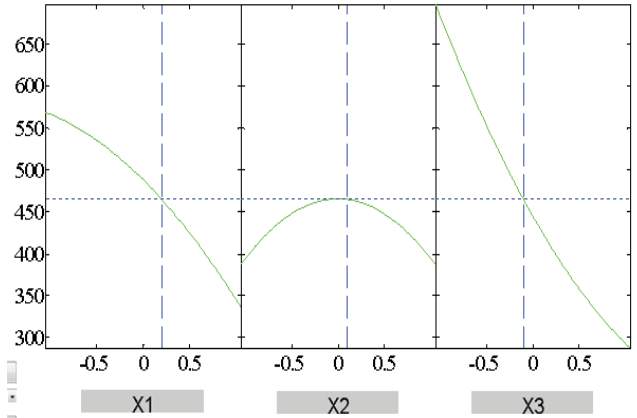
Fig. 14. The effect of each design parameter on system performance (total number of transmissions during one hour)

TABLE XIV
OPTIMISATION RESULTS

	Original design	Simulated Annealing	Genetic Algorithm
Microcontroller clock frequency(Hz)	4M	125k	125k
Watchdog timer wakeup time(sec)	320	160	480
Transmission time interval(sec)	5	1	1
Number of transmissions	405	869	809

VII. CONCLUSION

Wireless sensor networks are fast developing and energy harvester powered sensor nodes have attracted great research interest. In order to design energy efficient wireless sensor nodes, it is crucial to consider all the components in the context of energy consumption in a complete, autonomous wireless system. This paper presents such an HDL based modeling approach that links the system's energy generation and consumption with its analog parts as well as digital processes. Simulation and optimisation results of the developed HDL models match well with the experimental measurements and correctly reflect the changing energy flow when the digital processes are carrying out different operations. Future work will focus on the optimisation of both the energy harvester and digital control algorithms so that the system's overall energy efficiency can be improved. This paper also presents an approach to fast design space exploration based on a response surface model. The RSM has been used to optimise a complete wireless sensor node syste using SystemC-A and MATLAB. SystemC-A has been used to model the system's analogue components as well as the digital processes and MATLAB to generate and optimise the response surface model. As demonstrated by the optimisation results, the proposed technique leads to an efficient optimisation process by combining the power of SystemC-A in modelling multi-domain systems and the power of MATLAB in computation.

REFERENCES

[1] C. Alippi, R. Camplani, C. Galperti, and M. Roveri, "A robust, adaptive, solar-powered wsn framework for aquatic environmen-

tal monitoring," *Sensors Journal, IEEE*, vol. 11, no. 1, pp. 45–55, 2011.

[2] Q. Ling, Z. Tian, Y. Yin, and Y. Li, "Localized structural health monitoring using energy-efficient wireless sensor networks," *Sensors Journal, IEEE*, vol. 9, no. 11, pp. 1596–1604, 2009.

[3] A. Sapio and G. Tsouri, "Low-power body sensor network for wireless ecg based on relaying of creeping waves at 2.4ghz," in *Body Sensor Networks (BSN), 2010 International Conference on*, 2010, pp. 167–173.

[4] S. Roundy, P. K. Wright, and J. M. Rabaey, *Energy scavenging for wireless sensor networks: with special focus on vibrations*. Springer, 2004.

[5] M. P. Buric, G. Kusic, W. Clark, and T. Johnson, "Piezo-electric energy harvesting for wireless sensor networks," in *Wireless and Microwave Technology Conference, 2006. WAMICON '06. IEEE Annual*, 2006, pp. 1–5.

[6] S. Ergen, A. Sangiovanni-Vincentelli, X. Sun, R. Tebano, S. Alalusi, G. Audisio, and M. Sabatini, "The tire as an intelligent sensor," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 28, no. 7, pp. 941–955, 2009.

[7] P. Mitcheson, T. Green, E. Yeatman, and A. Holmes, "Architectures for vibration-driven micropower generators," *Journal of Microelectromechanical Systems*, vol. 13, no. 3, pp. 429–440, 2004.

[8] D. Zhu, J. Tudor, and S. Beeby, "Strategies for increasing the operating frequency range of vibration energy harvesters: a review," *Measurement Science and Technology*, vol. 21, no. 2, 2010.

[9] Energy Harvesting Systems: A Block Diagram (2010, July 16), "Holistic energy harvesting," September, 2011. [Online]. Available: http://www.holistic.ecs.soton.ac.uk/res/eh-system.php

[10] H. Boussetta, M. Marzencki, S. Basrour, and A. Soudani, "Efficient physical modeling of mems energy harvesting devices with vhdl-ams," *Sensors Journal, IEEE*, vol. 10, no. 9, pp. 1427–1437, 2010.

[11] L. Wang, T. Kazmierski, B. Al-Hashimi, A. Weddell, G. Merrett, and I. Ayala-Garcia, "Accelerated simulation of tunable vibration energy harvesting systems using a linearised state-space technique," in *Design, Test and Automation in Europe (DATE 2011)*, March 14-18, 2011, pp. 1267–1272.

[12] R. Torah, P. Glynne-Jones, J. Tudor, T. O'Donnell, S. Roy, and S. Beeby, "Self-powered autonomous wireless sensor node using vibration energy harvesting," *Measurement Science and Technology*, vol. 19, no. 12, pp. ISSN 1361–6501, 2008.

[13] M. G. Corporation, *SystemVision User's Manual*, ser. Version 3.2, Release 2004.3, July 2004.

[14] R. Torah, M. Tudor, K. Patel, I. Garcia, and S. Beeby, "Autonomous low power microsystem powered by vibration energy harvesting," *Sensors, IEEE*, pp. 264–267, 28-31 Oct. 2007.

[15] Recoil Ltd, UK, *http://www.recoilltd.com/index.htm*, Sept. 2008.

[16] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, Massachusetts: the MIT Press, 1996.

[17] I. A. Garcia, D. Zhu, J. Tudor, and S. Beeby, "Autonomous tunable energy harvester," in *PowerMEMS 2009*, 1-4 December 2009, pp. 49–52.

[18] D. Zhu, S. Roberts, J. Tudor, and S. Beeby, "Design and experimental characterization of a tunable vibration-based electro-magnetic micro-generator," *Sensors and Actuators A: Physical*, vol. 158, no. 2, pp. 284–293, 2010.

[19] H. Al-Junaid and T. Kazmierski, "Analogue and mixed-signal extension to SystemC," *IEE proc. Circuit Devices Systems*, vol. 152, no. 6, pp. 682–690, Dec. 2005.

[20] J. Jacquez, "Design of experiments," *Journal of the Franklin Institute*, vol. 335, no. 2, pp. 259–279, 1998.

[21] R. Unal, R. Lepsch, and M. McMillin, "Response surface model building and multidisciplinary optimisation using d-optimal designs," in *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on multidisciplinary Analysis and optimisation*, 1998, pp. 405–411.

A Simulation Study of Experimental GaInP/InGaAs/Ge Triple-Junction Solar Cell

Veljko Nikolić, Nebojša Janković

Abstract – In this paper, an experimental GaInP/InGaAs/Ge triple-junction (TJ) solar cell is described and fully simulated using the TCAD tool ATLAS/Silvaco. The major stages of the simulation process are explained and the simulation results are compared with experimental data. The simulated and measured results agree well with respect to I-V cell characteristics under the one sun irradiation. Also, low quantum efficiency is obtained from full cell simulation, but requires further cell design optimization.

I. INTRODUCTION

Multijunction solar cells made of III-V compound semiconductors are the most efficient devices for converting solar radiation into electrical energy. Today, these solar cells are widely used for powering satellites in space, and are started to be used for terrestrial applications through the use of photovoltaic concentrator systems.

Efficiencies have been dramatically improved in recent years. In 2010, Spire Semiconductor LLC produced a world record efficiency concentrator photovoltaic solar cell. The triple-junction solar sell achieved 42.3% conversion efficiency under a solar concentration of 406 suns [1].

Single solar cells can absorb only a small portion of the solar spectrum. To achieve a higher power density, solar cells with different I-V curves and bandgaps are stacked in order. The top cell has the highest bandgap, while the bottom cell has the lowest. This way, a cell absorbs the photons with energy higher than the bandgap and produces electric power. At the same time, it allows the lower–energy photons to pass through it. This method is called spectrum separation. Unfortunately, stacking the cells this way creates parasitic junctions between dissimilar regions of different cells reducing the current flow. To solve this problem tunnel junctions are used for separating individual cells. Tunnel junctions allow the photons to pass through as well as current with minimal voltage loss [2].

In this work, an experimental GaInP/InGaAs/Ge triple-junction (TJ) solar cell is described and fully simulated using the Technology Computer Aided Design (TCAD) tool ATLAS/Silvaco. In order to provide an accurate simulation, specific parameters needed to be inserted into the material related models. The exotic

Nebojša Janković is with the Department of Microelectronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia,
E-mail: nebojsa.jankovic@elfak.ni.ac.rs

materials used in this device were carefully studied and all their major electrical and optical parameters were researched or derived. Three individual cells were first developed and tested individually. This way, their efficiency was optimized and their performance characteristics were adjusted to the required range. Material and model parameters were optimized to reach a better approximation. This process was repeated several times until the results were accurate enough. Short circuit currents and open circuit voltages were obtained. J-V curves and frequency responses are also plotted. An identical process was done with the tunnel junctions. Finally, the complete triple-junction structure was simulated.

II. TJ CELL STRUCTURE

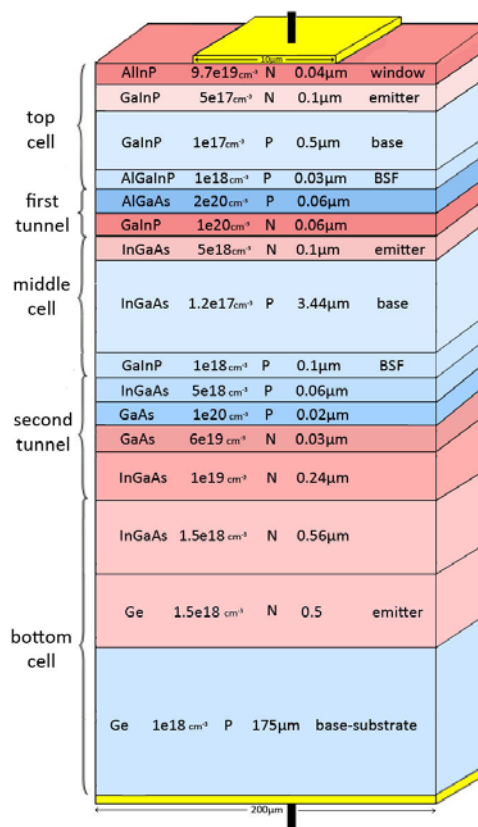The Simulated structure of fabricated experimental TJ collar cell [3] is schematically shown in Fig. 1.

Fig. 1. Triple junction solar cell

In the cell, the Ge substrate acts as a base for the bottom cell. On top of it are the Ge emitter and two layers of InGaAs which form a heterostructure. Next is the second tunnel junction consisting of two heavily doped n- and p-type GaAs layers. A thin InGaAs window layer comes after, followed by a GaInP BSF layer. InGaAs middle cell is stacked next, on top of the cell is the first tunnel junction formed by heavily doped layers of n-type GaInP and p-type AlGaAs. Then come AlGaInP BSF layer, the top cell from GaInP and an AlInP window. All contacts are made of gold.

A window layer is used in order to reduce the surface recombination velocity. Similarly, a BSF layer reduces the scattering of carriers towards the tunnel junction. Both layers have a heterojunction structure and are used for lattice matching. Also, they must be transparent to wavelengths absorbed by the next cell.

III. DEVICE SIMULATION

Silvaco TCAD software package is a large suite of highly sophisticated tools for design and development of all types of semiconductor devices. It contains ATLAS, a physically based device simulator which calculates electrical characteristics of the device's physical design and bias conditions [4].

The modules required for Solar Cell simulation include: S-Pisces, Blaze, Luminous, Device3D, Luminous3D. S-Pisces is an advanced 2D device simulator for silicon based technologies that incorporates both drift-diffusion and energy balance transport equations. Blaze simulates 2D solar cell devices fabricated using advanced materials. It includes a library of binary, ternary and quaternary semiconductors. Device3D simulates DC, AC and time domain characteristics for silicon and other material based technologies. Luminous and Luminous3D are advanced simulators specially designed to model light absorption and photogeneration in non-planar solar cell devices. Exact solutions for general optical sources are obtained using geometric ray tracing. This feature enables Luminous and Luminous3D to account for arbitrary topologies, internal and external reflections and refractions, polarization dependencies and dispersion. It also allows optical transfer matrix method analysis for coherence effects in layered devices. The beam propagation method may be used to simulate coherence effects and diffraction. [4]

When interference effects are important, such as in the presence of an antireflective coating or when the semiconductor absorption layers are thin, using traditional ray-tracing is useless because reflected rays in each layer need to be set to a very large number and that would exponentially increase the time of the simulation. The only way to achieve a good accuracy is to use the transfer matrix method. This approach relates total tangential components of the electric and magnetic fields at the multilayer boundaries. The structure of a multilayer completely determines the characteristic matrix of this multilayer. The transfer matrix also contains information about the media on both sides of the multilayer. [5]

There are five groups of physical models in ATLAS: mobility, recombination, carrier statistics, impact ionization, and tunnelling. They can be specified for each material. In Atlas simulations, we have included the Shockley-Read-Hall generation/recombination model, bandgap narrowing, Fermi-Dirac carrier statistics, and the non-local band to band tunnelling model.

Across highly doped junctions, the band-to-band tunnelling current can be very high and it depends very strongly on the band edge profile throughout the junction. In turn, the injected tunnel current creates a charge dipole and strongly affects the potential, and consequently the band energies, at and near the junction. This strong coupling can cause convergence problems in some situations. Non-local coupling and band-to-band forward models are used to mitigate these problems. Band to band optical generation/recombination (radiative) model is important for narrow gap semiconductors and semiconductors whose specific band structure allows direct transitions. [4]

The mesh is very important for the final accuracy of the simulation. The density needs to be high near regions such as junctions, material boundaries, or electrodes. If not, the results could be inaccurate and misleading. On the other hand, creating a thicker mesh increases simulation time. One part of the mash structure is shown in Fig. 2.
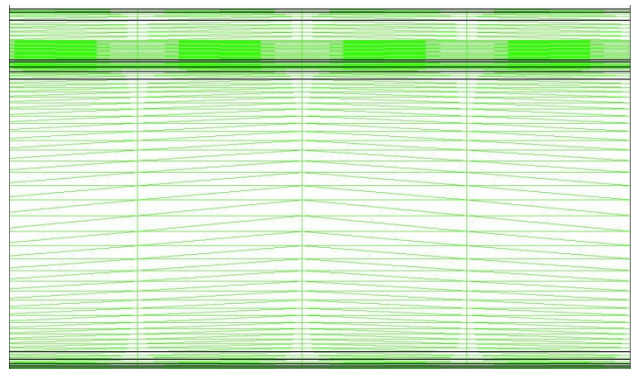
Fig. 2. Insight into the 2D Meshing of TJ solar cell

The tunnelling layer can be modelled as being one-dimensional in nature so that it can be calculated using a special rectangular mesh superimposed over and coupled to the ATLAS mesh. This mesh needs to include the junction region of interest and the direction of band-to-band tunnelling, which is generally perpendicular to junction interface.

Because of their electro-optical nature, the simulation of solar cells must necessarily take into account many detailed phenomena. From the optics point of view, the

simulator must allow the definition of standard solar spectra, such as AM1.5 (Fig. 3).
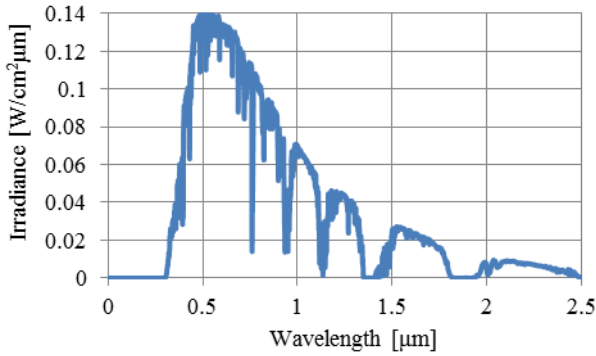
Fig. 3. Air mass spectrum 1.5

All simulations were performed at room temperature (27°C) and AM1.5 with one sun intensity.

A. Individual Cell Simulation

Current-voltage characteristics of the three individual solar cells are shown in Fig. 4. GaInP cell has an open circuit voltage (V_{OC}) of 1.405V, InGaAs has 0.945V and Ge has 0.347V. GaInP cell has the largest bandgap and should produce less current then the others, but in this case the additional layers contribute so the top cell has the highest current density. Fig. 4 shows that the InGaAs middle cell limits the short-circuit current in TJ cell and should be further optimized.

Quantum efficiencies of the solar cells under a concentration of one sun are displayed in Fig. 5. As previously mentioned, GaInP cell has the largest bandgap. Therefore, the top cell absorbs higher energy photons and lets the lower energy photons pass through. The lower energy photons continue to the middle cell, one number of photons is absorbed here and the rest pass through to the bottom cell. Ge has a low attenuation coefficient, so Ge layers are always built thick.
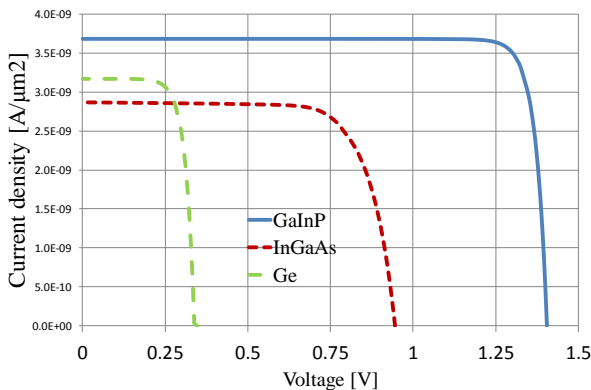
Fig. 4. J-V characteristics of single solar cells
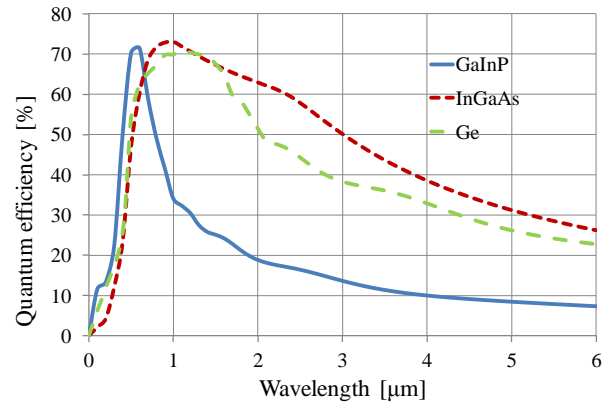
Fig. 4. Spectrum separation

The examples of real and imaginary optical indexes *n* and *k* of the cell's main materials, for wavelengths in the range from 0.2µm to 1.2µm, are shown in Figs. 6 and 7, respectively. These values and the indexes for other semiconductors in the TJ cell structure were inputted into ATLAS externally as separate material files. The data were collected from available literature. However, only a small number of data is published for ternary and quaternary materials, therefore some values had to be interpolated at higher wavelengths.
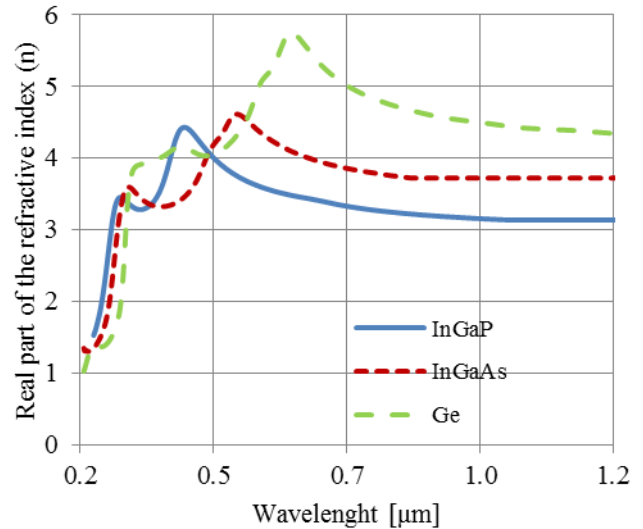
Fig. 6. Real part of the refractive index vs. wavelength
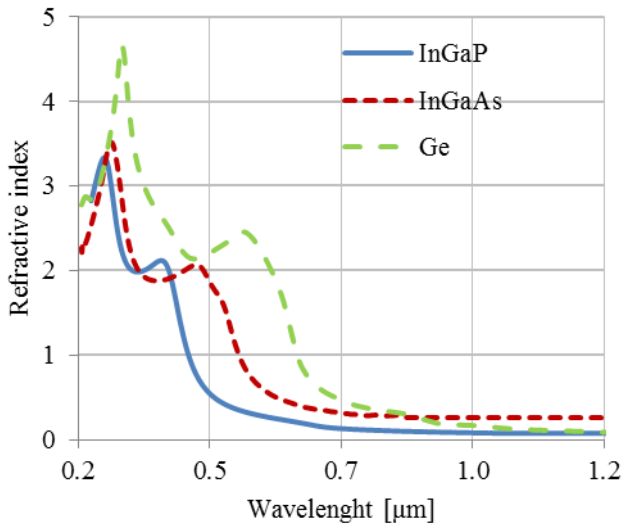
Fig. 7. Extinction coefficient vs. wavelength

B. Tunnel Junction Simulation

Tunnel junction is a very important layer, essential in the vertical stacking of more than one cell, in order to construct a multi-junction configuration. It should be optically transparent and connect the component cells in the multijunction structure with the minimum of electrical resistance. The optical absorption in the tunnel junction is minimised by using thin layers of wide bandgap materials, although care must be exercised as the tunnelling current decreases exponentially with increasing band-gap energy [4].

Fig. 8 shows J-V characteristics of both tunnel junctions, simulated without irradiance. It is found that the GaAs tunnel diode exhibits a smaller dynamic resistance in comparisons with the AlGaAs/GaInP tunnel diode.
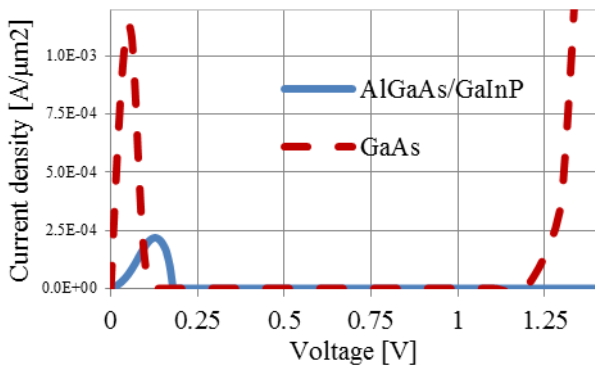
Fig. 8. Tunnel junction J-V characteristics

C. Complete Cell Simulation

Simulated J-V characteristic of the complete triple-junction solar cell is shown in Fig. 9. The predicted open

voltage of the TJ cell has a $V_{OC}=2.47V$, while the cumulative open circuit voltage of individual cells is 2.697V, obtained from Fig.4. This difference can be attributed to the voltage drop on the two tunnel junctions. The experimental data, obtained from IQE Europe Ltd. Manufacturing [3], are also displayed in Fig. 9.
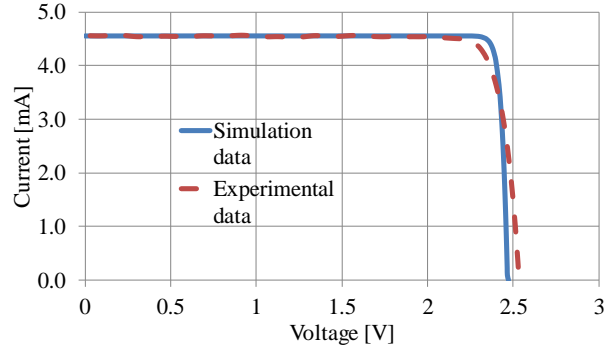
Fig. 9. J-V characteristic of the triple-junction solar cell
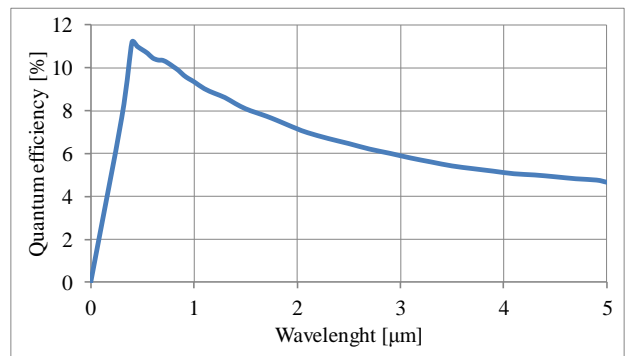
Fig. 10. Quantum efficiency of the triple-junction solar cell

The simulated external quantum efficiency of the complete TJ solar cell is shown in Fig. 10. The obtained cell efficiency is not satisfying, but at the time or writing this paper, the experimental data were not available and the simulation results need to be verified.

Fig. 11 displays the total current density flow of the top cell and the first tunnel junction. As shown in the figure, the maximum current density is located near the top contact, the cathode. Inside the tunnel junction, the current is not defined because the Ohm's law doesn't apply here, only quantum tunnelling.
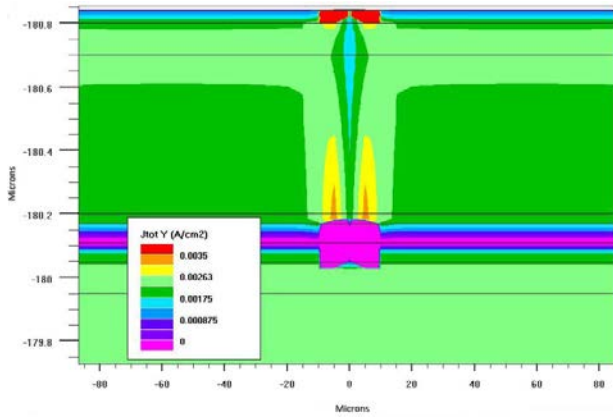
Fig. 11. Total current density flow
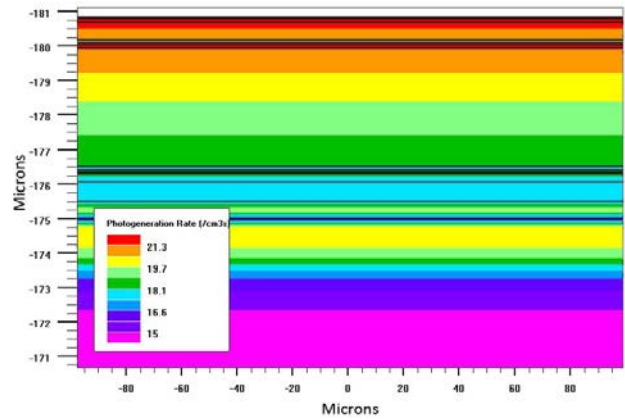
Fig. 13. Photogeneration rate

Photogeneration rate graphs are shown in Figs. 12 and 13. Individual solar cells produce current in different wavelengths, according to their frequency response. The highest photogeneration rate is in the upper layers because they are exposed to a greater number of photons. If the solar cells are observed individually, all emitters have a higher photogeneration rate then their bases.

Band energy diagram of the TJ solar cell with several heterojunctions is shown in Fig. 14.
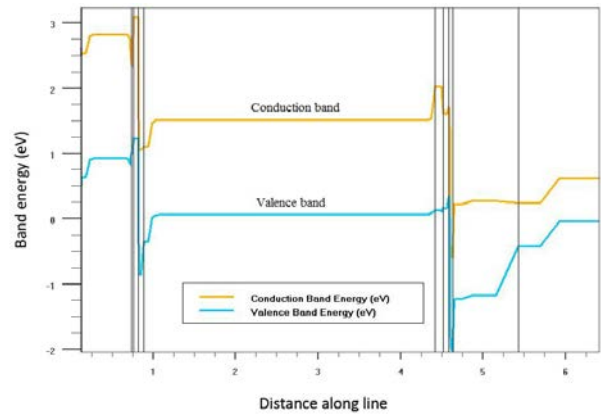
Fig. 14. Band energy diagram
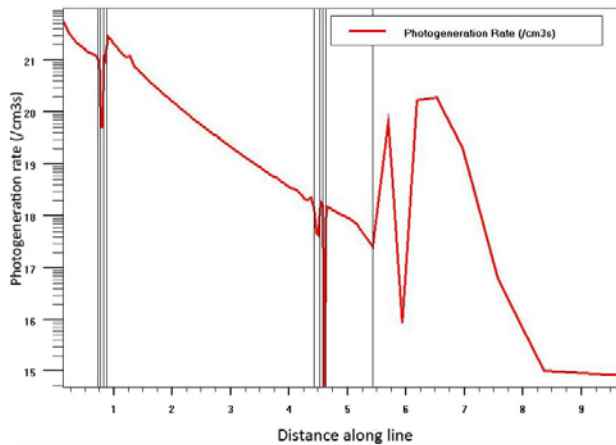
Fig. 12. Photogeneration rate through cross section

Finally, all materials used for building the triple-junction solar cell, with their parameters used in ATLAS simulation package are given in Table 1. The table illustrates all the complexity of the TJ cell simulation and the immense task for simulator calibration due to involving many physical, optical and material parameters that need to be carefully selected and verified.

Table 1
Material properties

Parameter	GaAs	GaInP	Ge	AlGaAs	AlInP	AlInGaP	InGaAs
	Material						
E_g [eV]	1.424	1.9	0.661	1.8	2.4	2.4	1.45
ε	12.9	11.62	16.2	12.3	11.7	11.7	11.7
$\frac{q\chi}{[eV]}$	4.07	4.08	4	3.54	4.2	4.2	4.05
N_c [cm^{-3}]	4.7e17	1.3e20	1e19	4.35e17	1.08e20	1.2e20	3.2e19
N_v [cm^{-3}]	9e18	1.28e19	5e18	8.16e18	1.28e19	1.28e19	1.8e19
c^{dir} [cm^{-3}s^{-1}]	7.2e-10	1e-10	6.41e-14	1.5e-10	1.2e-10	1e-10	7.2e-12
μ_n [cm^2/Vs]	8000	4000	900	2000	2291	2150	300
μ_p [cm^2/Vs]	400	1000	900	100	142	141	1
AUGN [cm^{6}s^{-1}]	5e-30	3e-30	1e-30	5e-30	9e-31	3e-30	/
AUGN [cm^{6}s^{-1}]	1e-31	3e-30	1e-30	1e-31	9e-31	3e-30	/
$\frac{m_n}{m_0}$	0.001	0.1	0.2	0.1	0.2	0.2	0.2
$\frac{m_p}{m_0}$	0.01	0.1	0.2	0.1	0.2	0.2	0.2

REFERENCES

[1] "Spire pushes solar cell record to 42.3%", 7.10.2010. http://optics.org/news/1/5/5

[2] T. Michalopoulos, "A novel approach for the development and optimization of state-of-the-art photovoltaic devices using Silvaco", 2002.

[3] "IQE Europe Ltd. Manufacturing". http://www.iqep.com

[4] "Silvaco ATLAS User's Manual ", 2010. http://www.silvaco.com

[5] M. Baudrit, C. Algora, "Modeling of GaInP/GaAs Dual-Junction Solar Cells including Tunnel Junction", 2008.

IV. CONCLUSION

In summary, a separated single junction GaInP, InGaAs and Ge solar cells were successfully simulated along with AlGaAs/GaInP and GaAs/GaAs tunnel junctions. Material and model parameters were optimized to achieve a better match with the experimental data. The single junction cells were then combined into a staked triple-junction solar sell which was also successfully simulated as a whole. Basic physical quantities such as band diagrams, optical photogeneration rate and current density are demonstrated. The simulated J-V characteristics of analysed TJ solar cell shows a close match with the experimental results obtained from the fabricated test TJ solar cell.

To the best of our knowledge, no published papers in conferences or journals reported a full simulation of a complete TJ multijunction solar cell, especially not the ones with simulations of its quantum efficiency. This conference paper is the first to show that the TJ collar cell with two tunnel junctions can be successfully simulated using Silvaco TCAD tools.

Operating Points and Topographic Dependence of the Thin Layer-Photovoltaic Cells as Relevant Characteristics for Modeling of the PV Cells

Duško Lukač, Miona Andrejević Stošović, Vančo Litovski

Abstract —The level of the DC voltage measureable in solar cells depends of the material of the semiconductor. Besides, the level of the produced open-circuit voltage is to a great extent independent from the occurring insolation that is to say from the intensity of the incident light insolation. Indeed, with an increasing light intensity rise the maximum current drawn and therefore also the output power, but not the electric voltage. Three most used PV technologies: modules from amorphous double-junction- silicon cells (A-Si), Cd-Te thin layer cells and conventional mono-crystalline silicon cells (SC-Si) have different behavior compared with the occurring sunlight and with it, with the spectral responsivity. Hence, the climatic differences between the topographic regions have direct consequences on the solar spectrum measurable at the ground (change of the Air – Mass factor AM) and with it on the spectral irradiance - E (λ). Those changes, influence in further the module performance depending of the spectral responsivity which signifies, that PV modules of different technologies point out, with different wavelength λ, a different change efficiency of photons in available load bearer pairs. This work explains the operating points and effects of the topography on the parameters of the PV cells.

Keywords —PV technologies, Topography, Operating Points, Modelling

I. INTRODUCTION

Solar cells consist of semi-conducting materials with different doping. In the transition between the n-semi-conducting layer and the p- semi-conducting layer with incidence of light, the separation of charge occurs, with which charge carriers are released. Via suitable contacts an electric tension can be measured - in a closed circuit it comes to an electric current flow. The attitude of the electric tension measureable in solar cells is depending on material of the semiconductor. With the cells used for photovoltaic (PV) arrangements made by silicon, electric tension measureable amounts to about 0.5 V. Besides, the

attitude of the generated open-circuit voltage is to a great extent independent of the occurring insolation that is to say from the intensity of the striking light insolation Indeed, with an increasing light intensity rise the maximum current draw and therefore also the output power, but not the electric voltage. The electric tension delivered in the normal mode is dependent basically on the load with a certain insolation, i. e. from the load current of the single solar cell or from the PV device. This is presented, according to [1, p.32] in the following figure.
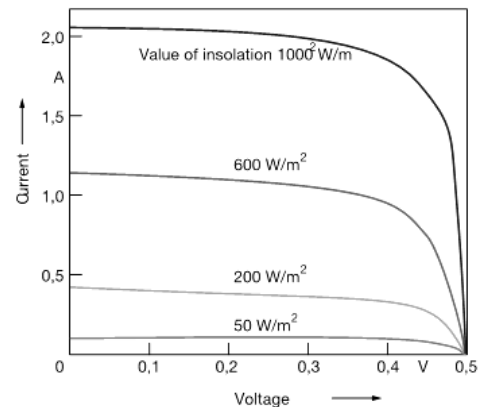
Fig. 1. Dependence of the electrical tension and current on the intensity of the occurring light insolation with solar cells.

II. RECOGNIZABLE TRENDS IN ICT DEVELOPMENT OPERATING POINTS OF THE THIN LAYER PHOTOVOLTAIC CELLS

In the practical case, three different operating points occur with the function of the PV cells. Those are: no-load operation (idle state), short-circuit state and regular operating state.

A. Characteristics of the no-load operation (idle state)

The electric voltage reaches in the idle state its maximum value, which is almost as high with a low illumination as with the strong illumination. The output power is zero because no electrical current flows.

B. Short-circuit state

The electrical current reaches its maximum value which depends on the intensity of the illumination as well as on the size and the efficiency of the respective solar cell. The electrical tension measured on the short-circuit cell is likewise as the output power zero valued.

C. Normal operating state

The working (operating) point is adjusted in that way, that a maximum output power by the solar cell is delivered. In addition electrical tension and current by the power

inverter are, according to the lighting situation, adapted constantly to the optimum values. Operating points are presented in the Figure 2.
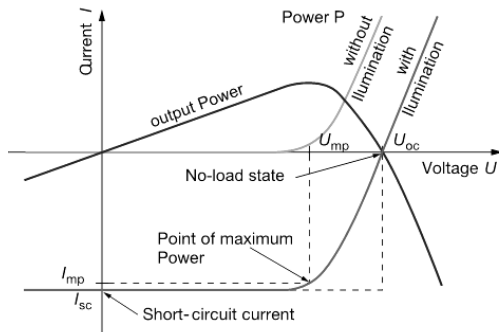
Fig. 2. Operating points of the PV cell.

Without illumination no charge carriers are released within a solar cell, so that no electrical current flow is possible. Also with low luminosities, e.g., with moonlight, firelight, road lighting or with application lighting of the fire department, extremely dangerous electrical tensions can appear in the PV devices. During the operation of the PV devices in the idle state, e.g., by disconnection of the DC cutting unit or by interruptions by ruined modules, it is absolutely possible that the electrical tensions reach the so called rated voltage of the PV device. However, the output power and the current will be only low on account of the low insolation intensity. Nevertheless, a danger for the application forces of the e.g. fire department or other rescue forces by the appearing electrical tensions is given; in any case, i. e. independent of the present lighting situation. The definition of the operating points is crucial information for the realization of the equivalent model of the PV cells.

III. TOPOGRAPHIC DEPENDENCE OF THE THIN LAYER-PHOTOVOLTAIC CELLS

There are different studies which show a strong dependence of the performance of the PV modules (cells) on changes in the insolation spectrum in module level as a result of changed environmental parameters as well as from the topographic location [2,3,4,5]. In order to be accurately measured, the free flied measurements needs to be carried out at least two places with the explicit topographical difference, which can be distinctive e.g. in terms of attitude (high/low) where the investigated PV modules are placed. For the measurement at different places, the same conditions have to be used, especially regarding the mounted degree of module inclination (in the regular case 30°-45°). The typical electric parameters of the PV modules for every technology as well as the influence of the surroundings parameters like insolation intensity, wind speed, temperature and cloudiness degree needs to be grasped in order to get accurate measurement. Also, in order to investigate the interconnection between the altitude of the sun and the PV module orientation the

modules need to be oriented in each case towards different direction as e.g. to the west, the south and the east. Different measures carried out shows that the basic behavior of the climatic circumstances is similar at the different locations. Determined can be increases of the insolation to midday, followed by a decrease up to the evening. Indeed, the daily course of the insolation factor (irradiance) E differs for the case that locations are differing considerable and based on it the topographical conditions are substantially different regarding their insolation range. Based on the study done by [5, p.60] such values for the locations of 170m over the MSL and for 1600m over MSL are presented in the next figures.
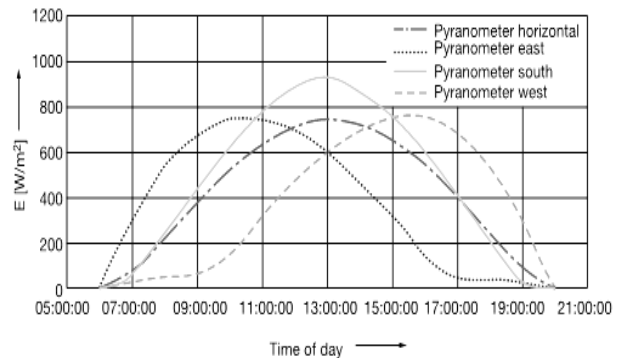
Fig.3. Values of insolation in dependence of orientation at 170m over the MSL
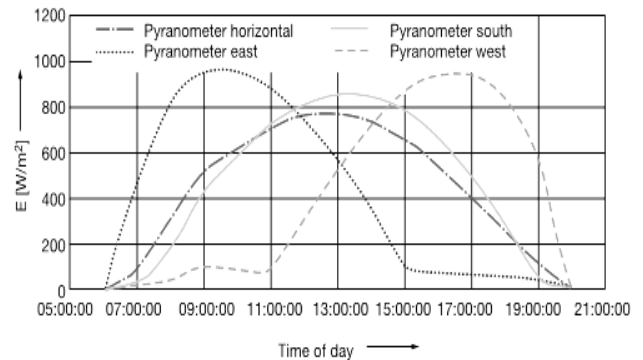
Fig.4. Values of insolation in dependence of orientation at 1600m over the MSL

The sensors and modules which are in the each case oriented to the west or to the east register for both locations in the morning or in the evening only vague insolation because of lacking direct sunlight. Only diffuse insolation becomes registered by the sensors. The climatic differences between the topographic regions, have direct consequences on the solar spectrum measurable at the ground, which can be explained with the change of the Airmass Factor and with it the occurring change of the spectral irradiation intensity in dependence of the wavelength - E (λ). Those changes influence in further result the module performance via spectral responsively which signifies, that PV modules of different technologies point with different wavelength λ, a different change efficiency of photons in available pairs of charge carriers. Air Mass (AM) explains that the solar radiation (insolation) is decreased on their way through the terrestrial atmosphere by reflection, absorption (by aerial

molecules and aerial particles) and dispersion. The decrease of the insolation grows, the longer the way of the radiation through the terrestrial atmosphere is. The factor AM declares how long the way of the solar radiation is thought the terrestrial atmosphere and it is given in proportion to atmosphere-thick. With the vertical solar state, the light takes the shortest way through the atmosphere, so that in that case is AM scored with "1". If the sun stands a little bit sloping, its way through the atmosphere is extended, so that in that case AM grows. By using of the Air Mass calculator [6] with the following settings (Rayleigh atmosphere (no aerosols), full multiple scattering, with refraction, spherical atmosphere, no clouds, US standard atmosphere) here are shown some examples for the Air Mass in dependence of solar zenith angle calculated by using of the different attitude (500m and 150mm) and Albedo (10 and 90) with the constant wavelength of 500 nm.
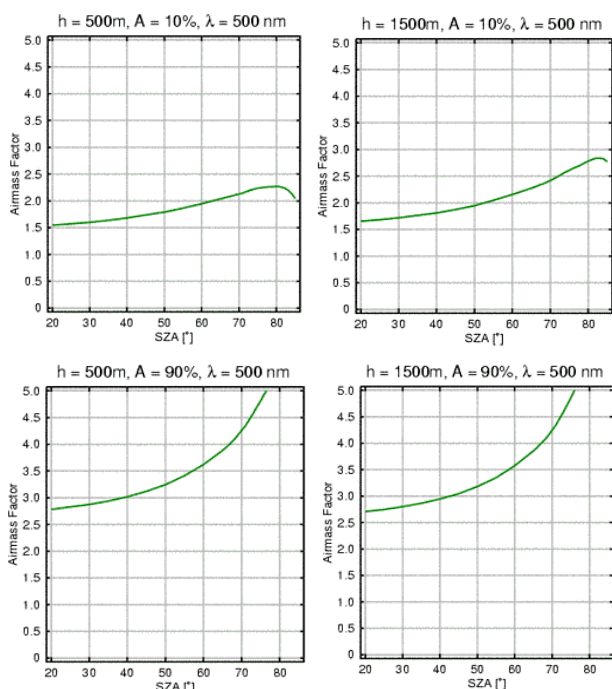
Fig.5. Air Mass (AM) factors at 500m (left) and 1500m (right) with different Albedo factors

Albedo factor describes the reflection of the sunlight on the earth surface. It is given in percent to the incident light and amounts on the earth surface to about 10-20% in the normal case, and with a reflective surface (e.g. snow) about 80%. By the calculation of the portion of a certain wavelength in the respective whole spectrum, of the spectral component $p(\lambda)$ different times of day, weather conditions and insolation intensities can be directly compared. Such comparisons has been made by e.g. [6,7]. The results show that on the basis of the processes of dispersion in the atmosphere which basically depends on the Rayleigh dispersion and Mie dispersion [8, 9]) the portion of the short-wavy light (<500 nm) lies with the integral insolation intensity E is higher than the portion of the long-wavy light (> 700 nm). By contrast is spectral

component for attitudes above 1,600 ms over MSL for all wavelengths and time of day are roughly consistently [2, 3, 5, 7]. Rayleigh scattering (dispersion) can be described as scattering in the small size factor regime. Dispersion from larger spherical particles is explained by the Mie theory for a random size parameter. For small size factor the Mie theory reduces to the Rayleigh approximation. Rayleigh dispersion explains the greater proportion of blue light (short-wavy light) scattered by the atmosphere relative to red light (long-wavy light). According to [5, p.61], comparing the short – circuit currents characteristics at 6h P. and M 8h A.M. (I_{SC}(6h p.m.) / I_{SC} (8h a.m.)) , by using the Modules from amorphous Double-Junction- Silicium cells (A-Si), Cadmium-Tellurid (Cd-Te) - thin layer cells and conventional mono-crystalline Silicium cells (SC-Si), as reference cells, the comparison of the short – circuit currents is presented in the following Table:

	East	South	West
>170m MSL			
SC-Si	1	1	1
A-Si	1.50	1.08	0.71
Cd-Te	1.29	0.92	0.94
>1600m MSL			
SC-Si	1	1	1
A-Si	1.26	1.04	0.70
Cd-Te	1.04	1.08	0.92

TABLE 1. Comparison of the short – circuit currents characteristics at 6h p.m. and 8h a.m (I_{SC}(6h p.m.) / I_{SC} (8h a.m.) of SC-Si , A-Si and Cd-Te modules

The figure show the behavior of the modules concerning the short circuit current I_{SC}. The I_{SC} values of the mono-crystalline silicon modules (SC-Si) in comparison with the thin layer modules (A-Si and Cd-Te) show only very small variations. For the attitude under 170 ms MSL, looking on the A-Si modules and Cd-Te modules oriented to the south, however, the I_{SC} value measured in the morning decreases in comparison to evening values less strongly than with the SC-Si modules. Looking at to the east oriented A-Si and Cd-Te modules; it can observed, that I_{SC} (6h p.m.) / I_{SC} (8h a.m.) shows even an increase. By contrast I_{SC} decreases with the A-Si-aimed to the west and Cd-Te modules in the day course. The changes directly arise from the different spectral components $p(\lambda)$ and the orientations. This effect is further strengthened by purely vague isolation, that is to say bigger portion in blue and ultraviolet light. Generally the effects become more evident in the areas of the smaller attitudes (170 m) in regard of the spectral components, than in the areas of higher attitudes (1,600 m). Looking on the Table 1 it can be observed that with the exception of the modules oriented to the west, all thin layer modules in comparison with the mono-crystalline modules at 6h p.m. shows higher currents than around 8h a.m.. The effect can be explained

with the adjustment, the topography of the locations and the day-temporal changes of the solar spectrum in the module level. That information can be used for modeling of the certain cells in regard of the topographic influence on the solar cells.

IV. CONCLUSION

For the modeling of the PV cells behind the general operating modes also topographical influences on the cells made by using of the different production technologies are important.

REFERENCES

[1] Baade, V. (2011), "Gefährliche Spannung bei geringer Lichtstärke", EP Photovoltaik, 3/4 2001,P. 32

[2] Ebert G. et all, (2009), "Four years of PERFORMANCE – was it worth the effort?", 24th European Photovoltaic Solar Energy Conference, September 21-25, 2009, Hamburg, Germany

[3] Friesen G. et all (2009), "Inter-comparison of different energy prediction methods within the European project "PERFORMANCE" – results of the 2nd round-robin" 24th European Photovoltaic Solar Energy Conference, September 21-25, 2009, Hamburg, Germany

[4] Pearsall N.and Atanasiu B. (2009) "Assessment of PV System Monitoring Requirements by Consideration of Failure Mode Probability",24th European Photovoltaic Solar Energy Conference, September 21-25, 2009, Hamburg, Germany

[5] M. Rennhofer, M. (2010), "Topographischer Einfluss auf Dünnschicht-Photovoltaik", EP Photovoltaik, 11/12 2010, P.60-62

[6] Airmass Factor Calculator (2011), http://www.doas-bremen.de/airmassfactors.htm, Available at 21.December 2011, University of Bremen

[7] Wagner, J.E. et all (2010), „Einfluss von Sonnenspektreum und Klima auf die Performance von c-si und CdTe Modulen", Proceedings to the "25. Symposium Photovoltaische Solarenergie" vom 3. bis 5. März 2010, Kloster Banz, Bad Staffelstein, (2010)

[8] Chakraborti, S. (2007), "Verification of the Rayleigh scattering cross section", American Journal of Physics, 75:9, pp. 824−826

[9] Sneep,M. and Ubachs,W.(2005),"Direct measurement of the Rayleigh scattering cross section in various gases", Journal of Quantitative Spectroscopy and Radiative Transfer,92:2005, pp.293-310

Solar Energy Harvesting for Wireless Sensor Nodes

Velibor Škobić, Branko Dokić, Željko Ivanović

Abstract – In this paper, we propose a system for collecting solar energy in the ambient environment. The system power supplies a module whose role is to measure the ambient temperature. For system supplying is used battery and small size of solar panel. To compensate the lost energy during the active period, measurements are performed at least every 95s. Characteristics of the used module and solar panel are shown in the paper. Simulation and practical results are also presented in this paper.

I. INTRODUCTION

In most aplications, *w*ireless sensor networks consist of a number of wireless sensor node in the network. A common task in network devices using sensors is to either obtain information about the measured values (temperature, pressure, etc..) which is afterwards processed and then forwarded through the network, or to manage small processes. ZigBee standard is just one of numorous technologies which is designed for applications of wireless sensor networks. ZigBee devices are based on low-power, small size and low cost, what makes them suitable for the realization of wireless sensor networks. They consist of a microcontroller and the transceiver with an antenna, and all these together are called ZigBee module. In order to carry out its tasks, the device consumes appropriate energy. Most applications require that the device within the network is mobile and battery powered. In this case it is necessary to replace a discharged battery with a new one. For some applications this process can be complicated while in others even impossible. In order to increase the lifespan of the batteries, there was a justifiable need for introducing alternative energy sources. The amount of energy needed to operate the device in the network is small, and that is why we are using alternative sources of energy, such as solar energy, temperature differences, vibrations, etc. [1]. From these sources we can collect enough energy to power the device within the network. Solar energy is the most common source of energy that is available throughout the day from sunlight or from the artificial light at indoor conditions. Vibrations, kinetic and mechanical energy generated by the object movement can also be converted into electrical energy and adequately stored. Thermal energy harvesting uses tempreature differences or gradients to generate electricity. Depending on what kind of energy

Velibor Škobić is student of second cycle studies, Faculty of Electrical Enginering, University of Banja Luka, Bosna i Hercegovina, E-mail: veliborskobic@gmail.com.

Branko Dokić and Željko Ivanović are with the Department of Electronics and Telecommunications, Faculty of Electrical Engineering, University of Banja Luka, Patre 5, 78000 Banja Luka, Bosna i Hercegovina, E-mail: {bdokic, zeljko}@etfbl.net.

source is available to us, we use an appropriate system for collecting energy. The system can also have a hybrid character in which the energy is collected from two or more sources [2]. In Table I values of these power sources are shown [3].

This paper introduces a way of power ZigBee modules using solar energy in the ambient environment. Power system consists of solar panels, boost converter, battery and super capacitor. The role of modules is to measure the ambient temperature, and then forward information about the measurement through the network. The second section provides a description of the solar energy and solar panel model. In the third section proposed scheme is for power supplying ZigBee module and theoretical analysis of the module energy consumption. In the fourth section are shown results of measuring the consumption of the system for temperature measurement and comparison with a theoretical consideration.

TABLE I
POWER DENSITIES OF ENERGY HARVESTING TECHNOLOGIES [3]

Sollar cells (outdoors at noon)	$15 mW / cm^3$
Piezoelectric (shoe inserts)	$330 \mu W / cm^3$
Vibration (small microwave oven)	$116 \mu W / cm^3$
Thermoelectric ($10^o C$ gradient)	$40 \mu W / cm^3$
Acoustic noise (100dB)	$960 nW / cm^3$

II. INDOOR SOLAR HARVESTING

The most common restorable source of energy in this kind of systems is solar energy. Depending on intensivity of the lightening, we can collect apropriate amounts of energy. Solar cells are elements which convert solar into electrical energy. There are two ways for collecting the energy, directly from the sun or indoor collecting- ambient light source. The energy from outdoor conditions, collected from sunlight, is than the one collected in indoor conditions (Table II). Solar panel whose main role is to convert solar energy into electricity does not give a constant amount of energy over time. During the office hours of the day, the panel has the ability to convert energy with more success because of relatively good lighting. After ending a office hours, there is no lighting rest in the room and no ability to collect energy anymore. In applications that use this method for collecting the energy time intervals have to be considered which give the information whether the source is available or not.

TABLE II
SURFACE DENSITY OF SOLAR ENERGY [2]

	Solar panel
Indoor condition	$100\,\mu W / cm^2$
Outdoor condition	$10\,mW / cm^2$

Consumption reduction can be provided by the device the network that is not constantly active, but there is a certain time interval when the device goes to the inactive state lowering the power consumption. In its active state module communicates with the other devices in the network and its current supply is significant, about $10\,mA$. In an inactive state, microcontroller and the transmitter module are excluded thus reducing consumption to several μA. Smaller dimension of solar panels can not provide enough power for continuous operation of modules in an active state. Consumption in the active period must be less than the amount of energy that is collected in periods of inactive time. Collected energy is used to carry out tasks and send data through the network. The time between battery replacements depends on the amount of sent and received data through the network [4].

In Fig. 1 a block diagram of a typical wireless sensor node, which is powered by combination of energy scavenging and battery technology, is shown [5]. The system consists of sensors which can observe environment, an analog-to-digital converter (ADC) which can quantize the analog signal from the sensors, a digital signal processing (DSP) core which can analyze and encode the quantized data and a transceiver (RF) so that the node can transmit and receive information. Light energy is converted to electrical energy trough a photodiode and mechanical vibrations are converted to electrical energy by an electromehanical transducer.
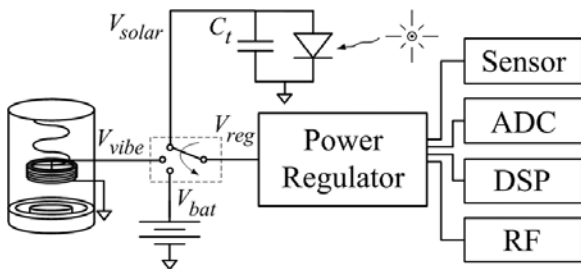
Fig. 1. Low power wireless system powered from energy scavengers and a batery. Energy sorces include solar, mehanical vibration and a battery. A multiplexer switches between the unregulated energy sources [4].

The solar panel is modeled (Fig. 2) with the electric generator I_{Light} as a function of light intensity, the diode with the inverse saturation current I_O and serial R_S and parallel R_P resistance [6].
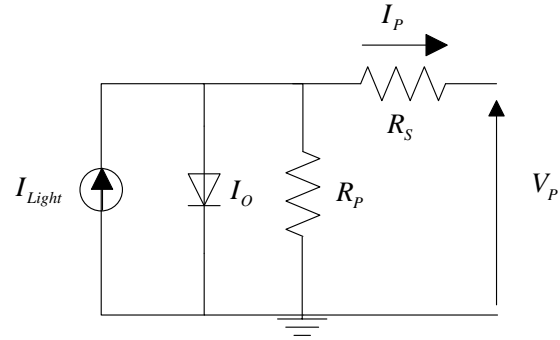
Fig. 2. Model solar cells

Dependence of output current I_P and output voltage V_P is given by the equation [6]:

$$I_P = I_{Light} - I_O (e^{\frac{V_P + I_P R_S}{kN_S T_C / q}} - 1) - \frac{V_P + I_P R_S}{R_P}, \quad (1)$$

where: q is electron charge, k Boltzman constant, T_C temperature potential and N_S number of cells in series.

The maximum energy utilization from solar panels is obtained by setting the working point of the cell to the place of the maximum power point (MPP) [3]. For this purpose, various algorithms have been developed to achieve optimum operating point, i.e. maximum utilization of solar cells [7].

III. SUPLLYING ZIGBEE MODULE

The functional block diagram for proposed solution which is providing power to the ZigBee module is shown in Fig. 3. It consists of solar panels, boost converter, super-capacitors and batteries. The role of solar panels and boost converter is to power ZigBee module. In the case of no solar light, the device is battery supplied. In a case that solar panel provides more energy than it is enough to supply module, the rest of the energy it is used to charge the battery.
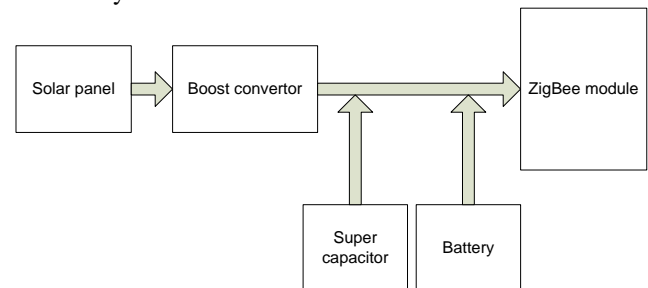
Fig. 3. Functional block diagram

The communication device mode is based on the ZigBee standard, ATMEL ATZB-24-A2 and consists of the microcontroller ATmega1281 and RF transceiver

AT86RF230. In active mode and idle mode, supply current is $I_A = 19mA$ and $I_N = 6\mu A$, respectively. Solar panel is used from the calculator "GENIE". Solar panel is based on four cells conected in series producing voltage of 2.4V at the output of the panel. The maximum measured power of solar panels in the ambient environment is $P_{Pmax} = 20\mu W$. The capacitance of the used super capacitor is $0.1F$.

A. Calculation of power consumption

In order to calculate power consumption, it is necessary to know, besides the values of I_A and I_N, the values of power supply module. The maximum and minimum power supply voltages are $V_{M\,max} = 3.6V$ and $V_{M\,min} = 1.8V$, respectively. The active mode time is $T_A = 10ms$. Voltage at the capacitor before the active period is 3.3V. After the module goes in active mode, voltage on the capacitor decreases. Required amount of charge for an active period is :

$$Q = I_A \cdot T_A = 190\mu C \qquad (2)$$

The capacitor voltage drops to a value in time T_A:

$$\Delta V = \frac{\Delta Q}{C} = \frac{190\mu C}{0.1F} = 1.9mV \qquad (3)$$

The maximum permited voltage amount change is:

$$\Delta V_{max} = V_{M\,max} - V_{M\,min} = 1.8V \qquad (4)$$

Thus the number of active periods in which the module is working properly for one charging capacitor, where voltage changed from $V_{M\,min}$ to $V_{M\,max}$, is:

$$N = \frac{\Delta V_{max}}{\Delta V} = 947 \qquad (5)$$

Maximum size of the physical layer header of ZigBee standard is 128 bytes, with the maximum useful payload 104 bytes. Taking this into consideration we find that the maximum number of transferred data for a single charging capacitor is:

$$N_B = N \cdot N_{kar} = 947 \cdot 104 = 98488 \qquad (6)$$

The output voltage of used solar cell is 2.4V and current depends on the lighting intesity. At maximum power consumption ($20\mu W$), through measurement we obtain the value of solar panel current as $I_{sol} = 8\mu A$ under indoor conditions. Thus, minimal required time needed to recharge the super capacitor is given as:

$$T_{ch} = \frac{\Delta V_{max} \cdot C_{super-capacitor}}{I_{sol}}$$

$$= \frac{1.8V \cdot 0.1F}{8\mu A} = 22500s = 6.25h \qquad (7)$$

As it is calculated, during an active period of time it is spent $190\mu C$ of charge, so we have the $I_{sol}(T_A + T_N) = T_A \cdot I_A + I_N \cdot T_N$, the time that solar panel spends to compensate the charge:

$$T_N = T_A \cdot \frac{I_A - I_{sol}}{I_{sol} - I_N} = 95s \qquad (8)$$

The result of simulation for the circuit in Fig. 4 is shown in Fig. 5, where the solar panel is replaced by current generator and a ZigBee module is replaced by current pulse generator with pre-calculated times of active and inactive modes and the corresponding currents.
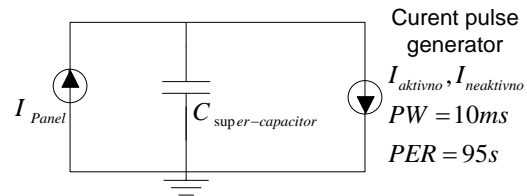
Fig. 4. Electrical circuit of supply mode
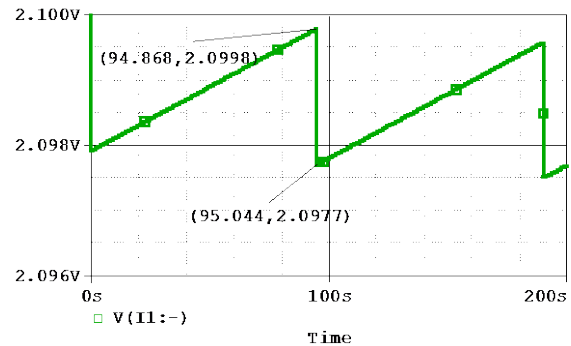
Fig. 5. The results of simulation of ZigBee module power consumption

The result shown in Fig. 6 proves theoretical analysis already given in this paper. There it is shown that 95s of the inactive working mode is enough to collect the charge lost during the active working mode. In this way we can increase the time between battery replacements, because the system supply is capable of collecting enough energy in the minimal time of 95s, for an active period of work time mode.

IV. EXPERIMENTAL MEASUREMENTS

In order to verify the preliminary examination, the prototype was implemented (Fig.6.) to measure temperature. It consists of super capacitor, module ATZB-24-A2 and circuits for measuring the temperature and voltage of power supply for module. The system for measuring temperature is implemented with using a thermistor. Thermistor is chosen to increase the speed of measuring a temperature and to reduce power consumption. From a voltage divider the voltage is assessed with using analog - digital converter (AD), in order to obtain information of the temperature. The power supply voltage module is not constant, so that the use of voltage dividers can yield information about the value of power supply voltage. If the supply voltage drops below a certain critical value, the module has the ability to send information about reducing the voltage, across the network and goes to idle mode longer time interval. Using the GPIO pins (General Purpose Input / Output) ensures that the circuit for measurement of voltage and temperature during the inactive period is turned off in order to reduce consumption of energy. After taking the information about voltage dividers, voltage information is then sent through the network. Further processing samples obtaines information about the ambient temperature. After sending data, module goes into inactive mode specific time interval.
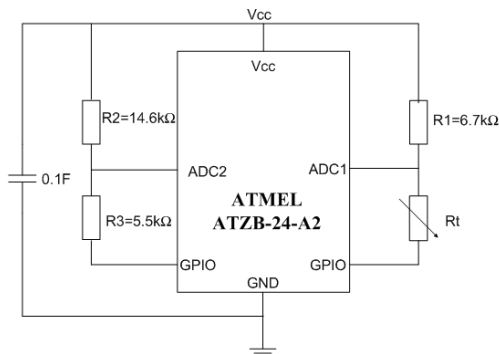
Fig. 6. Electrical cyrcut for mesuring a temperaure.

Realization of prototypes and the corresponding measurements yielded following results. The temperature was measured every 10s. Capacitor is charged to a value of 3.3V. For 170 active periods capacitor voltage fall to 2.8V. Taking into account the consumption of circuits for measuring temperature and supply voltage the expected theoretical value is 195 measurements. The result does not deviate much from theoretical considerations, taking into account the deviation values of capacitance super capacitor and the deviation of consumption circuit for measuring voltage and temperature.

V. CONCLUSION

A particular challenge in creating wireless sensor networks is the realization of power nodes in the network. In this paper it is shown that it is possible to provide adequate power for ZigBee modules using solar panels with small dimensions. The usage of this power supply type is limited to applications which do not require frequent processing and sending data. Solar energy is collected in the ambient environment that is considerably smaller than the energy available in outdoor applications. In order to exploit the maximal solar energy and more efficient energy storage it is recommended to use the boost converter.

REFERENCES

[1] Winston K.G. Seah, Zhi Ang Eu Hwee-Pink Tan, *"Wireless Sensor Networks Powered by Ambient Energy Harvesting (WSN-HEAP) – Survey and Challenges"*, Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference, pp. 1-5.

[2] Yen Kheng Tan, Sanjib Kumar Panda, *"Energy Harvesting From Hybrid Indoor Ambient Light and Thermal Energy Sources for Enhanced Performance of Wireless Sensor Nodes"*, IEEE transactions on industrial electronics, Vol. 58, No. 9. September 2011.

[3] Adel Nasiri, Salaheddin A. Zabalawi, Goran Mandic, *"Indoor Power Harvesting Using Photovoltaic Cells for Low-Power Applications"*, IEEE transactions on industrial electronics, Vol. 56, No. 11, November 2009.

[4] Swati V. Sanklap, Vishram Bapat, *"Comparison of Proposed Solar Energy Efficient MAC (SEHEE-MAC) with ZigBee and Preamble MAC for SHM in Wireless Sensor Network"*, International Journal of Engineering Science and Technology (IJEST), Vol.3 No.9 September 2011.

[5] Nathaneil J. Guilar, Travis J. Kleeburg, Albert Chen, Diego R. Yankelevich, Rajeevan Amirtharajah, *"Integrated Solar Energy Harvesting and Storage"*, IEEE transaction on very large scale integration systems, Vol. 17, No. 5, May 2009.

[6] Denis Dondi, Alessandro Bertacchini, Davide Brunelli, Luca Larcher, Luca Benini, *"Modeling and Optimization of a Solar Energy Harvester System for Self-Powered Wireless Sensor Networks"*, IEEE transaction on industrial electronics, Vol. 55, No. 7, July 2008.

[7] Cesare Alippi, Cristian Galperty, *"An Adaptive System for Optimal Solar Energy Harvesting in Wireless Sensor Network Nodes"*, IEE transaction on circuits and systems, Vol. 55 No. 6 July 2008.

Realistic Modeling and Simulation of The PV System - Converter Interface

Miona Andrejević Stošović, Duško Lukač, and Vančo Litovski

Abstract — In this paper we will present main problems and existing solutions concerning photovoltaic (PV) cells modeling and simulation. After several experiments performed by simulation we come to the conclusion that dynamic model of the PV cell is needed in order to get a realistic picture of its working conditions.

Keywords — modeling, photovoltaic cells, circuit simulation.

I. INTRODUCTION

In our recent proceedings [1] we made an overview of the existing applications of the photovoltaic (PV) cell models and the corresponding PV panel models. Our main interest was to search for application of dynamic modeling of the PV system. A large set of published results was consulted, to mention only a few of them [2-14], and we came to a conclusion that no dynamic circuit modeling was exercised at all. In fact, under dynamic modeling of PV systems thermal transient analysis was understood.

The reasons for such a situation, in our opinion, are several. First, the changes of the excitation to the PV system i.e. the light intensity are incomparably slower than the transients (local time constants) in it. Second, in existing applications, parallel to the PV system a capacitor with large capacitance is connected. It is assumed that its capacitance is at least by order of magnitude larger than the output capacitance of the PV system so suppressing any oscillations. Finally, it is a common practice to separately design the PV system and the DC to AC electronic conversion chain. In that way the interaction between the input of the converter and the output of the PV system is overlooked.

It is our intention here to put some more light to the electrical interface between the PV system and the DC/DC converter that is first encountered in the conversion chain. It will be shown by simulation that, due to the commutations within the converter, the output voltage of the PV system by no means is as simple as a DC voltage. In addition the properties of large capacitors will be exposed to show that inductive behavior may be expected

at the harmonic frequencies of the controlling signal of the converter. Finally, sudden faults and especially intermittent ones are expected to seriously disturb the DC levels within the PV panel which, we expect, will give rise to transients in which the dynamic properties of the PV cells come to the fore.

The paper is organized as follows. The common model of a solar cell will be introduced first. An equivalent Norton source will be extracted in order to simplify the proceedings. Then, a DC/DC converter simulation results will be given to show the signals at its inputs. A simplified model of the interface will be created and simulation results will be given to show the output voltage of the PV system (input voltage of the DC/DC converter) in different situation. After introducing the so called link (electrolytic) capacitor and its model final conclusions will be drawn related to the need of modeling the dynamic properties of PV cells.

II. PV CELL MODELS

A function of a PV cell is simple: it absorbs photons from sunlight and releases electrons, so when there is a load connected to the cell, electric current will flow. PV cells are based on a variety of light-absorbing materials, including mono-crystalline silicon, polycrystalline silicon, amorphous silicon, thin films such as cadmium telluride (CdTe) and copper indium gallium selenide (CIGS) materials, and organic/polymer-based materials.

A PV cell is usually represented by a light-induced current source in parallel with a diode, as shown in Figure 1. The output of the current source is proportional to the light flux falling on the cell. The diode determines the I-V characteristics of the cell.
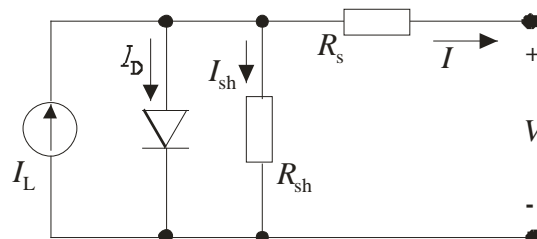
Figure 1. Circuit model of PV solar cell

Because of material defects and ohmic losses in the cell substrate material as well as in its metal conductors, surface, and contacts, the PV cell model also must include

Miona Andrejević Stošović and Vančo Litovski are with the University of Niš, Faculty of Electronic Engineering, 18000 Niš, Serbia. e-mails: (miona.andrejevic; vanco.litovski@elfak.ni.ac.rs).

Duško Lukač is with the University of Applied Sciences, Cologne, Germany, e-mail: lukac@rfh-koeln.de.

series resistance (R_s) and shunt resistance (R_{sh}), respectively, to account for these losses. R_s is a key parameter because it limits the maximum available power (P_{max}) and the short-circuit current (I_{sc}) of the PV cell.

The R_s of the PV cell may be due to the resistance of the metal contacts on the cell, ohmic losses in the front surface of the cell, impurity concentrations, or junction depth. Under ideal conditions, R_s would be 0 Ω. The R_{sh} represents the loss due to surface leakage along the edge of the cell or crystal defects. Under ideal conditions, it would have an infinite value and in most of the literature it is neglected in order to simplify the electrical model. But, in [15] it is shown that at very low irradiances, its value increases dramatically, i.e. the contribution of the apparent shunt resistance is only significant for cell voltages below about 0.45 V, and depends on irradiance.

The equations describing I-V characteristics of the PV cell based on equivalent circuit shown in Fig. 1, are usually expressed in the form given below,

$$I = I_L - I_0(e^{\frac{q(V+I\cdot R_S)}{k\cdot T}} - 1) - \frac{V + I \cdot R_S}{R_{sh}} \qquad (1)$$

$$I_D = I_0(e^{\frac{q(V+I\cdot R_S)}{k\cdot T}} - 1) \qquad (2)$$

where I is the cell current; q is the charge of electron; k is the Boltzmann constant; T is the cell temperature; I_L is the light generated current; I_0 is the diode saturation current; R_s and R_{sh} are the cell series and shunt resistances, V is the cell output voltage.

The electrical properties of the cell as a function of the ambient irradiance are captured within the expression of I_L while the cell temperature influence is mainly expressed through (2).

The suitability of this way of expressing of the I-V characteristic of the PV cell was discussed in [1] and alternative expressions were suggested allowing for circuit simulation of the PV cell in a complex electronic surroundings. Such an expression asks for a new variable to be introduced in the model, namely, the cell`s internal voltage V_i. If so, having in mind the notation of Fig. 2, the following nodal equations may be written

$$\begin{aligned} I_D + I_{sh} + I &= I_L \\ -I + I_{out} &= 0. \end{aligned} \qquad (3)$$
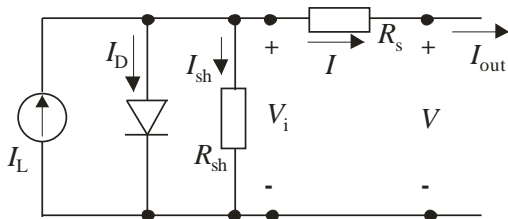
Figure 2. Modified interpretation of the PV cell model

All, I_D, I_{sh}, and I may represent models of nonlinear

voltage controlled elements as does I_D in Equ. (2). I_L is here considered as voltage independent. In the case, however, when some voltage dependent light emitting element is in the electronic circuit and it illuminates the PV cell, I_L may become voltage dependent and will be shifted to the left-hand side of the nodal equation.

For the most frequent case, when linear R_s and R_{sh} are expected, one may use the following nodal equations

$$\begin{aligned} \frac{1}{R_s}(V_i - V) + \frac{V_i}{R_{sh}} + I_D &= I_L \\ \frac{1}{R_s}(V - V_i) + I_{out} &= 0 \end{aligned} \qquad (4)$$

where

$$I_D = I_0(e^{\frac{q\cdot V_i}{n\cdot k\cdot T}} - 1), \qquad (2)$$

and: I_{out} is the load current (Most frequently $I_{out} = V/R_{load}$, where R_{load} is the load resistance), n is the p-n junction`s ideality factor.

Introduction of the cell`s nonlinear capacitances in the model is a straightforward task as shown in Fig. 3.

Using this concept, if model parameters available, simulation of photovoltaic systems containing virtually unlimited number of PV cells and electronic circuitry of any complexity may be simulated using standard electronic circuits analysis methods [15].
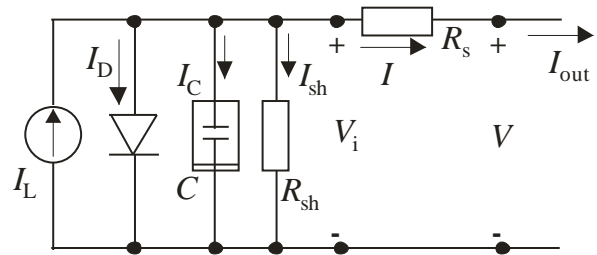
Figure 3. Modified interpretation of the PV cell model
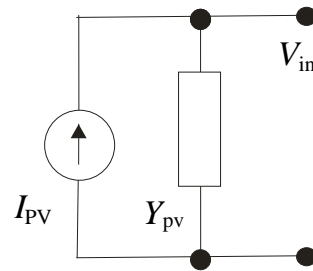
Figure 4. A nonlinear model of the PV system

III. THE CONVERTER

The output circuitry of a PV system, as complex as it can be [16], may be modeled as a current source I_{pv} (equivalent Norton) with internal admittance Y_{pv} as shown in Fig. 4. Note the admitance is to be nonlinear since it represents the nonlinearities of the diode(s), the junction capacitance(s) and the resistances. Here, however, the

purpose of modeling is to get a rough picture of the PV system-converter interface and no details will be given about the PV-model parts.
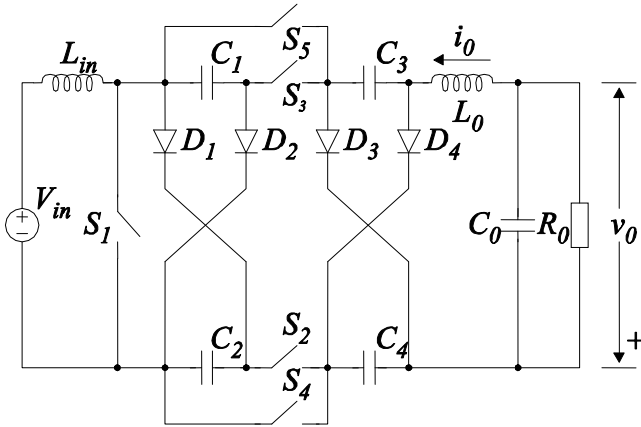
Figure 5. The Ćuk converter
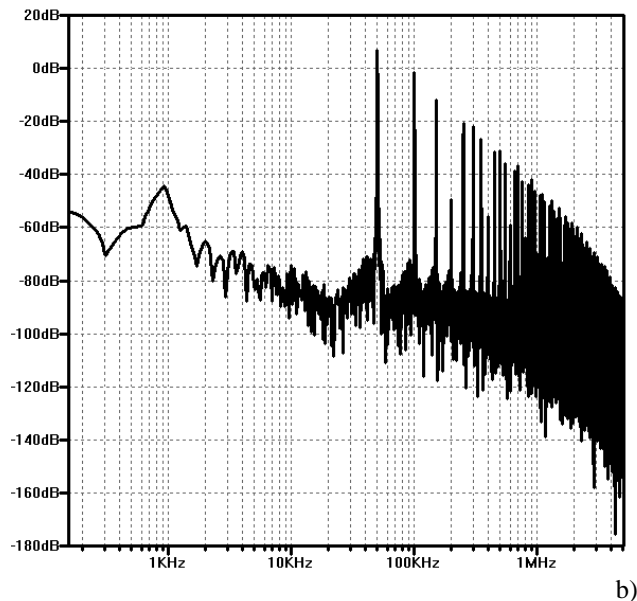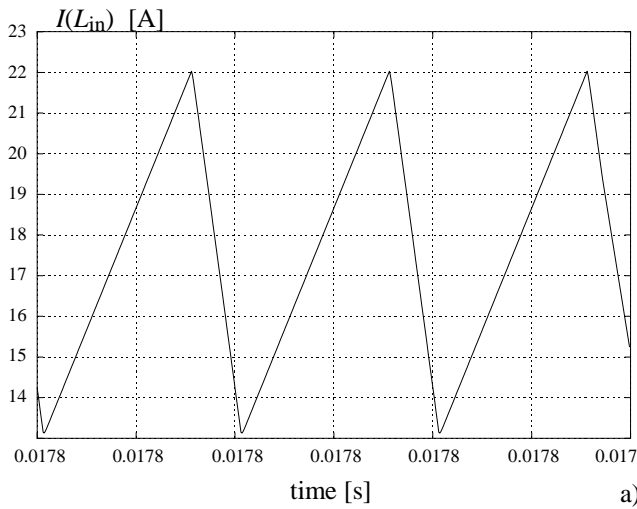
Figure 6. The input current of the Ćuk inverter after steady state (a) and its spectrum (b)

This circuit in most cases is driving a DC to DC converter. One, among many, variant of the DC/DC converter is the Ćuk converter shown in Fig. 5. Here

constant voltage excitation of 12 V is assumed while the switching frequency is 50 kHz. For this proceedings the input current trough the coil (L_{in}=20 μH) is of interest. It was obtained by simulation and part of the response is shown in Fig. 6a. The corresponding spectrum is depicted in Fig. 6b.

It may be observed that the input current, in addition to the DC component, has an AC component reach of harmonics. Accordingly, when connected to the PV system such a converter will draw alternating current in addition to the DC power which was targeted.

IV. MODELING THE COMPLETE SYSTEM

The simplified schematic of the complete system is depicted in Fig. 7 which is considered self-explanatory. Fig. 8 represents a model of the whole system where the PV system, as modeled in Fig. 4, is loaded by the input resistance (R_{in}) of the converter and excited by alternating current (depicted in Fig. 6) labeled by J_{AC}.
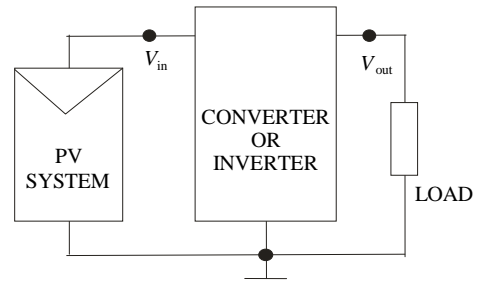
Figure 7. A simplified representation of the PV system to load connection
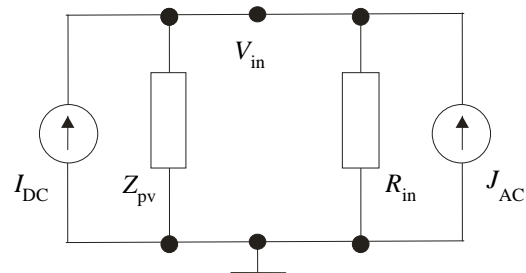
Figure 8. A simplified model of the PV system - converter interface

Since J_{AC} may be considered as a form of feed-back to the PV system, stability is becoming an important concern. Namely, a question arises as to whether the oscillations may persist and influence the quiescent working point of the PV system. To check for that an additional experiment was performed.

A new circuit was created consisting of the Ćuk converter [17] and a capacitor charged by 12 V. The capacitor is considered as a substitution to the constant voltage source. Relatively small capacitance was used to simulate the PV system output capacitance. As expected,

oscillatory discharge of the capacitor was observed as shown in Fig. 10. Having in mind the DC power comming from the PV system, in more realistic situations, one may expect sustained oscillations at high frequencies at the PV system to converter interface.
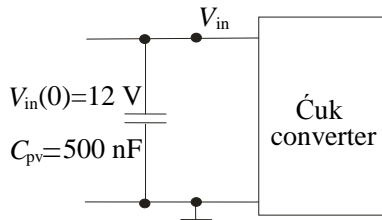
Figure 9. Experiment to check for oscillations
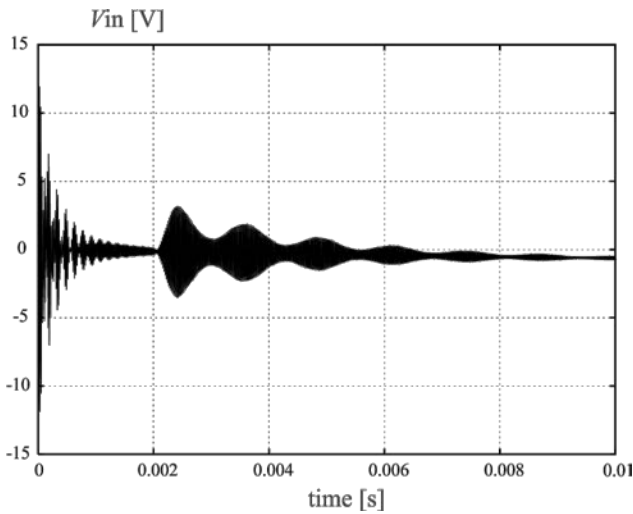
Figure 10. The input voltage as a function of time for the circuit of Fig. 9.

To avoid such oscillations a capacitor of large capacitance is usally inserted at the interface as shown in Fig.11.

To verify whether this is a solution to the oscillations problem we performed an additional simulation in which the PV system was modeled by a single DC current source as shown in Fig.12. The simulation results are shown in Fig. 13. As can be seen, alternating current of large amplitude may arise at the input of the converter. That current, in the circuit of Fig. 12, flows through the capacitor only since here no realistic model of the PV system was implemented. That however does not affect the conclusion that the capacitor is not a solution to the oscillation problem at the input of the converter. Namely, even in this case, the capacitor voltage still contains a significant AC component (about 2Vpp) that is driving backwards to the PV system. One is to add to these considerations the fact that real electrolytic capacitors of large capacitances suffer of relatively large series resistance and inductance which may becomes dominant at high frequencies i.e. at the frequencies of the harmonics depicted in Fig. 6b.

V. COMMENTS ON THE PV MODELING

Summarizing the analysis of the PV system to converter interface we may draw a general conclusion that in any case oscillations will remain at the interface. The amplitude of the time varying voltage at the output of the PV system will depend on several factors such as the switching frequency of the converter, the type and structure of the converter, the capacitance value, quality, age, and temperature of the electrolytic line capacitor, and the output capacitance of the PV system.
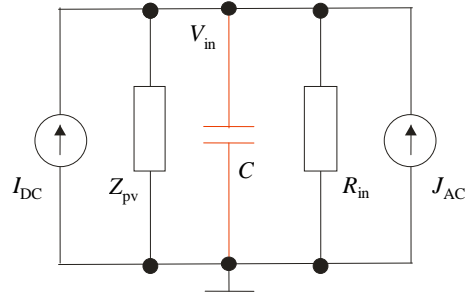
Figure 11. Model of the interface with a capacitor inserted
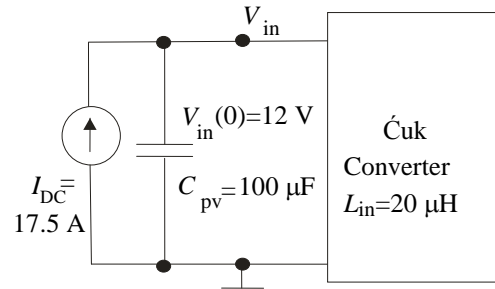
Figure 12. Simplified model of the PV system used to excite the Ćuk converter with large capacitance in between
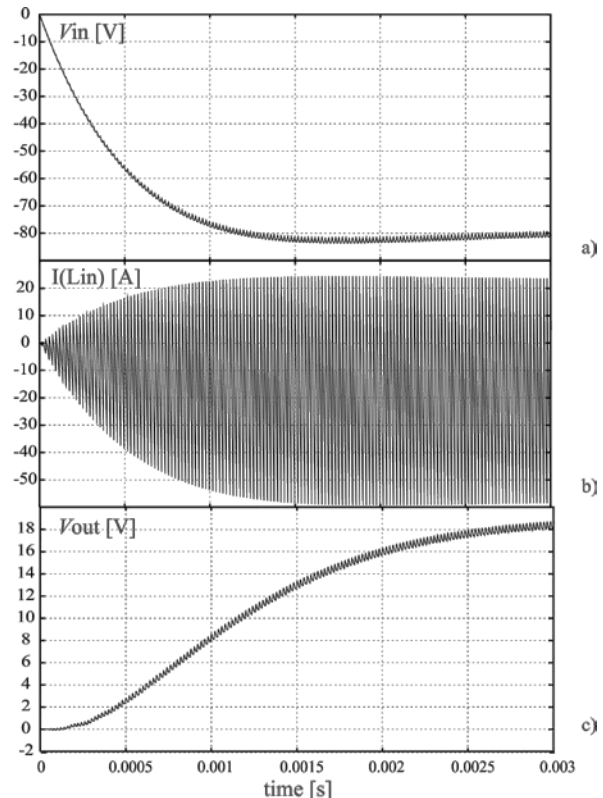
Figure 13. Simulation results for the circuit of Fig. 12. a) input voltage, b) input current and c) output voltage of the converter

The last claim is not as obvious as other ones. Namely, one may expect that the output capacitance of the PV system being equivalent of the PV cell`s junction capacitance, is much smaller than the line capacitance so being of no influence. That however is to be taken with caution since in between these two capacitances we meet several circuit elements such as the (nonlinear) diode, the shunt and series resistance of the PV cell and the parasitic elements of the line capacitor.

In any case the alternating component of the PV system output voltage is distributed downwards to all cells affecting all quiescent working points. If one wants to get realistic picture of the working condition of the PV cell one needs to take into account that component which implicitly means that one needs to use a model of the PV cell that exhibits dynamic behavior like the one of Fig. 3.

VI. CONCLUSION

The state-of-the-art in modeling PV cells was investigated. Properties of the existing models and simulation concepts were established. It was concluded that, based on the presumption that the output of the PV system may be characterized as a DC circuit, no dynamic simulations were performed and reported in the literature.

After a set of simulations on simplified models of the PV system to converter interface a conclusion was drawn that the output voltage of the PV system that is driving the converter contains a significant time varying component that is due to the switching in the converter. That alternating component is reduced but not suppressed by the line capacitor.

Accordingly, the main result of these investigations was the conclusion that one needs a realistic dynamic model of the PV cell in order to establish knowledge on its real working conditions. Simulations based on such a model including more realistic model of the electrolytic capacitor will shed clearer light on the working conditions of the PV system and the PV cell itself.

ACKNOWLEDGEMENT

REFERENCES

[1] Andrejević Stošović, M., and Litovski, V., „Modeling and circuit simulation of photovoltaic cells – an overview", 7th Int. Symp. Nikola Tesla, Belgrade, Serbia, Nov. 2011, pp. 83-92. ISBN 987-86-7466-420-9.

[2] G. H. Yordanov, O.-M. Midtgård, "Physically-consistent Parameterization in the Modeling of Solar Photovoltaic Devices," *PowerTech, 2011*, Trondheim, pp. 1-4.

[3] K. Leban, E. Ritchie, "Selecting the Accurate Solar Panel Simulation Model," NORPIE/2008, Nordic Workshop on Power and Industrial Electronics, June 9-11, 2008, pp. 1-7.

[4] H. I. Cho, S. M. Yeo, C. H. Kim, V. Terzija, Z. M. Radojević, "A Steady-State Model of the Photovoltaic System in EMTP," Int. Conf. on Power Systems Transients (IPST2009) in Kyoto, Japan, June 3-6, 2009.

[5] S. Liu, R. A. Dougal, "Dynamic Multiphysics Model for Solar Array," University of South Carolina, Faculty Publications, 2002.

[6] H. –L. Tsai, C.-S. Tu, and Y.-J. Su, "Development of Generalized Photovoltaic Model Using MATLAB/SIMULINK," Proceedings of the World Congress on Engineering and Computer Science 2008, WCECS 2008, October 22 - 24, 2008, San Francisco, USA.

[7] M. Azab, "Improved Circuit Model of Photovoltaic Array," Int. J. of Electrical Power and Energy Systems Engineering 2:3, 2009, pp. 185-188.

[8] *IV and CV Characterizations of Solar/Photovoltaic Cells Using the B1500A*, Agilent Techologies, Application Note B1500A-14, 2009.

[9] P. Maffezzoni, D. D'Amore, "Compact Electrothermal Macromodeling of Photovoltaic Modules," *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 56, No. 2, Feb. 2009, pp. 162-166.

[10] T. O. Saetre, O.-M. Midtgård, G. H. Yordanov, "A new analytical solar cell I-V curve model," *Renewable Energy*, Vol. 36, 2011, pp. 2171-2176.

[11] L. Zhang, Y. F. Bai, "Genetic algorithm-trained radial basis function neural networks for modelling photovoltaic panels," *Engineering Applications of Artificial Intelligence*, Vol. 18, No. 7, Oct. 2005, pp. 833-844.

[12] M. AbdulHadi, A. M. Al-Ibrahim, G. S. Virk, "Neuro-fuzzy-based solar cell model," *IEEE Transactions on Energy Conversion*, Vol. 19, No. 3, Sep. 2004, pp. 619-624.

[13] S.I Sulaiman, T.K Abdul Rahman, and I. Musirin, "Partial Evolutionary ANN for Output Prediction of a Grid-Connected Photovoltaic System," *International Journal of Computer and Electrical Engineering*, Vol. 1, No. 1, April 2009, pp. 40-45.

[14] S.I Sulaiman, T.K Abdul Rahman, I. Musirin and S. Shaari, "Performance Analysis of Evolutionary ANN for Output Prediction of a Grid-Connected Photovoltaic System," World Academy of Science, Engineering and Technology, Vol. 53, 2009, pp. 1023-1029.

[15] V. Litovski, M. Zwolinski, "VLSI circuit simulation and optimazation," Chapman and Hall, 1995.

[16] -, Photovoltaic Technologies for the 21st Century, Report of the Steering Committee for Advancing Solar Photovoltaic Technologies, NIST Published: December 2010, www.ieeeusa.org/communications/ebooks/govdocs

[17] Ćuk, S., Middlebrook, R. D., "A General Unified Approach to Modelling Switching-Converter Power Stages", *Proceedings of the IEEE Power Electronics Specialists Conference,* Cleveland, OH., June 8, 1976, pp.73–86.

Pspice Analysis of Parallel Operation of Two IGBT Inverters

Miroslav Lazić, Boris Šašić, Dragana Petrović and Dragan Stajić

Abstract — Two full-bridge inverters are connected in parallel in order to increase power of a programmable AC source. Lossless current sharing by adding balancing inductors was investigated. Effects of IGBT parameter tolerances and temperature variations were analysed through ORCAD 9.2 PSPICE (including Monte Carlo analysis). It was found that manufacturing tolerances of balancing inductors have greater effect on the current sharing than IGBT parameter variations. Adding the inductors in series with both power rails, positive and negative, of the bridges reduces required inductance and improves current sharing. Results of the analysis will be used to build an experimental circuit.

Keywords — Programmable AC Source, Current Sharing, Inverter Bridge.

I. INTRODUCTION

Current Sharing Analysis of Parallel operation of two H-Bridges in ADC Accessory are represented in this paper. Main goal of this analysis was to improved manufacturing efficiency by increasing power of the power electronics equipment (used in the magnetron sputtering applications for thin film deposition of semiconducting materials).

II. REQUIREMENTS

The existing system consists of a standard high power DC source followed by a full bridge inverter. The inverter is capable of generating unipolar and bipolar pulses of various frequencies and duty cycles. An example of unipolar and bipolar outputs is shown in Fig. 1.
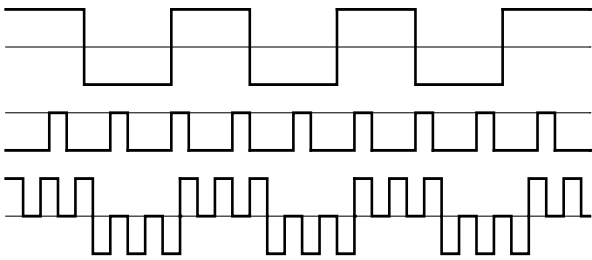
Fig. 1. Examples of output waveforms

Output power can exceed 10kW, with very wide, process dictated, ranges of output voltage and current – up to 1700V and 300A. Frequencies of interest are in a very wide range of 50Hz to 25kHz. A dedicated digital circuit is used to control the output in response to the system requirements.

In order to increase power, while avoiding major redesign and introducing risk to the established manufacturing process, it was decided to parallel two inverter stages. Requirement for current sharing was determined to be within 15%, not a very strict requirement. Due to the complexities of digital control

and utilized digital feedback loop compensation, standard current sharing schemes were not deemed practical, due to the implications on project timing and risk assessment. It was decided that a simpler approach is analyzed – a possibility of driving both inverters with identical drive signals and adding series inductors in line with the two paralleled inverters. Sensitivity analysis to temperature changes (affecting IGBT parameters), inductance values and tolerances and, finally, sensitivity to combined effects was the critical part of the project.

III. SIMULATION MODEL

Initial circuit, as modeled in PSpice is shown in Fig. 2. The model includes some of the relevant parasitic elements, values of which were estimated based on the existing inverters.
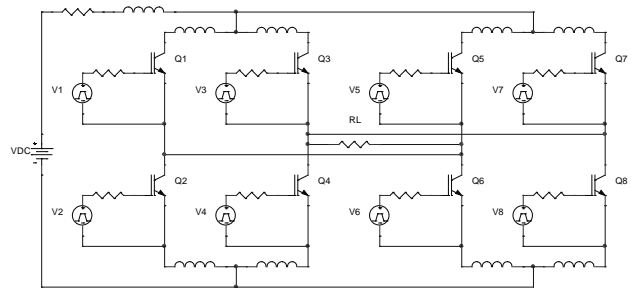
Fig.2. Initial simulation model

IGBT parameters suitable for PSpice simulation model are listed in Table I. The two bridges use IGBT modules SEMIKRON SKM400GB176D.The modules are built based on INFINEON part number SIGC186T170R3. After longer research and communication with INFINEON's applications engineers it was found that adequate replacement, for which simulation models are available is EUPEC's (formerly SIEMENS and now acquired by INFINEON) part number BSM150GB100D.

After running Monte Carlo analysis and varying for IGBT parameters and operating temperatures, the obtained results were expectedly poor, as shown in Table II. For brevity, results for only two switches, in identical positions, are shown. Simulated load current is 200A.

As we are using computer-generated random numbers for the analysis (the random number seed), it is important to note that, in reality, these are pseudorandom numbers, due to the deterministic nature of the computers. If the seed number is repeated, identical random numbers will be repeated as well. For multiple trials, different random number seeds were used, as presented in the Table II.

The Table II lists only the worst case current through the switch for each set of simulations. Highlighted are the worst case deviations. Ideally, current through each IGBT would have been 100A – large deviations from the ideal

number prove the need for forced current sharing.

TABLE I.
IGBT MODEL PARAMETERS

Model Parameters	Symbol	Value
AREA (area of the device)	A	1.858 cm2
AGD (gate-drain overlap area)	A_{GD}	1.4823 cm2
KP (MOS transconductance)	K_p	3.54 A/V2
KF (triode region factor)	K_f	
CGS (gate-source capacitance per unit area)	C_{GS}	10.7 nF/cm2
COXD (gate-drain oxide capacitance per unit area)	C_{OXD}	59.3 nF/cm2
VT (threshold voltage)	V_T	5.8
TAU (ambipolar recombination lifetime)	T	8×10^{-6} cm
WB (metallurgical base width)	W_B	36.7385×10^{-3} cm
NB (base doping)	N_B	0.1651×10^{14}/cm^3

TABLE II
CURRENT SHARING OF THE ORIGINAL CIRCUIT

No.	Seed number	ICQ2 (A)		ICQ6 (A)	
		min.	max.	min.	max.
1	default	86.5	130.1	68.7	109.1
2	100	109.5	122.7	76.2	89.5
3	1000	92.6	126.4	72.5	101.1

Figure 3 shows four current sharing inductors added to each leg of both paralleled inverters. The results for added 10µH inductance are summarized in Table III. The inductance was selected based on the excellent results (5% deviations) when temperature effects on IGBT parameters are neglected.
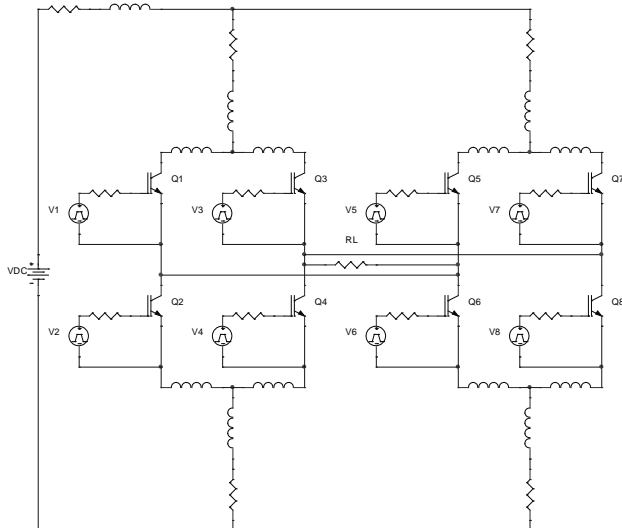
Fig. 3. Model with current sharing inductors

TABLE III
CURRENT SHARING WITH 10µH BALANCING INDUCTORS

No.	Seed number	ICQ2 (A)		ICQ6 (A)	
		min.	max.	min.	max.
1	default	102.6	126.5	88.4	96.2
2	10000	102.3	120.6	78.1	96.5
3	30000	107.7	124.7	74.1	91.2

The results indicate that, when IGBT parameters are taken into account, achieved results are modest, at best.

IV. VARIATIONS OF KEY PARAMETERS

It is obvious that seed number also plays significant role in the final outcome. Of course, this is little related to the actual operation of the circuit and warrants a closer look. Table IV summarizes I_{CQ2}-I_{CQ6} values for different seed numbers.

TABLE IV
DEPENDENCE OF CURRENT IMBALANCE ON SELECTED SEED NUMBER

Seed number	I_{CQ2}-I_{CQ6} (A), max.	
	No inductors	10uH inductors,
default	61.24	22.14
10	41.91	28.08
100	46.57	33.78
1000	31.6	27.05
10050	48.18	**36.50**

It is interesting to note that selected seed number has much larger effect on the circuit without balancing inductors then on the one with 10uH inductors. Relative to the load current of 200A, worst case imbalance is 36.5A, or 18.2%.
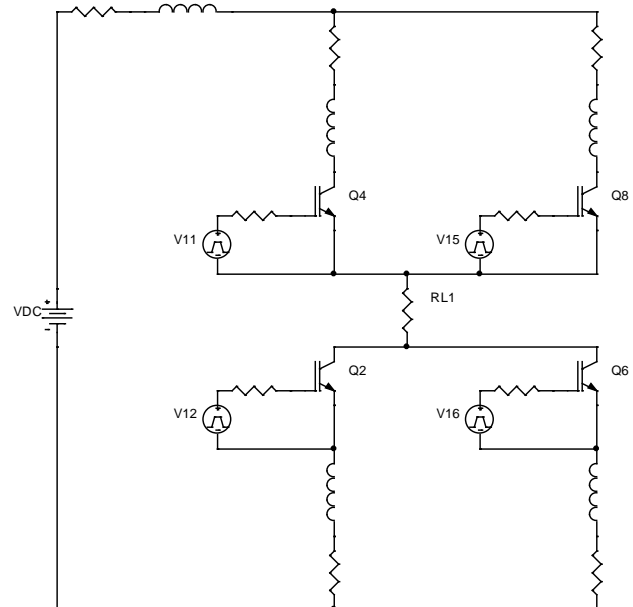
Fig. 4. Simplified circuit

So far, adding balancing inductors to decouple the two inverter bridges did not result in significant improvements. Further improvements were investigated. It was found that adding balancing inductors in series with emitters of the IGBTs with grounded emitters (low

side switches Q2, Q6, Q3 and Q7) has significant effects. Rather than running the initial model (Fig. 1) and change inductor values, the schematic was simplified in order to minimize convergence errors. Only IGBTs connected in parallel are shown. The simplified circuit is shown in Fig. 4 and summary results in Table V.

TABLE V
CURRENT SHARING WITH BALANCING EMITTER INDUCTORS

Added inductance	ICQ2 (A)	ICQ6 (A)
1uH	111.4	87.5
5uH	108.4	90.5
10uH	**105.6**	**93.2**
20uH	103.2	95.8

The table outlines worst case results of several simulations with different inductance values and seed numbers. Upon examination of results, and comparison with those shown in Tables III and IV, improvements are significant, deviation from ideal current sharing is less than 7%. It is interesting to note that adding the same inductor values into the collector circuits does not yield any improvements.

One example of simulation resulsts are given in Figure 5. econg example can be seen in Figure 6 (illustrates simulation results ICQ2-ICQ6 for L3=L4=L5=L6=10uH and Iload=50A).
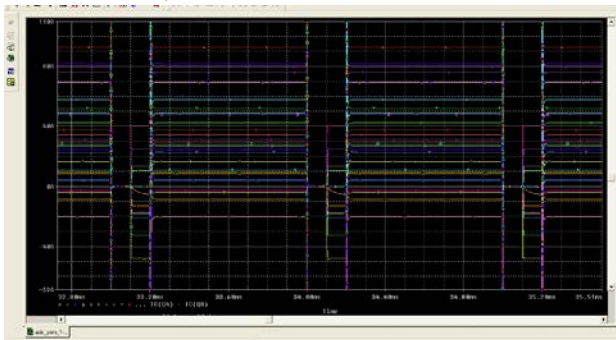
Figure 5: Differences in collector currents between Q4 and Q8
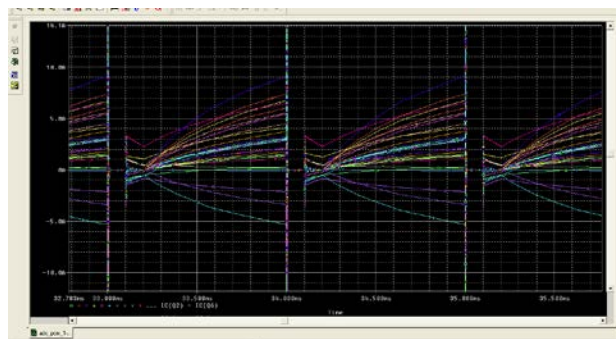
Figure 6: Differences in collector currents between Q2 and Q6

V. INDUCTOR TOLERANCES

The paper further discusses analyses for various duty cycles and load currents, finding and addressing the worst case scenario. Final inductance value of 70uH is identified as acceptable, and analyzed for the manufacturing tolerances.

Table VI summarizes worst case analysis, combining effects of IGBT parameter variations and inductor tolerances.

TABLE VI
CURRENT SHARING RELATIVE TO 70uH INDUCTOR TOLERANCES

	20%	10%	5%	3%	1%
I_{CQ2avg}/ I_{CQ6avg}	1.45	1.22	1.13	1.11	1.08
I_{CQ4avg}/ I_{CQ8avg}	1.33	1.17	1.12	1.09	1.07

Here we are looking at ratio of currents through IGBT's in identical positions. 5% tolerance allows meeting the requirement for current sharing within 15%.

VI. PRACTICAL OBSERVATIONS

Based on the previous analyses the following observations can be made:

1. The very basic current balancing scheme with one balancing inductor per H-bridge was evaluated against somewhat more complex scheme using two inductors per bridge, one in the upper leg and one in the lower leg of each paralleled bridge. The latter was proved to be more effective and results in significantly smaller required inductance

2. Variations of IGBT parameters have much lesser effects than balancing inductor tolerances.

3. Tolerances of the balancing inductors should be within ±5% (from each other) in order to ensure current sharing within 15%.

4. It is desired to have IGBT parameters vary within 10% (DEV=10% and LOT=5%). This combined with inductor tolerances of DEV=10% results in adequate current sharing. It proved difficult to obtain exact parameter distributions from component vendors and it may be impossible to establish any type of control over the parameters, however, manufacturer's applications engineers feel that normal distribution falls well within the desired tolerances.

5. Inductor values of 70uH are adequate for switching frequencies of fs≥1kHz. For low switching frequencies inductor values need to increase (inversely proportional increase seems like a reasonable approximation).

6. It is important to reiterate that

a. Simulations did not take into account positive temperature coefficient of chosen IGBTs. It was assumed that there is a fixed difference in Vces due to temperature differences: this is the worst case and in practical circuit Vces of the two IGBTs operating at different temperatures will tend to drift toward each other minimizing the difference and improving current sharing.

b. Snubber circuits were not simulated due to convergence problems. Resulting reduction in current rise time (di/dt) is neglected, which again results in somewhat worse current sharing than could be expected in a practical circuit.

c. Analyses were performed with duty cycle of D=0.8 and D=0.1. Presented results are given for the worse of the two cases (constant D=0.8), which again may lead to somewhat exaggerated imbalances.

7. In summary, the proposed current balancing scheme uses two inductors, one in the upper and one in the lower leg of the H-bridge. Target current sharing can be achieved by using reasonable inductance of 70uH, assuming IGBT parameter values variations of 10% or less and inductor tolerance of ≤5%.

VI. Conclusion

In order to achieve significant increase in output power of the high power programmable waveform AC source, without major redesign effort, it was decided to connect two inverter bridge circuit in parallel. Given sufficient power reserve, current sharing requirements were set at 15% over the range of output loads.

Computer simulation and Monte Carlo analysis was used to determine the worst case operating conditions and determine minimum required inductance value and tolerance.

The basic lossless current sharing scheme, by using one inductor per inverter bridge was compared with the concept using two inductors, one in positive and the other one in negative leg of each paralleled bridge. The latter was proved to be more effective requiring significantly smaller required inductance. Adding inductors in series with IGBT emitters further improves current sharing, albeit at added cost and complexity.

Variations of IGBT parameters have less effect on the current sharing than tolerances of added inductors. Tolerance of $\pm5\%$ is sufficient to allow current sharing within 15%.

An experimental circuit is currently being built for laboratory and field evaluations.

References

[1] Chibante, R., Araújo, A., and Carvalho, A., *"A Simple and Efficient Parameter Extraction Procedure for Physics Based IGBT Models",* Proceedings of 11th International Power Electronics and Motion Control Conference (EPE-PEMC'04). Riga, Latvia 2004.

[2] Protiwa, F.-F., Apeldoom, O., and Groos, N., *"New IGBT Model For Pspice",* Proc. of the Fifth European Conference on Power Electronics and Applications, September 1993, Brighton, UK, Vol. 2, pp. 226-231.

[3] -, Semikron, *"SKM400GB176D",* Datasheet

[4] -, Infenion, *"SIGC186T170R3",* Datasheet

[5] -, Eupec, *"BSM150GB100D",* Datasheet

[6] -, Cadence, *"Pspice Manual",* Orcad 9.2

Simulation as the optimization tools for the complex logistic systems

(business, technical, IT and control systems)

Milosav Georgijević, Vladimir Bojanić, Goran Bojanić, and Sanja Bojić

Abstract - If we conclude that the simulations are the most modern tools for optimization, it is question for real multidisciplinary problems: what kind of simulations can describe the problem?

Since logistic processes are part of larger business processes that satisfy certain society demands, in every process of planning or reengineering the ideas must be searched for at the higher level of occurrences, at the level of the stock market. At this level the simulations are applicable in the social-economic domain.

At the large corporation level (LSS - Large Scale Systems) the simulations are applicable at the level of business processes. Results of these simulations can be used to derive the project tasks for technical systems, which is subject of this paper. The first step is to analyse the material flow and logistic goodness parameter of the system that is in the planning or reengineering phase.

This analysis is the basic framework of project requirements for designing of mechanization and equipment, control systems, management of manufacture and material flows. After each level, by means of feedback, the information about how practical the whole idea and if there is a need for corrections at the previous level is received. Application of different simulation tools at the abovementioned levels will provide recommendations for further development.

Keywords - Simulation, Logistic, Optimization.

I. INTRODUCTION

Motto: A picture speaks more than 1,000 words; one model shows more than 1,000 pictures [6].

Until about 15 years ago, it was arguable to talk about simulations as an optimization method, because strictly speaking there is no goal function which streams to a certain value (max. or min.), which could give optimal values of one or a countable small number of parameters. Requirement to analyze more influential parameters or

Milosav Georgijević (Professor) University of Novi Sad, Faculty of Technical Sciences, E-mail: georgije@uns.ac.rs

Vladimir Bojanić (Assistant) University of Novi Sad, Faculty of Technical Sciences, E-mail: ydad@uns.ac.rs

Goran Bojanić (Researcher on project Ministry of Education and Science) University of Novi Sad, Faculty of Technical Sciences, E-mail: gbojanic@uns.ac.rs

Sanja Bojić (Researcher on project Ministry of Education and Science) University of Novi Sad, Faculty of Technical Sciences, E-mail: s_bojic@uns.ac.rs

process models and systems that are closer to reality gives an insolvable system of equations (aim functions) and even since before last millennium, simulations have been accepted as an optimization method even with the strict authorities. Today, simulations are not just analyses of system operation and processes in the time and time related domain (for example. frequency analysis, etc.), but multiple repetitions as well, e.g. counting FEM analyses as simulations.

II. FROM GLOBAL TO LOCAL (SUBJECT AND AIM)

If we accept globalization as a process, simulations from the social-economic level support decisions that are made at the company level, and they are related to the business as well as technical system, which is the subject of this work (Fig. 1).

The aim of this paper is to point out the connection of simulations in the frame of the technical system, valuation of capability of the modern software, and especially the unsolved problems in their application.

From the commanded information level or project assignment are coming down on lower levels [8]. In technical systems there are material (goods) flows, which are the input (and the output) in/out of the business-technical system where simulations are applicable in design of logistic systems.

Parts of these systems are machines (and equipment) with information and control systems which should be projected exactly in accordance with application demands, where simulations are applied again in the designing process. As a reversible connection in all cases, hierarchically higher levels are getting affirmation of demands which are dropped to lower levels, or notification that with unfounded (extreme) demands it is not going to be possible to get a qualitative or economically lucrative solution.

In the real conditions, the decision and construction and control of machines and equipment is followed by simulations of: production, montage, functioning of new technical system or, with products of mass application of market simulation, distribution and sale etc.
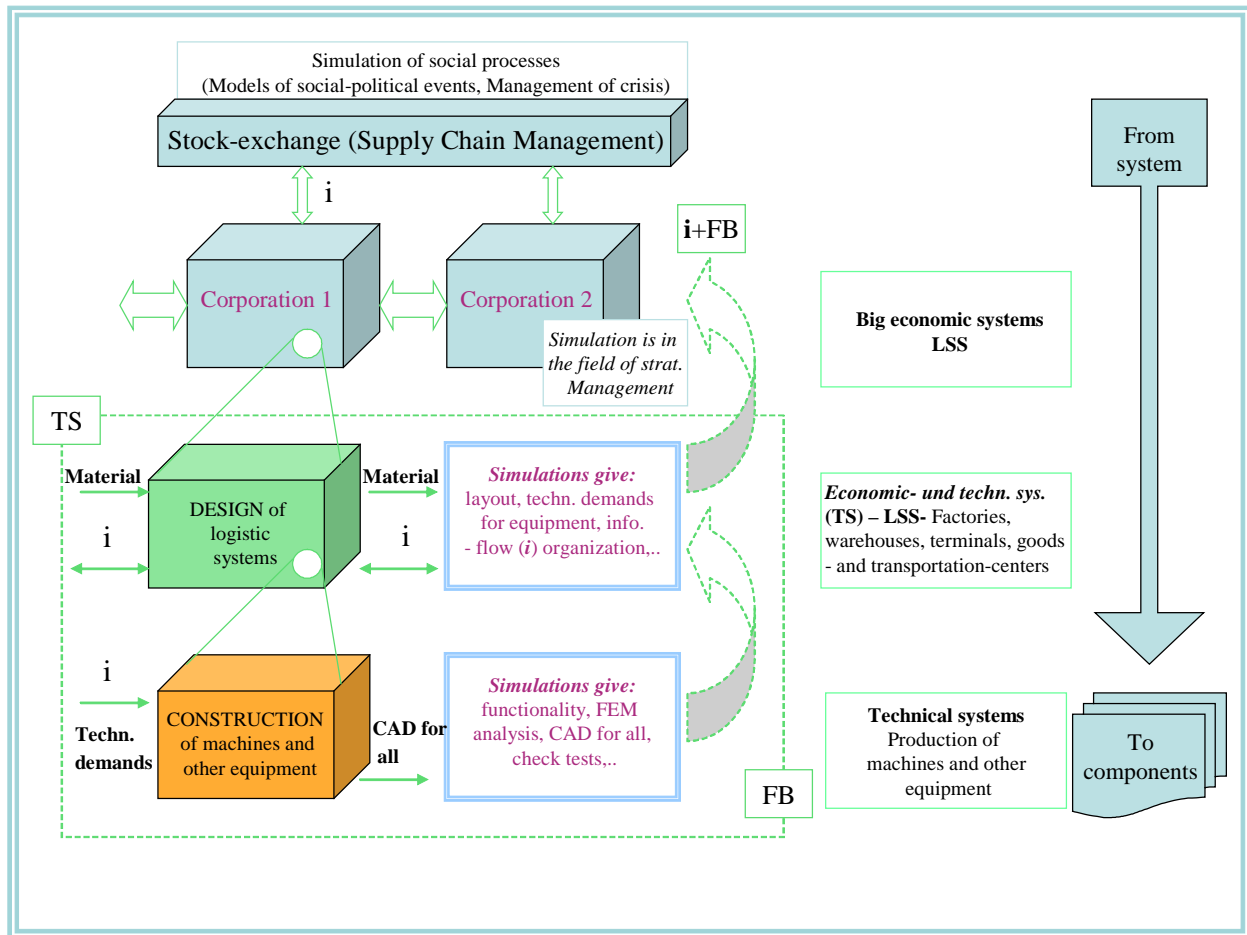
Fig. 1. Hierarchical levels of simulations (TS-Technical systems, i-information, FB-feed back, CAD-Computer Aided Design, FEM-Finite Element Method)

Since modelling is the first step in the real system analysis, Fig. 2 gives [7] key points of simulation studies, where each individually requires special research that starts from the technology of some process, followed by data analyses and the process (or machine) model, all of which need validation (comparison with previous experiences, comparison with experiments on real systems etc.) in order to be possible to approach the research - optimization of the process (or machine) through simulation experiments. Proper interpretation calls for validation of all previous operations and demands new rounds of simulation experiments until the satisfying solution is achieved in the optimization process

A. Application –Location problem and its simulation

As a consequence of permanent increase in cargo flows, there is a trend of constructing logistic canters to reduce transportation time and costs and to improve customer service.

However, costs associated with the projected technical, management and IT systems, maintenance and operation of the logistic canters increase as more of them are built. Thus determination of optimal number of logistic canters and their locations in a region can significantly contribute to the savings of transportation and storage costs, maintenance and, at the same time, high level of customer service.

Location problems consider locating a set of new facilities in a way that the transportation cost from the facilities to customers is minimized

Having in mind that a model should, as precise as possible, represent the real problem, the location model of a logistic canter should be defined in accordance with the existing transport network, demand for logistic services - both in terms of quantity and location, existing infrastructural resources and cost structure of entire SC. Therefore we propose models that could be classified in a group of capacitated, network, fixed costs, location – allocation models to determine an optimal number, location and capacity of logistic canters.
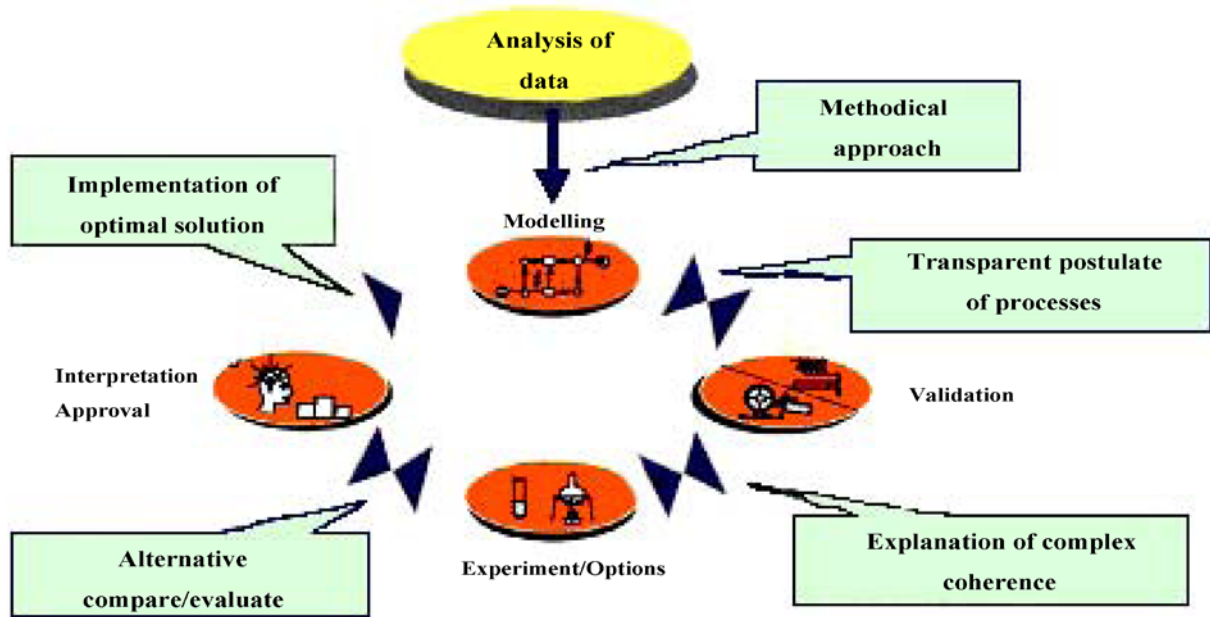
Fig. 2. Steps in a simulation's study

Based on the following inputs: i – index for supplying nodes; j – index for potential logistic centers; k – index for allocated demand nodes; w_{ij} – demand of customers in node j from the supplying node i, w_{ik} – demand of customers in node k from the supplying node i; s_j – capacity of potential logistic centres j; c_j – costs for opening a center in node j per unit of capacity; f_j – capacity costs per unit of product in center j; d_{ij} – distance between nodes i and j; d_{jk} – distance between center j and node k; p – number of considered centers to be opened; Y_{jj} - binary decision variable {1 if center is opened at node j, 0 otherwise}, Y_{jk} - binary decision variable {1 if node k is allocated to opened center j, otherwise 0} the logistic centers location problem can be formulated as follows:

$$\min F = \sum_{i \in I} \sum_{j \in J} d_{ij} w_{ij} Y_{jj} +$$
$$+ \sum_{j \in J} \left(c_j s_j + f_j \sum_{i \in I} w_{ij} \right) Y_{jj} +$$
$$+ \sum_{i \in I} \sum_{j,k \in J} (d_{ij} + d_{jk} + f_j) w_{ik} Y_{jk}$$

subject to:

$$\sum_{j \in J} Y_{jj} = p \tag{1}$$

$$Y_{jk} - Y_{jj} \le 0, \qquad \forall j, k \in J, \tag{2}$$

$$\sum_{i \in I} w_{ij} Y_{jj} +$$
$$+ \sum_{i \in I} \sum_{k \in J, k \ne j} w_{ik} Y_{jk} \le s_j, \qquad \forall j \in J, \tag{3}$$

$$Y_{jk} \in \{0,1\}, \qquad \forall j, k \in J, \tag{4}$$

$$Y_{jk} \in \{0,1\} \qquad \forall j, k \in J, \tag{5}$$

$$w_{ij}, d_{ij}, c_j, s_j \ge 0$$

The objective function minimizes total distribution costs of cargo flows between suppliers, logistic canters and demand nodes. The first part of the objective function considers the aggregated transportation costs from supplying nodes to the logistic canters, the second one considers the storage costs, and the last one the distribution costs from opened public logistic canters to remaining demand nodes. Constraint (1) defines that the number of opened canters equals p and constraint (2) that node k cannot be served from a non-opened canter j. Constraint (3) is a capacity constraint assuring that total demand at node j and demand of all associated nodes does not exceed the capacity of canter j. Constraints (4) and (5) are integral constraints.

Modelling the capacitated network location models with a fixed cost approach can be a computationally very difficult combinatorial optimization problem. This was the reason for intensive investigation of possibilities of heuristic solution procedures which run in reasonable
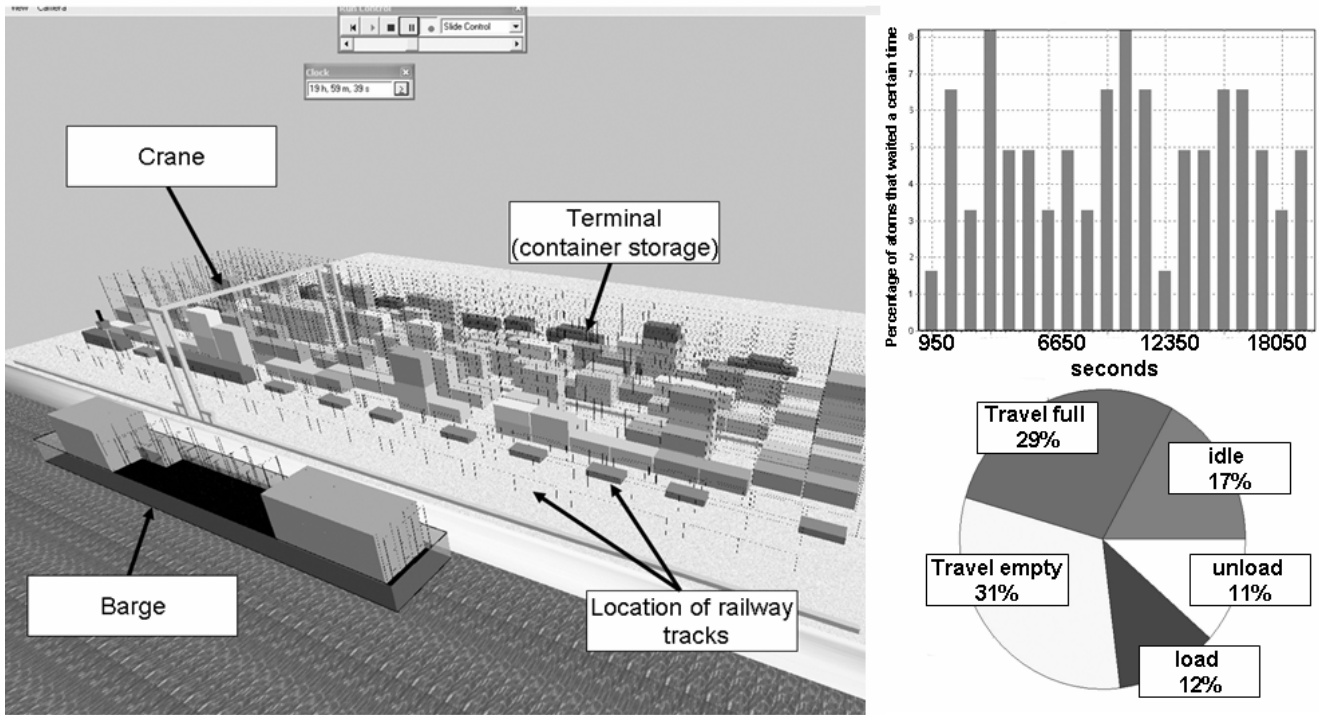
Fig. 3.Simulation of the model container terminal and output results

computer time and yield solutions of acceptable quality. Having in mind the complexity of the model we suggest application of genetic algorithms for solving the location problem. Given model is used for analysis of logistic canter locations in Serbia.

B. Application - System simulation

After determination of optimal location and basic parameters of the logistic canter, next step requires detailed analysis and design of that logistic canter with the simulation of flows of information for the given conditions and concepts of IT systems. The first step is designing the system with simulation of material (or goods) flows, where working places are black boxes with setup and processing time, while transports and storages are the principal process parameters.

Inland container terminals are local logistic hubs for container transport to/from the costumer/manufacturer. A container terminal represents a complex system with highly dynamic interactions between the various handling, transportation and storage units, and incomplete knowledge about future events. There are many decision making problems related to logistics planning and control issues of and they can be assigned to three different levels: terminal design, operative planning, and real-time control [8]. There are a large number of scientific papers that are dealing with these problems. Most of them focus on optimization of specific processes only. But optimization of each element of the larger system does not always lead to the optimization of the whole system. Nowadays, for design of

such a system, simulations are commonly used tool. By creating the simulation model, it is possible to analyze the whole system with different stochastic input and system parameters in the real time.

Here is presented an example of simulation application for determining the concept for the container terminal in river port. Performances of the models were studied under a different set of variable values (input, output and system parameters). Modelling and experimentation is done in software Enterprise Dynamics 7.

Several concepts were examined with different length of container terminal (100, 150 and 200m) and different handling mechanization (cranes, straddle carriers, reach stackers) and the level of automation in several different tasks (combination of reloading of the barges, trains and trucks). As the result, truck waiting times, container throughput, cycle times and reloading times were obtained. This data is basis for determination of the terminal design, as layout, IT and management system, capacity, where port operator has the final word. And this data is also an input data for the design of the reloading mechanization inside the terminal with its control systems.

C. Application - Construction of handling and storing equipment

Based on defined location problem as well as the design of the logistic centre, requirements for mechanization intended for projected logistic canter operation are obtained. Construction and optimization of handling and storing equipment are possible to achieve with the help of high value software tools for complex system static and

dynamic analysis, such as KRASTA (KRAnSTAtik) and ADAMS (Automatic Dynamic Analysis of Mechanical Systems). Based on models developed in these software packages it is possible to obtain the necessary data needed for construction calculation and machine optimization, such as: forces, stress, velocity, acceleration, movement, all required in control systems as well as many other parameters.

For Control systems and database for control in the container cranes in this paper is presented the influence of kinematics of ropes upon the change of oscillating (sway) periods of container and this can be the base for choice and interaction of the control system for driving systems of trolleys and crane. If the whole crane is modelled, then on the basis of known working cycles it is possible to make database for control system. By varying of various expected masses of containers and various working cycles the concrete values are reached of electric power parameters of the driving engines of trolleys and crane, which also generate dynamic parameters and spectra of loads.

In the Figure 4 is given the model of crane of the river container terminal with following technical parameters: Span – 25 m; Outreach on both sides – 27.5 m; Lifting height – 25 m; Approximate weight – 290 t (imputing trolley of 50 t mass); Moving velocity of trolley - 2 m/s; Moving velocity of crane – 2 m/s; Lifting velocity – 0.5m/s.
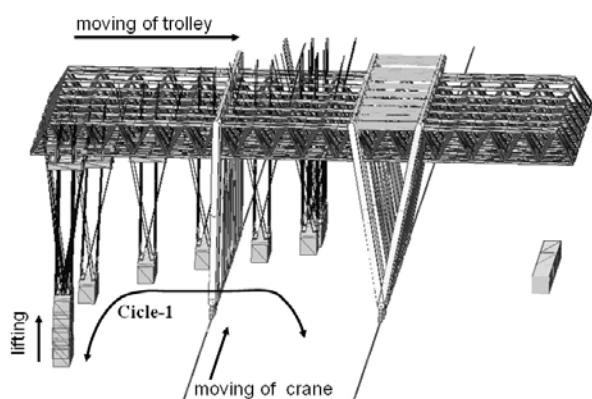
Fig. 4.Simulations of different cycles for the purpose of creation of database for control system

In Figure 5 is applied adaptive PID control for one working cycle. In the same way could also be done simulation for other cycles of work and with other masses of containers with an aim to obtain database for the control system. Due to the large number of influential parameters that are of stochastic nature, but more than 20 years individually in control systems fuzzy logic is introduced, because only these control methods can take into account the wind effect on the swing - the positioning of the container and the impact position and rocking of the ship in relation to coast [11].

After defining of the control system, by means of simu-

lations which are subject of this paper also are obtained spectra of loads for all parts of the construction for the foreseen working cycles, which result in corresponding dynamics (dynamic coefficients), which are the base for calculation of construction according to life time (fatigue).

III. CONCLUSION

Simulations as the most modern tool for optimization have plenty of possibilities, but a long-term problem remains in how to connect a number of different types of simulations that are follow one idea (problem). It is demonstrated with one example for problems in material flow logistics, starting from the location and concept of container terminals to machines operating in it, what is the basis for techno-economic analysis and what is the impact of the feedback of socio-economic analysis preceding the technical system. This approach leads to optimization of time that is spend on the developing process of a product [4], and can be named system projecting logistics.

For connection of more different simulations, required by the nature of the problem, interfaces that could transform the outgoing parameters from one simulation into incoming parameters for new following simulations are required. Further research should result in improvements of high value software, so that the large number of real cases from the practice can be included, which unavoidably includes simulations as a method for optimization - not just in the sense of technical performances, but also in the sense of rapid response to market demands for new systems, facilities or machines.

ACKNOWLEDGEMENT

REFERENCES

[1] Caricato, P., Grieco, A., *" Using Simulated Annealing to Design a Material-Handling System",* IEEE Intelligent Systems, July/August, Vol. 20, No. 4, 2005, pp. 26-30.

[2] Daskin, M., *"What you should know about location modelling",* Naval Research Logistics,Vol. 55, No 4, 2008, pp. 283–294.

[3] Drezner, Z., Hamacher, W., *"Facility Location: Applications and Theory"*, Springer, Berlin 2004.

[4] Gebhardt, A., *"Rapid prototyping",* Hanser Verlag, Muenchen, 2000.
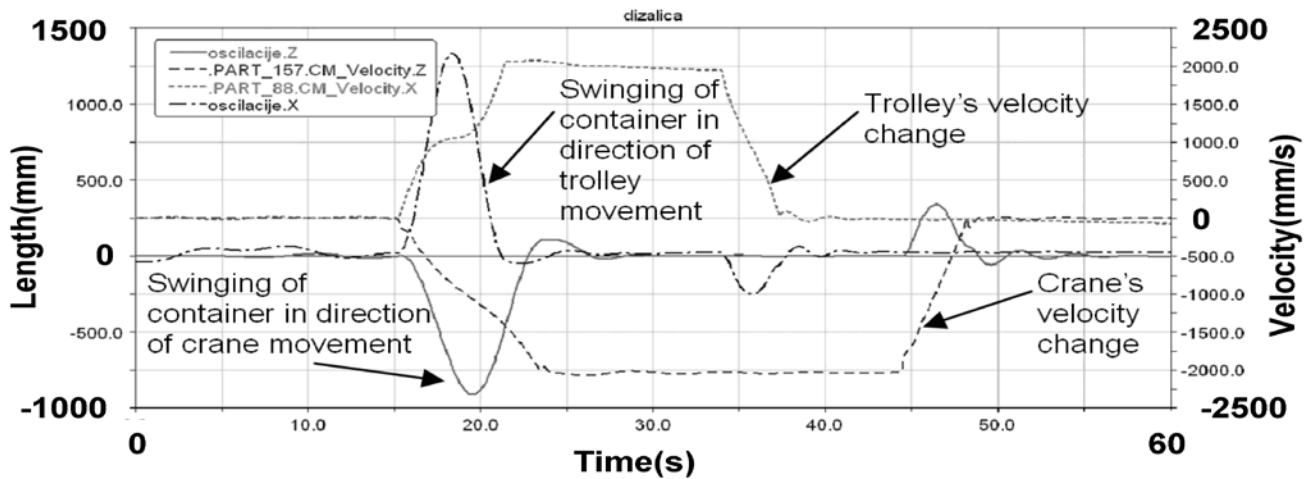
Fig. 5.Velocities and swinging of containers in relation to the trolley in real working cycles with lifting, moving of trolley and moving of crane portal (cycle in Figure 4)

[5] Georgijevic, M., Radanovic, R., *"Simulation komplexer Systeme und Optimierung"*, 9. Symposium Simulation als betriebliche Entscheidungshilfe 2004, Goetingen – Braunlage, pp. 307-321.

[6] He, F., Chen, Y., Zhao, S., *"Application of Fuzzy Control in the Stacker Crane of an AS/RS"*, FSKD '08 Proceedings of the 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery – Vol. 03, pp 508-512.

[7] Hellmann, A., Wloka, J., *"Mehrwert durch Simulation"*, Logistik fuer Unternehmen, Vol. 19, No.9 2005, pp.62-64.

[8] Kim, K.H., Günther, H.O., *"Container Terminals and Cargo Systems: Design, Operations Management, and Logistics Control Issues"*, Springer-Verlag, Berlin Heidelberg, 2007.

[9] Krame, U., *"Simulationstechnik"*, Hanser Verlag, München, 1998.

[10] Mertins, K., Rabe, M., Jaekel, F-W., *"Distributed modelling and simulation of supply chain"*, International Journal of Computer Integrated Manufacturing, Vol.18, No.5, 2005, pp. 342-349.

[11] Xu, W., Gu, W., Shen, A., Chu, J., Niu, W., *"Anti-swing control of a new container crane with fuzzy uncertainties compensation"*, IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2011, pp. 1648-1655.

Single Event Latchup Power Switch Cell Characterisation

Vladimir Petrovic, Marko Ilic, Gunter Schoof

Abstract - In this paper are described simulation and measurement processes of a power switch cell used for single event latchup protection of a digital fault tolerant application specific integrated circuit. The standard IHP 250 nm simulation models of components are used for the performed analog simulation using the Virtuoso$^{©}$ Cadence tools.

Keywords - Single event effects, fault-tolerance, power switch, ASIC design methodology

I. INTRODUCTION

The development of a design flow methodology for the fault tolerant application specific integrated circuits (ASIC), based on using dual modular redundancy [1] [2], has opened new questions about functional design verification before an ASIC is produced. As it is known, an ASIC can be designed as an analog chip, digital chip or as a mixed-signal chip. The new design methodology [1], is related to production of the fault tolerant ASIC and describes a way how a complex digital ASIC can be designed by using the standard tools in order to provide a digital system resistant to single event effects (SEE). The most known effects in aerospace microelectronics are: single event upsets (SEU), single event transients (SET) and single event latchups (SEL) [4]. The mentioned fault tolerant design methodology is based on using the dual modular redundancy (DMR) instead of using the standard triple modular redundancy (TMR) [3] [8]. The protection against SET and SEU can be done with known techniques on the system level. For the protection against SEL it was necessary to develop new power control circuit.

During the development process of the mentioned new design methodology [1] for a fault free ASIC, we proved that SEUs and SETs can be implemented as fault models into the VHDL code by using standard fault injection methods. The verification of the system functionality, related to the SEU and SET faults is done by the standard digital simulators. From another side, the SEL faults still need to be simulated in analog environment [4]. The basis of this work is verification of the system functionality when a SEL occurs in a digital ASIC and the SEL power switch characterization, in order to provide all needed

Vladimir Petrovic and Gunter Schoof are with the IHP, Im Technologiepark 25, 15236 Frankfurt Oder, Germany, e-mail: {petrovic, schoof}@ihp-microelectronics.com.
Marko Ilic is with IHP (PhD Student Intern), Klacevica, 35250 Paracin, Serbia, e-mail: markoilic2211@gmail.com.

information for the automated design process [1].

Test designs are implemented in the standard, non-radiation hard IHP's 250 nm CMOS technology [7].

The paper is organized through three sections:
1) Simulation environment of the SPS cell
2) SEL power switch cell description
3) Characterization of the SPS cell

II. SIMULATION ENVIRONMENT OF THE SPS CELL

During the development process of the SPS cell it was important to define an appropriate simulation environment, in order to provide the accurate functional verification. The simulation environment of the SPS cell consists of the three main parts:

a) Latchup generator
b) Control block
c) Digital block supplied by the SPS

In the Fig. 1 is represented the simulation environment, which is also used as a measurement environment for the SPS cell.
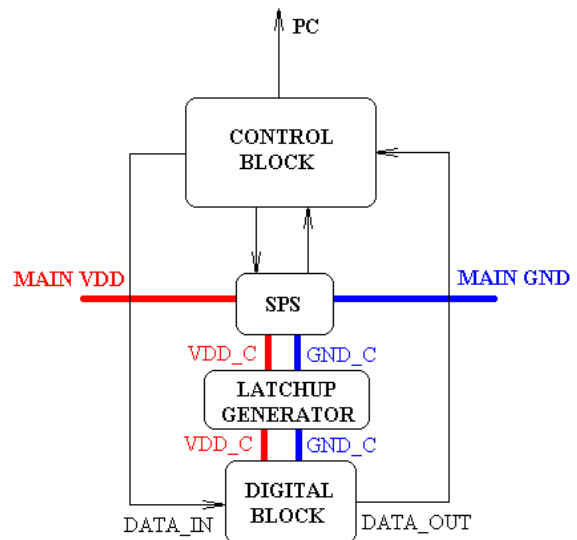
Fig. 1. Simulation environment of the SPS cell

The physical process which describes how a latchup is induced in a CMOS pair is used during the development of the latchup generator. As it is known [4] - [6], the latchup effect is based on a parasitic thyristor component, formed

in the CMOS pair. In order to save the time required for designing a technologically dependent thyristor, a simple switch controlled by voltage (VCSW) is used in the simulation process. This also provides easier hardware realisation of the latchup generator, required for measurements.

The control block is used to provide input control signal of the SPS cell, depended on the output signals generated by SPS cell. The outputs of the SPS are related to the latchup detection and current value of the controlled power supply provided by SPS, while the input signals are controlling the activation of the latchup protection (normal or latchup). The control block is also used for the digital block tests. It provides the stimulus for the digital block during the latchup test. The hardware realisation of the control system is based on the microcontroller system and signal generators.

The digital block, supplied by the SPS cell, is a simple DMR based digital system, which consists of the flip-flops (FF), multiplexers (MUX) and NAND gates. The digital block tests are used to provide information what is exactly happening with the data in the moment when the latchup effect occurs, during and after it. The digital block is implemented in the IHP 250 nm standard cells library.

The simulation process starts by defining the normal operational conditions for the SPS cell, in order to provide the power supply for the digital block. Next step in the simulation process is the digital block validation. The control block provides the data input, clocks and other control signals for the digital block. The control block verifies the correctness of the digital block using the "data out", generated by the digital block. When the digital block is verified, the latchup generator provides a shortcut between VDD_C and GND_C. The VDD_C is controlled voltage supply (VDD) by the SPS and GND_C is controlled ground (GND) line by the SPS (Fig. 1). After the shortcut, the control block should detect a latchup and provide the required control signals in order to test the functionality of the SPS in the latchup mode.

In the next sections of paper are provided the details of the SPS cell functionality, implementation and measurements required for the characterization process. Presented simulation environment is used also as the test environment of the implemented hardware.

III. SEL POWER SWITCH DESCRIPTION

The system protection against single event latchup effect requires design of a special power switch. The SEL power switch cells (SPS) are used in the unity with latchup sensors and power switches for the logic where a latchup was detected. The SPS is a new standard power cell of the IP library [7]. In the design process, the SEL power switch cells should be placed exactly under the power stripes-row crossover points, instead of "filler" cells as usual. The power stripes and power rows at the points where a SPS is placed are connected only through the SPS. In the Fig.2 the power stripes are horizontal power supply lines and the power rows are presented as vertical lines. A SPS has one output – the controlled power supply line, used for one of the redundant circuits. This requirement is based on the concept of having separate power supplies for the two netlists used for the DMR.
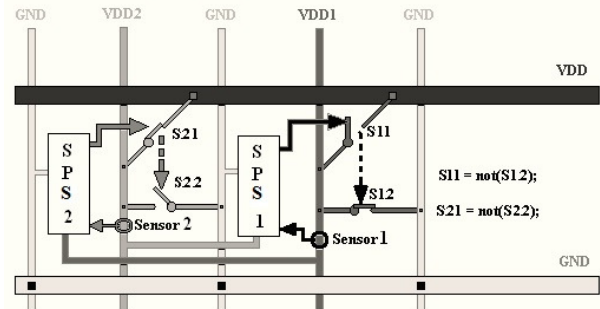
Fig. 2. SEL power switch cells for redundant circuitry

The SPSs, connected on the power supply lines of a redundant circuitry, are presented in Fig. 2. When a latchup is detected, for example on VDD1, the sensor 1 detects much more current then usual and at the same time the drain voltage of the switch S11 decreases to the ground level. Therefore, switch S11 will be opened and switch S12 closed. In this period, when switch S11 is open, the controlled logic is disconnected from the main power supply line VDD1. A sensor is a specially designed transistor for the fast reaction in the latchup handling process. The circuit, where latchup is detected should be disconnected from the main supply line for a period of time defined by control logic (timer or neural network), in order to stop the current flow through the parasitic PNPN structure in CMOS transistor pair, shown in Fig. 3 [A Hastings].
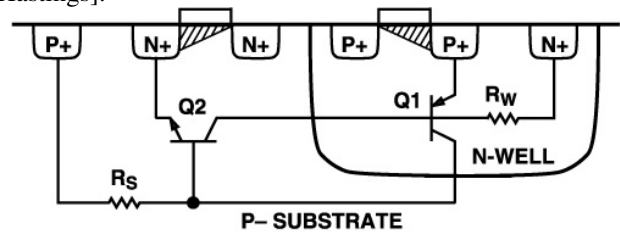
Fig. 3. The parasitic PNPN structure

The redundant circuits are connected to the separated power supply lines - we can call them "domains". As is shown in Fig. 2, SPSs are working in such a way that the SPS, which is connected to one power domain (e.g. VDD1), always controls another power supply domain (e.g. VDD2). This scheme for the power supply connection is done for self-protection of the SPS from the latchup effect. The control logic in a SPS needs to have a continuous and independent power supply in order to make accurate control of the power supply lines for the logic suffering a latchup condition.

In this way, a SPS can react quickly to the latchup and simply makes a disconnection of the harmed part of a system.

IV. CHARACTERIZATION OF THE SPS CELL

A. Introduction

The characterization process of the SPS cell is done through the three main steps. In the development process, the simulation was first important step in order to scale the W/L ratio of the transistors and to provide enough current for the correct functionality of the SPS. The second important step in the characterization process was the implementation of the SPS based circuits. There are three implemented circuit types: the SPS cell, the drive transistors (T5 and T6, Fig. 4) and the SPS cell with small digital system. The measurement was the third important step in the characterization process.

B. Simulation

In Fig. 4 is presented the simplified schematic of the SPS used for simulation. As the SPS cell is used to provide protection on the induced latchup (short-circuit), the simulation (functionality) description is based on this effect.
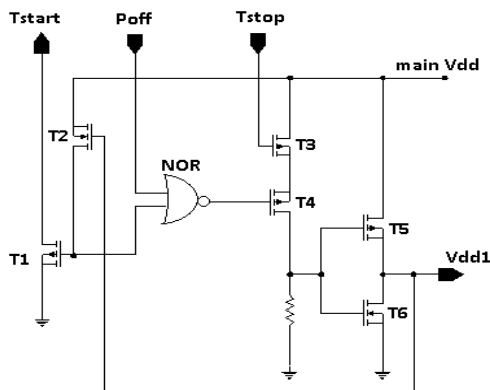
Fig. 4. SPS schematic

In case that output pin Vdd1 is short-circuited, the transistor T5 conducts more current than usual and the voltage between source and drain is higher. That means - the voltage on drain of the PMOS (T5) transistor is being lowered. The time required to set output pin Vdd1 on the zero voltage (ground level) is defined as a power off time (POFT in Fig. 5). Feedback line from the Vdd1 pin causes transistor T2 to activate when mentioned voltage is under the threshold voltage. Automatically, the transistor T1 will trigger Tstart (low active) output pin. Required time to trigger Tstop after latchup is defined as a latchup recognition time (LRT in Fig. 5).

Working condition for the T5 transistors to operate in the linear range and to provide enough current for the digital circuit was one of the most challenging points. From one side, the transistor T5 should be sensitive on the higher current flows, but from another side, the digital system (controlled by the SPS) also has some current fluctuations

during the operation time. The solution was to design the transistor which will work on the border between the saturation and the linear range.

In order to wake up the power switch circuit (SPS) from the latchup protection mode, it is required to provide an impulse on the Tstop pin. The minimal length of the Tstop impulse is defined as a minimal activation time (MAT in Fig. 5). This impulse should stop the current flow through the transistor T3 and set the gates of transistors T5 and T6 on the low voltage level. The transistor T5 should activate and provide power supply on Vdd1 output pin. The time required for this process is defined as a power on time (PONT in Fig. 5). The feedback line is deactivating the transistors T2 and T1, where Tstart pin should be set on the high voltage level, whereby latchup protection sequence is finished. The time required for this process is defined as a protection deactivating time (PDT in Fig. 5).

C. Implementation

In order to prove the functionality of the SPS, we have implemented three groups of test circuits in standard IHP 250 nm process [6]. The first group of test circuits is based on the main functionality tests of the SPS and timing measurements in the moment when latchup occurs. The second group of test circuits is used for measuring the burn off time of the output transistor (T5) in the latchup (short-circuit) mode. Transistors are designed with three different W/L ratios – the smallest, middle and the biggest W/L ratio. The third group of test circuits is used for the functional analysis when the SPS cell is integrated in the small digital system.

D. Measurements

Timing measurements are done in order to characterize the SPS as a standard power switch cell for usage in the automated design flow. In Fig. 5 is presented waveform diagram of the mentioned signals and required timings. The strobe signal, in Fig 5, presents the latchup activation signal. The Table I presents simulated and measured signal timings. The timing measurements are done on the 50% voltage level of the signal transitions.
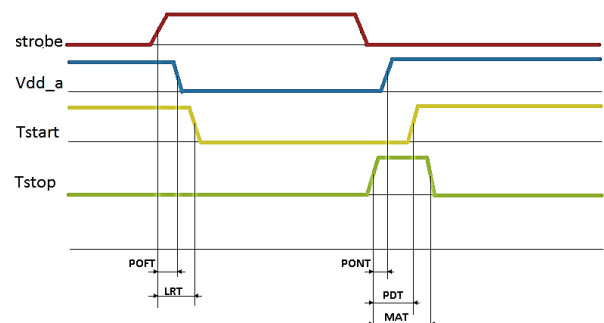
Fig. 5. SPS timing diagram

Before we discuss simulated and measured SPS timings, it is important to note that measurements are done with standard equipment. Measured values are normalized by correction factors for each pad on the test circuit due parasitic capacitances in cables and connections.

TABLE I
SPS TIMINGS (WITH CORRECTED MEASURED VALUES)

	POFT	LRT	PONT	PDT	MAT
Simulated	55.19 [ps]	76 [ps]	487 [ps]	786.5 [ps]	700 [ps]
Measured	120 [ps]	440 [ps]	1.42 [ns]	1.71 [ns]	2 [ns]

It is important also to note that in case of permanent short circuit on the Vdd1 output pin, SPS will automatically be in the protection mode. It is also possible to control activity of the SPS by the Poff pin.

Maximal current tests are based on the longer time shortcut and measurements when the output transistor of power switch cell will be destroyed. The burn off test is done for the output PMOS transistor T5 in order to prove that transistor will survive high current flow in the moment when latchup occurs. Fig. 6 presents the simulation result of the voltage and current for the transistor T5 in the moment when latchup triggers.
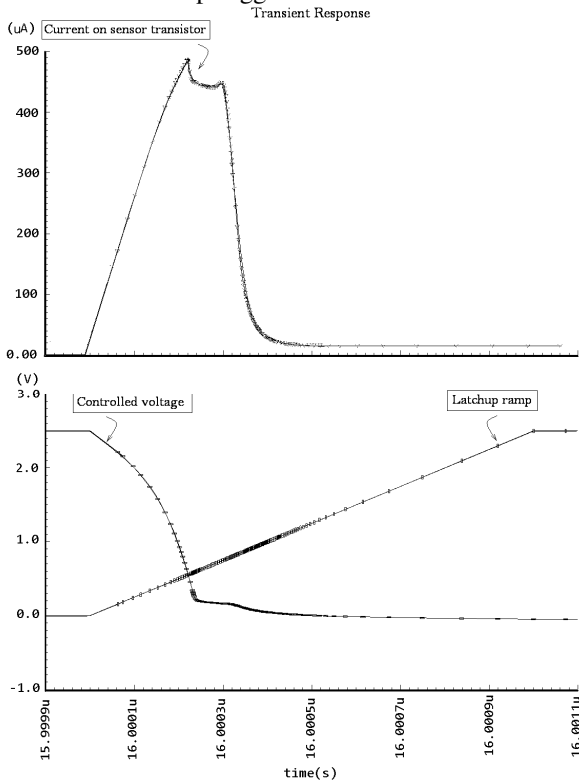
Fig. 6. The current and voltage diagrams after latchup occurrence

In the Table II are presented burn off timings of different test transistors which can be used as transistor T5 (Fig. 4).

TABLE II
BURN OFF TIMINGS (WITH CORRECTED MEASURED VALUES)

W/L	20.83	41.67	208.33
$t_{burn-off}$ [ns]	20	30	38

The measured power consumption, of SPS itself, is about 500 µW in normal conditions and it goes up to 1.25 mW when stimulated latchup occurs. A simulated value of the SPS power consumption is 75 pW in normal conditions and 1.16 mW in the latchup mode. This power consumption difference, between two mentioned modes, is due to the pull-down resistor shown in Fig. 4.

The system used for measurements of the SPS cell, which is integrated in digital circuit is presented in Fig. 7.
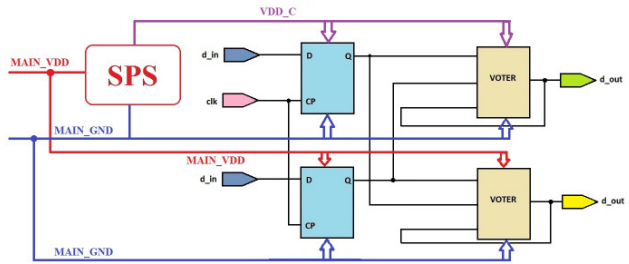
Fig. 7. Digital dual modular redundant circuit with power supply short-circuit protection

After the measurements of the SPS cell, next step was including the protection cell (SPS) in a small digital circuit and testing the functionality when the controlled power supply is short-circuited (VDD_C on Fig. 7). The measurement results are shown in the Fig. 8 [9].
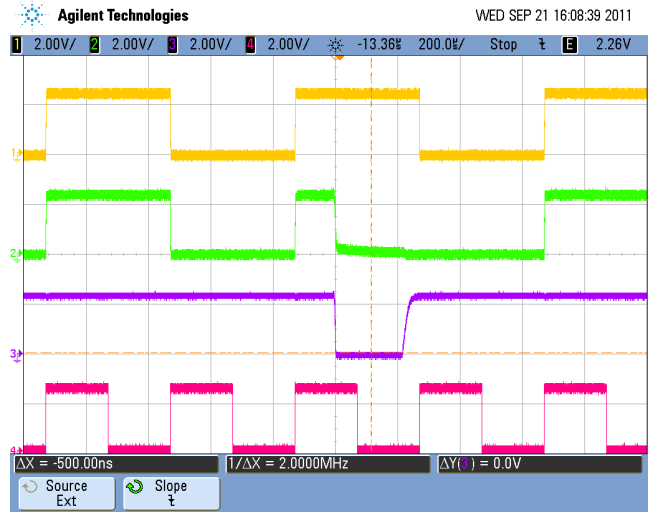
Fig. 8. Measurements of the SPS cell included in digital circuit

In Fig. 8, the third signal presents the power supply line where the latchup is triggered. The second signal in Fig. 8 is the data output of the logic which is supplied by controlled power supply (SPS). The usage of a redundant digital circuit (DMR in this case) provides recovering of the data after latchup is relaxed. From Fig. 8 it is possible

to see that short-circuit protection is active as long as the latchup effect exists.

IV. Conclusion

In order to protect an ASIC against SEL effects, we have developed, implemented and tested a custom protection cell – the power switch (SPS).

During tests, we have noticed a problem. The problem is based on the setup time violation in the moment when the latchup protection is deactivated near the active clock edge. Future work will be based on the development of a protection delay and synchronized components in order to avoid such a problem.

It is important to notice that the simulation process of the system with latchup protection still needs to be done in the analog environment in order to verify the complete functionality – the functionality of the digital system and the functionality of the protection system.

The tests of SPS cell, integrated with simple DMR system, have shown correct behave in the latchup period, what provides the adequate usage of the new design flow methodology [1]. The result comparison between simulated and measured values has shown the correctness of the IHP's transistor models used for development and implementation of the SPS cell. The measurement of the SPS driving transistor burn-off time has also shown that during the latchup effect all components in the SPS will not be destroyed.

Further development will be based on the design flow automation for the SPS cells in the layout process and the possibility to implement a neural network for self-testing and control of the power switches (SPS). Radiation tests of described test structures and tests on another technologies (IHP25RH and IHP13) [7] are planned for the next year.

References

[1] V. Petrovic, G. Schoof „Design Flow Approach for Reliable ASIC Designs", the 7th International New Exploratory Technologies Conference (NEXT 2010), Turku, October 19 - 21, 2010, Finland

[2] J. Teifel, "Self-Voting Dual-Modular-Redundancy Circuits for Single-Event-Transient Mitigation", IEEE Transaction on Nuclear Science, Vol.55, No. 6, December 2008

[3] R. E. Lyons, W. Vanderkulk, "The Use of Tripple-Modular Redundancy to Improve Computer Reliability", IBM Journal April 1962

[4] Ragaie, H.; Kayed, S., "Impact of CMOS device scaling in ASICs on radiation immunity", 2002. (EWAED), The First Egyptian Workshop on Advancements of Electronic Devices, 2002

[5] Fault Tolerant Design of Digital Systems, Available: http://faculty.ksu.edu.sa/musaed/CEN491Doc/Fault%20 Tolerant.doc

[6] Alan Hastings, "The Art of Analog Layout - Second Edition", Chapter 4 – Failure Mechanisms, Page: 171

[7] Institute for High Performance Microelectronics – IHP, Frankfurt Oder, Germany. Available: www.ihp-microelectronics.com

[8] G. Schoof, M. Methfessel, R. Kraemer, "Fault Tolerant ASIC Design for High System Reliability", Advanced Microsystems for Automotive Applications 2009 VDI-Buch, 2009, Part 4, 369-382, DOI: 10.1007/978-3-642-00745-3_24

[9] Agilent Technologies, MSO6104A Mixed Signal Oscilloscope: 1 GHz, 4 scope and 16 digital channels, Available: http://www.home.agilent.com/agilent/product.jspx?pn= MSO6104A&cc=DE&lc=ger

Adiabatic Digital Circuits Based on Sub-threshold Operation of Pass-transistor and Slowly Ramping Signals

Aleksandar Pajkanović, Tom J Kazmierski, Branko Dokić

Abstract - An overview of pass-transistor logic and subthreshold operation of transistors is given in this paper. Benefits of combining these two design principles from overall energy consumption point of view are discussed. A simulation to prove extremely low supply voltage and low power consumption operation of thus designed digital circuits is performed. The results obtained imply that it is possible to preserve energy by slowing down transitions between input signal logic levels.

Keywords – pass-transistor logic, sub-threshold operation, low power consumption, energy efficiency, energy harvesters.

I. INTRODUCTION

Logic circuits in standard applications are optimized in such way that their operation yields minimum delay. The operation point which is targeted in such optimization is known as the minimum-delay operation point (MDP). Since minimum delay means maximum speed and maximum speed implies maximum power consumption, it is not possible to use this approach when energy efficiency is an issue. This issue appeared with the emergence of the applications that require ultralow energy levels, such as energy-harvester powered wireless sensors. In such systems the energy consumption, rather than speed, is the most important design concern. In order to address the energy concern, the opposite of MDP, i.e. the minimum-energy operation point (MEP) also became a very interesting part of the energy-delay space. This has led to a completely changed design approach: to design a circuit with the power consumption minimized, initial design point is MEP, not MDP [1].

It is well known that MEP occurs in the sub-threshold operational region of the MOS transistors and that its value is set by leakage [1]. The operation at MEP has been demonstrated and proven to be possible [2]. Of course, as there were disadvantages in operating at MDP, there are some when operating at MEP. The delay at MEP is at least three orders of magnitude larger than at MDP, as illustrated in Fig. 1. Besides, sub-threshold logic has to be ratioed to be functional and sensitivity to parameter variance is

Aleksandar Pajkanović and Branko Dokić are with the Department of Electronics, Faculty of Electrical Engineering, University of Banja Luka, Patre 5, 78000 Banja Luka, Republic of Srpska, Bosnia and Herzegovina, e-mail: {aleksandar.pajkanovic, bdokic}@etfbl.net

Tom J Kazmierski is with University of Southampton, Southampton SO17 1BJ, UK, e-mail: tjk@ecs.soton.ac.uk.

increased in the weak inversion area [1].

Ever since the pass-transistor logic (PTL) was introduced over two decades ago by K. Yano in his papers (e.g. [3]), this technology was sure to yield improvement on the field of power consumption, speed and area when logic circuits are in question [4]. The main concept behind PTL is the utilization of nMOS pass-transistor network for logic organization, instead of source-grounded nMOS trees in the conventional differential logic. The main reason that PTL achieves higher speed and lower power dissipation is that its input capacitance is about half that of the conventional CMOS configuration. Also, PTL is able to realize complex Boolean functions efficiently in a small number of MOS transistors, thus reducing the area and delay time [3]. The first design methodology intended for PTL was introduced by Yano e.a. [4], thus solving the most important reason for PTL not being able to capture a major role within logic circuits design.
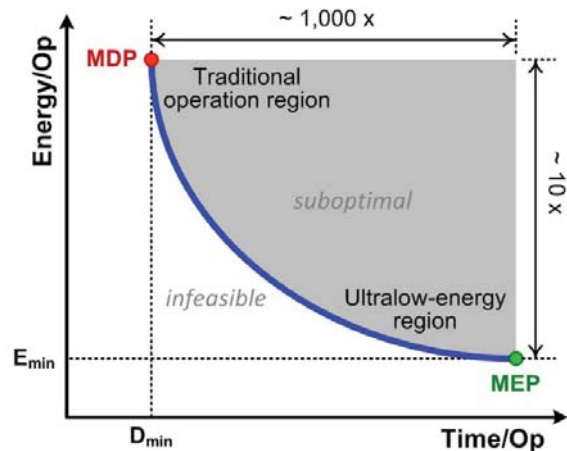
Fig. 1. Energy-delay space, MDP and MEP marked [1].

It has also been pointed out [5] that all previous researches on PTL examine the behaviour of full-adders. In that paper it is claimed that a full-adder structure is very easy to design using any of PTL approaches and is the least efficient structure for CMOS. Also, it is said that PTL styles have difficult layout and require greater design efforts. These conclusions come from the claim that CMOS is the most widespread logic style at the point of time. Since [5] has been published almost fifteen years ago, there

is no reason to oversee all the other aforementioned advantages of PTL [6].

Two different approaches for minimising energy consumption have been mentioned until now. Each of these has been subject of many different papers, thus they are both very thoroughly covered in literature, but their combination is not. Since PTL is considered low power logic, operating it in sub-threshold region, which implies extremely low energy consumption, would yield an ultralow power digital system. Because of the minimal voltage supply (a condition to operate in sub-threshold region), such circuits would have limited performance in terms of speed. They would therefore be suitable for applications where performance is not of primary importance, such as energy harvester applications. Namely, the conventional method uses a general purpose MCU with performance capabilities much more advanced than what the system requires. These MCUs are able to work in a wide range of higher clock frequencies, they have many fast and accurate input/output peripherals and terminals. In contrast to the above mentioned standard MCUs, in energy harvester applications more energy efficient circuits and modes of operation are required. These conditions are to be yielded by the combination of the following two design methods: PTL and sub-threshold operation. Thus, two main benefits for energy harvesters would arise: very low voltage supply and ultralow power consumption of the system [6].

In this paper a simulation of PTL digital circuits operating in sub-threshold region is performed. Throughout the simulation the rise and fall time intervals of the input voltage are varied, thus slowing down the change of the input voltage from one logic level to the other. The results of the simulation showed that the power consumption decreases as the rise and fall time intervals increase until a certain point where it reaches its minimum. Afterwards, power consumption increases again. Thus, in order to preserve energy, changes of the input voltage should be slower.

To the authors' knowledge, accept for [6], the literature on PTL operating in sub-threshold region is very scarce.

The rest of the paper is structured as follows. In sections II and III short overviews of operation in sub-threshold region and PTL are given, respectfully. Section IV presents the performed simulation and the results thusly obtained. In section V the results are discussed.

II. OPERATION IN SUB-THRESHOLD REGION

When operating in a standard set-up, a MOSFET has the gate voltage greater than the threshold voltage, $V_{gs} > V_T$. This is operation in strong inversion region. Ideally, for $V_{gs} < V_T$ MOSFETs do not conduct current. However, there is a number of carriers which create a current between drain and source. Thus, MOSFET is in the weak inversion region, because the inversion layer of carriers is not yet formed. This region is also known as the sub-threshold

region of operation [7].

In the sub-threshold region the gate current is negligible relative to the sub-threshold current because it decreases much faster with V_{DD}. Other leakage components such as the gate induced drain leakage and pn-junction leakage are also negligible in sub-threshold. Thus, the following analysis justifiably equates the total current to the sub-threshold current for V_{DD} in the sub-threshold region [7]:

$$I_{sub} = I_0 e^{\frac{V_{gs}-V_t}{n\varphi_t}} , \qquad (1)$$

$$I_0 = \mu_0 C_{ox} \frac{W}{L}(n-1)\varphi_t^2 , \qquad (2)$$

where n is the sub-threshold slope factor ($1+C_d/C_{ox}$) and φ_t is kT/q.

An equally important parameter is the delay of logic gates. Eq. (3) shows the propagation delay of a characteristic inverter with the output capacitance C_g in sub-threshold [7]:

$$t_d = \frac{KC_g V_{DD}}{I_{o,g} e^{\frac{V_{gs}-V_{t,g}}{n\,\varphi_t}}} \qquad (3)$$

where K is a delay fitting parameter. The expression for the current in the denominator of Eq. (3) models the *on* current of the characteristic inverter, so it accounts for transitions through both nMOS and pMOS devices. Unless the pMOS and nMOS are perfectly symmetrical, the terms $I_{o,g}$ and $V_{T,g}$ are fitting parameters that do not correspond exactly with the MOSFET parameters of the same name [7].

The operational frequency is simply [7]:

$$f = \frac{1}{t_d L_{DP}} \qquad (4)$$

where L_{DP} is the depth of the critical path in characteristic inverter delays [7].

A mathematical model for the total energy consumption per cycle is further developed [7], yielding expressions for calculation of the optimum supply voltage and optimum threshold voltage for a given performance condition – frequency. These expressions are [7]:

$$V_{DDopt} = n\,\varphi_t[2 - lambertW(\beta)] \qquad (5)$$

$$V_{Topt} = V_{DDopt} - n\varphi_t ln\left(\frac{fKC_g L_{DP} V_{DDopt}}{I_{o,g}}\right) \qquad (6)$$

where:

$$\beta = \frac{-2C_{eff}e^2}{W_{eff}L_{DP}KC_g} > -\frac{1}{e}; \qquad (7)$$

C_{eff} is the average effective switched capacitance of the entire circuit, including the average activity factor over all of its nodes; W_{eff} estimates the average total width relative to the characteristic inverter; Lambert W function gives the solution to the equation $We^W = x$ [7].

The constraint given by Eq. (7) shows that there is a maximum achievable frequency for a given circuit in the sub-threshold region [7].

A way to even more improve energy efficiency while operating in sub-threshold region is described in [8]. Namely, dynamic threshold MOS (DTMOS) technique and its advantages and disadvantages are shown in this paper. Using this technique, the transistor threshold becomes dependent on the gate voltage – because gate is connected to the body. Thus, low leakage ($V_g = V_b = 0 \rightarrow V_T$ high) and high drive ($V_g = V_b = V_{DD} \rightarrow V_T$ low) are obtained. In the paper it is demonstrated that DTMOS can be used for a broad range of supply voltages. DTMOS delay and efficiency are superior to traditional designs as the voltage is reduced and the loading is increased. A drawback of this technique is that it is limited to 0.5V supply voltage, because forward biasing the source-body pn junction would lead to excessive gate current. Also, the problems that appear are area penalty and process complexity [8].

III. PASS-TRANSISTOR LOGIC

In contrast to classic static CMOS logic, in PTL two input logic signals are applied at the gate and at the drain of a MOS transistor, as shown in Fig. 2. Considering an ideal case, without load, the transistor shown in Fig. 2 is saturated when its gate and drain voltages are equal to V_{DD}. Thus, the source voltage is $V_{DD}\text{-}V_T$. However, the source will be in high impedance state for 0V gate voltage, no matter what the drain voltage is. Therefore, another nMOS is added and sources are connected to a single node to ensure that the logic function is valid for both values of B. This is illustrated in Fig. 3. This is an effective method to realize the logic AND function, because it requires only two nMOS transistors, whereas classic static CMOS would use six transistors. Additionally, it is possible to alter the logic function of the circuit only by changing the wiring of the input signals [6].
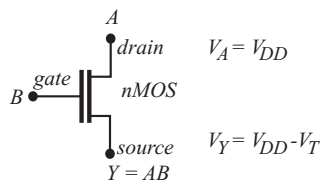
Fig. 2. NMOS transistor operating as pass-transistor.
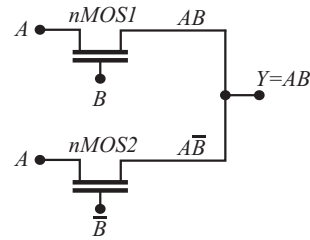
Fig. 3. Two nMOS transistors creating logic function AND.

A significant disadvantage of PTL is that the output voltage is lower than the input and that it does not allow series connections of large numbers of transistors. The addition of a static inverter, Fig. 4, recovers the voltage swing to appropriate values [6].
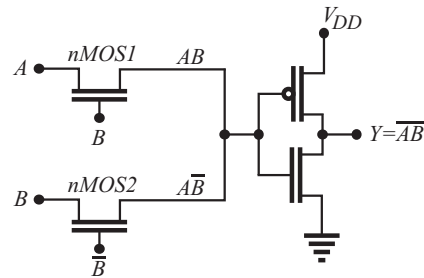
Fig. 4. Voltage swing recovery with static inverter.

There have been several different approaches to PTL design, some of which are: complementary pass-transistor logic (CPL), double pass-transistor logic (DPL) and dual value logic (DVL). All of these have their advantages and disadvantages, but CPL will be used within this paper, since it has been shown to result in high speed and high logic functionality [9].

CPL consists of complementary input/output, nMOS pass-transistor logic network and CMOS output inverters. A pMOS latch can also be added to CPL, as shown in Fig. 5, in order to decrease static power consumption and return the full swing [10]. Arbitrary Boolean functions can be constructed from the pass-transistor network by combining four basic circuit modules: AND/NAND, OR/NOR, XOR/XNOR and a wired-AND/NAND module [3].

There have been reports [11-13] which prove the functionality of CPL. The results shown in these papers justify the usage of CPL since it is said that the effect of parasitic capacitances is decreased [11] and delays can be reduced by 30% [12]. Circuit synthesis is possible using multiplexers and inverter only as components of the PTL cell library [13].
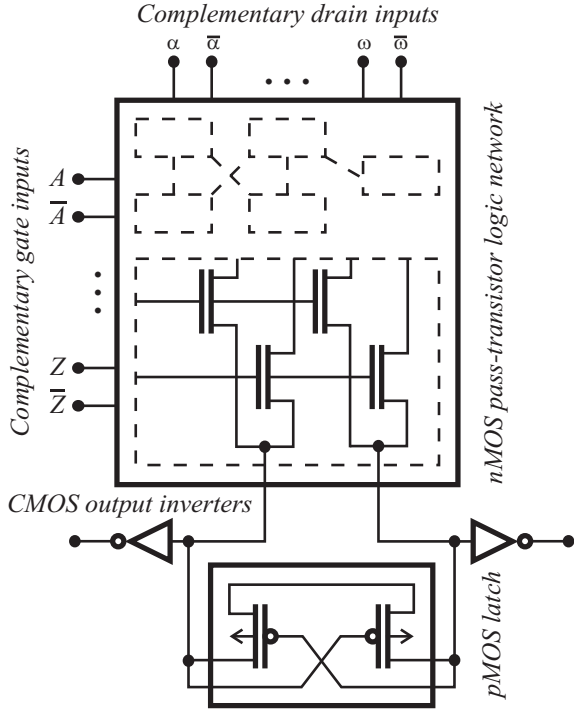
Fig. 5. Basic circuit configuration in CPL [3].

IV. SIMULATION RESULTS

In order to show the influence of the input signal switching speed on the behaviour of PTL operating in sub-threshold region concerning power consumption, a set of simulations have been performed and are presented in this paper.

Two digital circuits designed using aforementioned low-power principles have been simulated – a NAND gate and a full-adder. For both of these the supply voltage is $V_{DD} = 0.3$V. The NAND circuit was simulated as shown in Fig. 4 and full-adder was simulated as shown by Ivanov [6]. MOSFET models used in these simulations are BSIM3 and the technology process is 130nm. The BSIM model shows good physical behaviour in the sub-threshold region [14]. Since Ivanov [6] uses MOSFET models of a 350nm technology process, in this paper the channel widths are scaled down proportionally. Throughout the simulations all the inputs, for both circuits, are short-circuited since the input voltage set-up causes a logic state change at the output for every clock interval. Thus, the worst power consumption case is observed. In order to obtain more accurate results, a simulation was performed over one hundred cycles of the clock signal. The power consumption is calculated as:

$$W = V_{DD} \cdot \int i_{DD}(t) dt. \qquad (8)$$

Since this is the power consumed over one hundered

cycles, the power per cycle is given as:

$$W_{pc} = W / 100. \qquad (9)$$

Throughout the simulations the rise, fall, high and low intervals of the input signal are as follows. The rise interval t_r is defined as the time that the input signal takes to change between logic zero to logic one. The fall interval t_f is the time taken to change from logic one to logic zero. The high interval t_h is the time during which logic one is held at the input and during the low interval t_l logic zero is held. The interval values for which the simulations were carried out are:

$$t_h = t_l = \{5, 50, 500\} \text{ } [\mu s] \text{ and}$$

$$t_r = t_f = \{0, 10, 20, 50, 100, 200, 500, 1000\} \text{ } [ns].$$

All the combinations of the two listed sets were simulated and the results are shown in Fig. 6 and Fig. 7 for the NAND circuit and the full-adder, respectively.
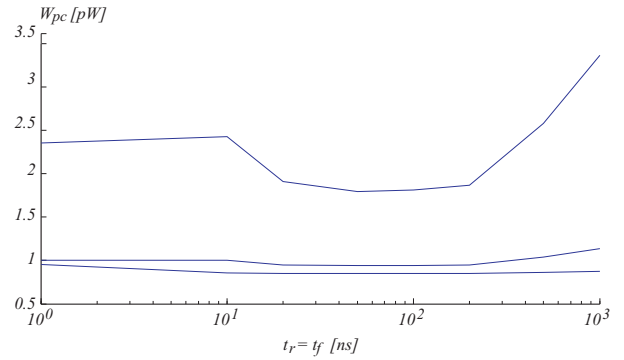
Fig. 6. Dependence of power consumption of sub-threshold PTL NAND circuit on rise and fall time of input signal.
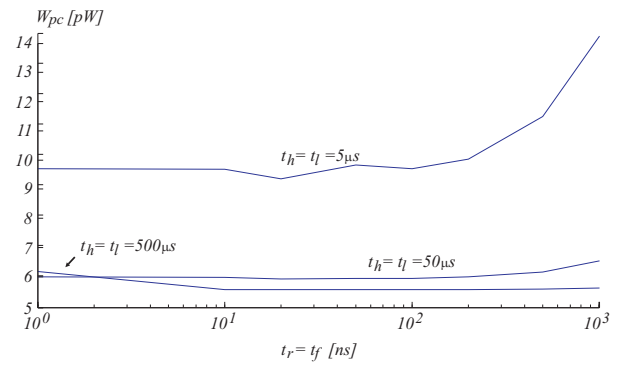
Fig. 7. Dependence of power consumption of sub-threshold PTL full-adder circuit on rise and fall time of input signal.

In Table I the rise and fall intervals at which the minimum power consumption is reached and the percentage reduction achieved relative to the instantaneous switching are shown.

$t_h = t_l$ [µs]	NAND		full-adder	
	$t_{rmin} = t_{fmin}$ [ns]	power saved [%]	$t_{rmin} = t_{fmin}$ [ns]	power saved [%]
5	50	15.84	20	3.71
50	50	6.05	20	1.35
500	50	31.57	50	30.15

Simulation files are available at the authors' webpage[1].

V. DISCUSSION

The static component of the transistor sub-threshold current is given by eq. (1). Since this current does not depend on the input voltage rise and fall time intervals, from the results obtained in section IV it is obvious that through the supply voltage source, there exists another component of the sub-threshold current, namely the dynamic component. This current, I_{dyn}, is explained as the influence of the parasitic capacitances during the transition periods t_r and t_f. The capacitances which are greatly responsible for this current are shown in Fig. 8.
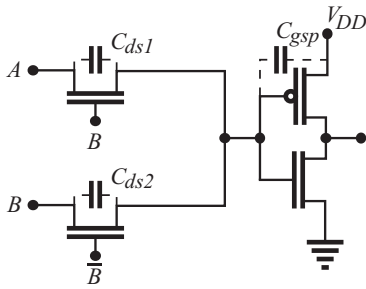
Fig. 8. Capacitances through which I_{dyn} flows.

Since $C_{gsp} << C_{dsp}$, the dynamic current flows for a shorter time than that for the current that flows through the output inverter, obtained from (1) and expressed as [10]:

$$I_{static} = I_{sp} = I_{sn} = I_0 e^{\frac{V_{gsn} - V_{tn}}{n\varphi_t}} = I_0 e^{\frac{V_{gsp} - V_{tp}}{n\varphi_t}}. \quad (10)$$

Also, the equation that describes dynamic component is dervided as:

$$I_{dyn} = f\left(C \cdot \left(\frac{du}{dt}\right)\right). \quad (11)$$

Further, the power consumed by the flow of this current is as:

$$W_{dyn} = V_{DD} \frac{1}{t_r} \int_0^{t_r} f\left[C \cdot \left(\frac{du}{dt}\right)\right] dt, \quad (12)$$

where C represents the effective capacitance of the whole circuit given in Fig. 8, and u is the voltage over this capacitance. An analogous equation can be derived for the input voltage fall time interval.

From eq. (12) it is obvious that the dynamic current decreases as time interval, during which the voltage difference appears, increases. However, this decrease in power consumption by increasing the time intervals of switching levels can happen only to a certain point. Namely, at the same time as the average value of I_{static} increases the above mentioned time intervals increase. With this increase, I_{static} becomes the dominant component, thus disabling further power consumption decrease.

Even though a concrete equation which utterly explains the dynamic component is not given in this paper, the previous discussion explains the behaviour of the power consumption shown if Figs. 6 and 7. At first, while I_{dyn} is dominant, the power consumption decreases with the increase of the switching time intervals. It reaches a minimum; and, when I_{static} becomes greater, it starts to increase.

VI. CONCLUSION

A summary of PTL and subthreshold ultralow energy operation has been given in this paper. Within sections II and III the advantages and disadvantages of these technologies have been described. Benefits of combining these two low power consumption design approaches have been pointed out.

Through a simulation of digital circuits based on sub-threshold operation of PTL it has been proven that this combination of technologies yields a very low energy consumption system. Also, according to the simulation results, it is possible to decrease the energy consumption by slowing down the switching time of logic levels at the system input. In other words, if the rise and fall time intervals are increased, the dynamic component of the supply voltage current is decreased, thus lowering the energy consumption. These results are shown in section IV, and discussed and explained in section V.

In future work, more detailed measurements will be made so that the presented effects can be further examined. Also, a detailed mathematical model for the dynamic component will be developed.

REFERENCES

[1] D. Markovic, C.C. Wang, L.P. Alarcon, L. Tsung-Te, J.M. Rabaey, "*Ultralow-power design in near-threshold region*," IEEE Journal of Solid-State Circuits, Feb. 2010, pp. 237-252.

---

[1] pajkanovic.netne.net/2012/ssss

[2] B. Zhai et al., "*A 2.60 pJ/inst subthreshold sensor processor for optimal energy efficiency*," IEEE J. Solid-State Circuits, vol. 40, no. 9, pp. 1778-1786, Sep. 2005.

[3] K. Yano, T. Yamanaka, T. Nishida, M. Saitoh, K. Shimohigashi, A. Shimizu, "*A 3.8 ns CMOS 16x16 multiplier using complementary pass transistor logic*," Proceedings of the IEEE Custom Integrated Circuits Conference*,* San Diego, May 1989, pp. 10.4/1-10.

[4] K. Yano, Y. Sasaki, K. Rikino, K. Seki, "*Top-down pass-transistor logic design*," IEEE Journal of Solid-State Circuits, vol. 31(6), Jun. 1996, pp. 792-803.

[5] R. Zimmermann, W. Fichtner, "*Low-power logic styles: CMOS versus pass-transistor logic*," IEEE Journal of Solid-State Circuits, vol. 32(7), Jul. 1997, pp. 1079-1090.

[6] I. Ivanov, "*Sub-Threshold Pass-Transistor Logic for Ultra-Low Power Processor Suitable for Energy Harvester Application*", MSc in Microelectronics System Design, University of Southampton, Faculty of Physics and Applied Science, School of Electronics and Computer Science, September, 2011.

[7] B.H. Calhoun, A. Wang, A. Chandrakasan, "*Modeling and sizing for minimum energy operation in subthreshold circuits*," Proceedings of the IEEE*,* Sep. 2005, pp. 1778-1786.

[8] N. Lindert, T. Sugii, S. Tang, C. Hu, "*Dynamic threshold pass-transistor logic for improved delay at lower power supply voltages*," IEEE Journal of Solid-State Circuits*,* Jan. 1999, pp. 85-89.

[9] D. Markovic, B. Nikolic, V.G. Oklobdzija, "*General method in synthesis of pass-transistor circuits*," Proc. of 22nd International Conference on Microelectronics*,* Nis, May 2000, pp. 695-698.

[10] B. L. Dokić, *Integrisana kola - digitalna i analgona*, Elektrotehnički fakultet, Banja Luka, "Glas srpski", 1999, ISBN: 86-7122-011-7.

[11] V. M. Srivastava, R. Patel, H. Parashar, G. Singh, "*Reduction in parasitic capacitance for transmission gate with the help of CPL*," International Conference on Recent Trends in Information*,* Telecommunication and Computing*,* Kochi, Mar. 2010, pp. 218-220.

[12] M. Suzuki, N. Ohkubo, T. Shinbo, T. Yamanaka, A. Shimizu, K. Sasaki, "*A 1.5-ns 32-b CMOS ALU in double pass-transistor logic*," IEEE Journal of Solid-State Circuits, vol.. 28(11), Nov. 1993, pp. 1145-1151.

[13] S.-F. Hsiao, M.-Y. Tsai, C.-S. Wen, "*Efficient pass-transistor-logic synthesis for sequential circuits*," IEEE Conference on Circuits and Systems*,* Dec. 2006, pp. 1631-1634.

[14] *BSIM MOSFET Model User's Manual*, Tanvir Hasan Morshed, Darsen D. Lu, Wenwei (Morgan) Yang, Mohan V. Dunga, Xuemei (Jane) Xi, Jin He, Weidong Liu, Kanyu, M. Cao, Xiaodong Jin, Jeff J. Ou, Mansun Chan, Ali M. Niknejad, Chenming Hu, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720.

Full-swing Low Voltage BiCMOS/CMOS Schmitt Trigger

Branko Dokić, Tatjana Pešić-Brđanin, Aleksandar Pajkanović

Abstract - Full-swing low voltage BiCMOS Schmitt trigger which ensures full logic amplitudes in two different ways is described. In this paper full-swing Schmitt trigger circuits have two complementary outputs: BiCMOS (inverting) and CMOS (noninverting). Proposed mathematical model to determine threshold voltages is verified by simulation. Voltage hysteresis depends on supply voltage, MOS transistor threshold voltages and geometry ratio of input MOS transistors and transistors within the positive feedback loop. As this ratio increases from 0.3 to 2, voltage hysteresis changes from $0.2V_{DD}$ to $0.54V_{DD}$.

Keywords - low voltage, BiCMOS/CMOS, Schmitt trigger, full-swing, threshold voltages.

I. INTRODUCTION

Schmitt trigger is a circuit with a hysteresis shaped transfer characteristic. Its application is very wide, both in mixed signals circuits and in digital ones. As digital circuits, they are usually referred to as Schmitt logic circuits. Those are circuits with standard elementary logic functions (inverter, NAND and NOR) and a hysteresis shaped transfer characteristic [1]. Because of this, Schmitt logic circuits have larger noise immunity than standard circuits. Second advantage is that all output level changes are influenced by positive feedback loop, thus their transfer characteristic are almost ideal (width of the transitional region is neglible), so that noise margin and noise immunity are equal. On the other hand, low frequency signals at the input Schmitt trigger transforms into pulses with very short rise and fall times. Therefore, Schmitt trigger is often used as input circuit of standard MSI or VLSI integrated circuits.

Schmitt logic circuits are often used to design pulse generators. For example, astable multivibrator consists of a Schmitt inverter, a resistor and a capacitor. It is possible to vary RC constant within eight orders of magnitude, so these generators work in a very wide frequency range, from several Hz to several hundreds of MHz. If a current generator is placed instead of the resistor, a simple function generator is achieved. In the same way triangle voltage generators are constructed. Such a voltage, for example, is used as auxiliary voltage within pulse-width modulators. Mixed signals integrated circuits which contain Schmitt trigger are described in [1-3].

Branko Dokić, Tatjana Pešić-Brđanin, Aleksandar Pajkanović are with the Department of Electronics, Faculty of Electrical Engineering, University of Banja Luka, Patre 5, 78000 Banja Luka, Republic of Srpska, Bosnia and Herzegovina, e-mail: {bdokic, tatjanapb, aleksandar.pajkanovic}@etfbl.net

BiCMOS, and the other one is CMOS, from whence comes BiCMOS/CMOS in the paper title. Basic circuit, which generates hysteresis shaped transfer characteristic, is the same for both solutions. The difference is at the output stage which allows full-swing at BiCMOS output. Namely, logic amplitude of BiCMOS circuit with the standard output is: $\Delta V_o = V_{DD} - 2V_{BE}$, where V_{BE} is base-emitter voltage of the conducting output bipolar transistors. The decrease of the logic amplitude for $2V_{BE} \approx 1.5V$ limits minimum supply voltage to about 3V. In very low-power CMOS or BiCMOS circuits supply voltage is less than 3V. Schmitt triggers, described in this paper, operate reliably at $V_{DD} < 1V$.

II. STATIC ANALYSIS

There are more ways to implement CMOS Schmitt triggers [2-4]. In this paper a principle of two different logic threshold voltages of cascode input CMOS transistors, proposed in [5], is used. Other MOS transistors provide optimum operation of bipolar transistors and full-swing at BiCMOS output (Figs. 1 and 2).

Basic circuit of BiCMOS Schmitt trigger (Fig. 1) consists of MOS transistors M_{ni} ($i=0,\ldots,4$) and M_{pj} ($j=0,\ldots,3$) and two npn bipolar transistors at the output. Transistors M_{n0} and M_{n1}, i.e. M_{p0} and M_{p1} ensure the difference of the output level changes thresholds, during positive and negative input voltage change (hysteresis shaped transfer characteristic), and M_{n3} and M_{n4} actively turn off transistors T_1 and T_2, respectively, in static states. CMOS inverters I_1 and I_2, connected as a latch circuit, besides complement CMOS output $\overline{Q}$, ensure full logic swing at the output Q.

Dominant influence to the static transfer characteristic have transistors M_{ni} and M_{pi} ($i=0,1,2$). It is optimal when the corresponding pairs of these transistors are symmetric. This means equal constants k and threshold voltages of M_{n1}, M_{n2}, M_{p1} and M_{p2} at one side, and M_{n0} and M_{p0} on the other side, i.e.:

$$k_n = \frac{\mu_n \varepsilon_{ox}}{2t_{ox}} \frac{W_n}{L_n} = k_p = \frac{\mu_p \varepsilon_{ox}}{2t_{ox}} \frac{W_p}{L_p},$$

$$k_{n0} = \frac{\mu_n \varepsilon_{ox}}{2t_{ox}} \frac{W_{n0}}{L_{n0}} = k_{p0} = \frac{\mu_p \varepsilon_{ox}}{2t_{ox}} \frac{W_{p0}}{L_{p0}}$$

$$V_{tn} = V_{tni} = |V_{tp}| = |V_{tpi}|, \quad i=0,1,2, \tag{1}$$

where: μ_n and μ_p are the mobility of the electrons and the holes, ε_{ox} is oxide dielectric constants, t_{ox} the oxide

thickness and W and L are width and length of the transistor's channel. In digital circuits it is common to have equal channel lengths for all transistors, which is a specificity of the technology process. Therefore, it is assumed: $L_{ni}=L_{pi}$, $i=0,…,4$. Further, it is assumed: $k_n=k_{n1}=k_{n2}$ and $k_p=k_{p1}=k_{p2}$.

Let input voltage increase from 0 to V_{DD}. At $0{\leq}V_i{\leq}V_{tn}$, transistors M_{n1}, M_{n2}, M_{n3}, M_{p0}, T_1 and T_2 are turned off, and M_{n0}, M_{n4}, M_{p1}, M_{p2} turned on. The Q output is on the high voltage level, $V_{OH}=V_{DD}$. Without inverters I_1 and I_2, transistor T_1 would conduct, thus: $V_{OH}=V_{DD}-V_{BE1}$. Because of the latch circuit at the output, $V_{OH}=V_{DD}$. As it is: $V_3=V_{DD}$, T_1 is turned off. Since gate and drain of M_{n0} hold the same voltage (V_{DD}), this transistor is in saturation, thus $V_1=V_{DD}-V_{tn0}$.
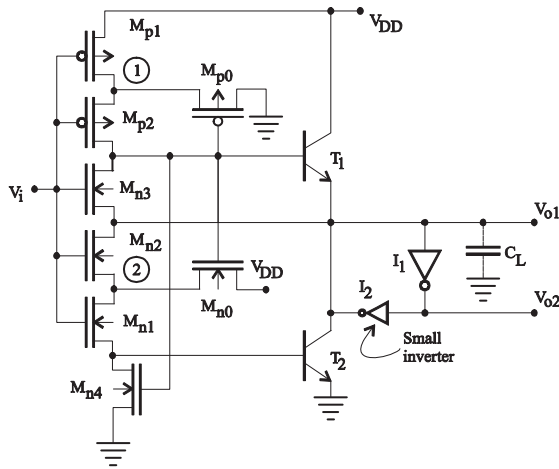
Fig. 1. Schmitt trigger with a latch circuit at the output.

For $V_{tn1}<V_i<V_1+V_{tn2}$, M_{n2} is turned off, and M_{n1} conducts in saturated region. Also, M_{n4} conducts until the state change at the output appears. Its influence to transistor's drain current can be taken into account by replacing M_{n1} and M_{n4} with one equivalent transistor [1] with twice as large channel length, i.e. constant $k_{ne}=k_n/2$. Equalizing drain currents in saturated area of this equivalent transistor and of M_{n0}, it is obtained:

$$V_1 = V_{DD} - V_{tn0} - \sqrt{W_n / 2W_{n0}} \left(V_i - V_{tn1}\right), \qquad (2)$$

where $W_n/(2W_{n0})=(k_n/2)/k_{n0}$, because of the assumption of equal channel lengths for all transistors. Therefore, V_1 linearly decreases as input voltage increases. Up to $V_i<V_1+V_{tn2}$, M_{n2} is turned off, thus there is no output voltage change, i.e. $V_o=V_{OH}=V_{DD}$. When:

$$V_i{\geq}V_1+V_{tn2}, \qquad (3)$$

M_{n2} conducts, so voltage V_3 will decrease as V_1 increases. M_{n0} plays the role of source follower, transferring negative

change of V_3 to 1, which accelerates M_{n1} turning on. Thus, positive feedback loop is created, which leads to step change of the outputs Q and $\overline{Q}$. Input voltage at which this change happens is the high threshold voltage of the Schmitt trigger, V_{TH}, and it is calculated from condition $dV_3/dV_i=$ $=-1$. The obtained equation is not explicitly solvable. However, it is possible to show that the state of the outputs changes immediately after M_{n1} starts to conduct. This means that approximate value of V_{TH} can be obtained from (2) and (3) by equalizing them and changing $V_i=V_{TH}$. This yields:

$$V_{TH} = \frac{V_{DD} + \sqrt{\left(W_n / 2W_{n0}\right)V_{tn}}}{1 + \sqrt{W_n / 2W_{n0}}} \qquad (4)$$

Therefore, the high threshold voltage V_{TH}, besides of supply voltage V_{DD} and nMOS transistor threshold voltage V_{tn}, depends on channel widths ratio of cascode input nMOS transistors and feedback transistor M_{n0}.

When $V_i=V_{DD}$, M_{n1}, M_{n2}, M_{n3} and M_{p0} are turned on, and M_{p1}, M_{p2}, M_{n4}, T_1 and T_2 are turned off. Inverter I_2 holds voltage at Q output at 0, i.e. $V_{OL}=0$V. M_{p0} is saturated, so $V_2=|V_{tp0}|$. When input voltage decreases, M_{p1} is turned on first, at $V_i=V_{DD}+V_{tp1}$, and then M_{p2} at:

$$V_i=V_2+V_{tp2}. \qquad (5)$$

For $V_2+V_{tp2}<V_i<V_{DD}+V_{tp1}$, M_{p1} is saturated, so equalizing drain currents of M_{p1} and M_{p0} yields:

$$V_2 = \sqrt{W_p / W_{p0}} \left(V_{DD} + V_{tp1} - V_i\right) - V_{tp0} \qquad (6)$$

Immediately after M_{p2} starts to conduct (Eq. (5)), through M_{p0}, positive feedback loop is established, thus step change appears at the output from 0 to V_{DD}. Combining Eqs. (5) and (6) and replacing $V_i=V_{TL}$, low threshold voltage of Schmitt trigger is obtained:

$$V_{TL} = \frac{\sqrt{\left(W_p / W_{p0}\right)\left(V_{DD} + V_{tp}\right)}}{1 + \sqrt{W_p / W_{p0}}} \qquad (7)$$

Therefore, the low threshold voltage also, besides V_{DD} and V_{tp}, depends on channel widths ratio of pMOS transistors M_{p1} and M_{p0}.

III. IMPROVED CIRCUIT

As it is already emphasized, output latch circuit with inverters I_1 and I_2 ensures full logic amplitude at BiCMOS output. Since inverter I_2 output is in parallel with BiCMOS output, it increases logic delay, because it decreases charge and discharge currents of the capacitive load. Thus,

transistors of that inverter should be up to about ten times smaller than the other MOS transistors. On the other hand, positive feedback loop, added by the latch, can distort static transfer characteristic under certain conditions.

Output of the improved Schmitt trigger (Fig. 2) with inverter I and transistors M_{n4} and M_{p2} does not bring in listed limitations, because BiCMOS output is applied at the input of the inverter I. Since circuit shown in Fig. 2 differs from the one in Fig. 1 only in output stage which enables full-swing, threshold voltages equations are the same.
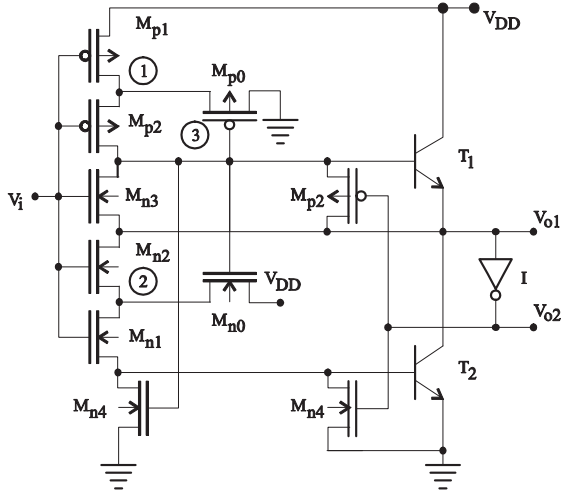
Fig. 2. Improved BiCMOS Schmitt trigger.

Analytic model is confirmed by simulation, in which BSIM MOSFET [6] and Gummel-Poon bipolar transistor models are used. Transistor parameters of 0.13µm CMOS technology process are used. Threshold voltages of transistor are $V_{tn}=|V_{tp}|=0.25$V. Channel lengths of all MOS transistors are 0.13µm, while nMOS and pMOS channel widths are $W_n=2$µm and $W_p=6$µm. In Fig. 3 transfer characteristic obtained by simulation at supply voltage $V_{DD}=1.25$V is shown.
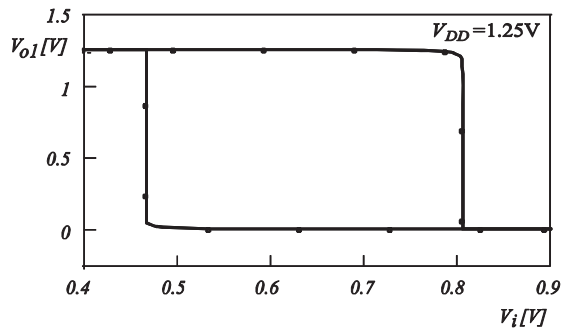
Fig. 3. Transfer characteristic obtained by simulation.

In Fig. 4 threshold voltages dependency on geometry ratio of input cascode transistors and feedback transistors

M_{n0} and M_{p0} at $V_{DD}=1.25$V is shown. Simulation shows high fidelity of the simplified analytic model to calculate thresholds V_{TH} and V_{TL}. Expectedly, calculated values are somewhat less than those obtained by simulation, because Eqs. (4) and (7) are input voltages at which M_{n2} and M_{p2} start to conduct, respectively. Only after this happens, a positive feedback loop can be established and output change appears. Somewhat more correct, and a lot more complicated analysis, quantitatively would not yield any new results. Actually, it would lose obviousness of the parameters which have the greatest influence. Both analytical model and simulation confirm that supply voltage and W_n/W_{n0} and W_p/W_{p0} ratios are dominant in determining V_{TH} and V_{TL}. Circuit designer, based on the required voltage hysteresis, from Eqs. (4) and (7) very easily can determine dimensions of the transistors M_{n0} and M_{p0}. Other MOS transistors geometry is determined by a standard procedure for CMOS digital integrated circuits. When input transistors are dimensioned so that their constants k are equal, i.e. $k_n=k_p$, ratios changes $W_n/W_{n0}=W_p/W_{p0}$ from 0.3 to 2, for example, voltage hysteresis can regulate from $0.2V_{DD}$ to $0.54V_{DD}$.
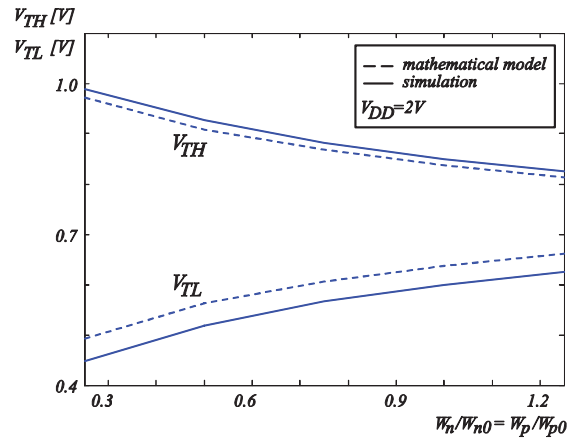
Fig. 4. Threshold voltages dependency on geometry ratio $W_n/W_{n0}= W_p/W_{p0}$.

It is known that transfer characteristic of Schmitt trigger is optimal when V_{TH} and V_{TL} are symmetric around $V_i=V_{DD}/2$. Because of the influence of transistor Mn4, transfer characteristics in Fig. 3 are not symmetric around $V_{DD}/2$. Symmetry is obtained when channel width of M_{p0} is twice as large as of M_{p1}, i.e. $W_{p0}=2W_p$. Then, for $V_{DD}=2$V, for example, threshold voltages, based on Eqs. (4) and (7): $V_{TH}=1.28=V_{DD}/2+0.2$V, and $V_{TL}=0.72$V$=V_{DD}/2-0.2$V. Therefore, transfer characteristic is symmetric with voltage hysteresis $V_H=V_{TH}-V_{TL}=0.56$V. Symmetry is verified by simulation (Fig. 5) at following parameters: $V_{DD}=2$V and $W_{p0}=12$µm. Other parameters are the same as those used for Fig. 3.
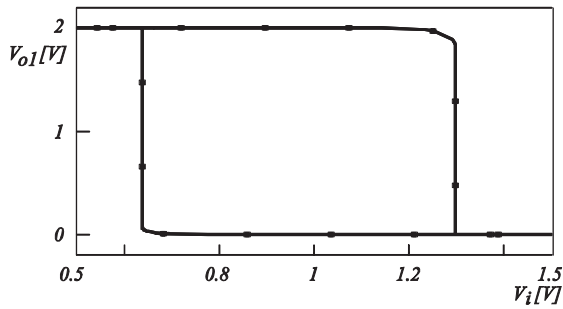
Fig. 5. Transfer characteristic obtained by simulation at $W_{p0}=2W_p$ and $V_{DD}=1$V.

Dependency of propagation time on ratios $W_n/W_{n0}= W_p/W_{p0}$, with capacitive load at the output $C_L=5$pF and $V_{DD}=2$V and $V_{DD}=1.5$V, is shown in Fig. 5. Change sensitivity of t_p is less in area $W_n/W_{n0}= W_p/W_{p0}>1$. The same stands for change sensitivity of V_{TH} and V_{TL} (Fig. 4). Increasing ratio $W_n/W_{n0}= W_p/W_{p0}$, both logic delay and voltage hysteresis are decreased. As V_H decreases, noise immunity decreases. Thus, to achieve a compromise between small propagation time and high noise immunity, it needs to be $W_n/W_{n0}=W_p/W_{p0}=1$.

Transistors of the input inverter I are defined based on the required strength of CMOS output. If this output is not used, the same geometry of transistors, as of those at the input, is recommended. If small hysteresis is required, output inverter should be dimensioned so that its threshold voltage is less than $V_{DD}/2$.
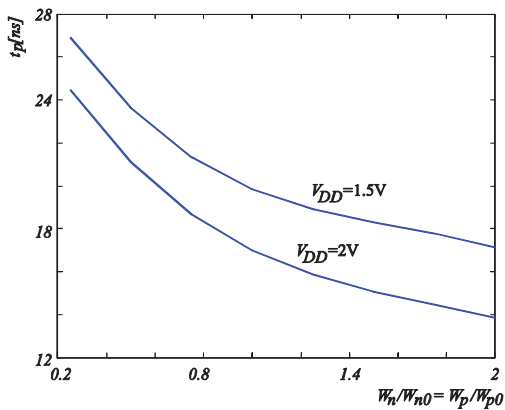
Fig. 6. Propagation time versus $W_n/W_{n0}= W_p/W_{p0}$, with capacitive load $C_L=5$pF and $V_{DD}=2$V and $V_{DD}=1.5$V.

IV. CONCLUSION

Application of the Schmitt trigger in Fig. 2 is recommended since it has better static stability and less propagation time. Reliability of operation is ensured in a wide range of supply voltages. Simulation shows that the circuit operates even at $V_{DD}=0.7$V. However, the influence of bipolar transistors is neglible for $V_{DD}<1$V, so appropriate CMOS solutions are recommended for this area of V_{DD}. Voltage hysteresis, besides on supply voltage and MOS transistors threshold voltages, depends on geometry ratios of input cascode MOS transistors and feedback ones. Changing this ratio from 0.3 to 2, voltage hysteresis can be adjusted within limits from $0.2V_{DD}$ to $0.54V_{DD}$. However, ratio 1 is optimal choice among two opposite demands: less propagation time and greater voltage hysteresis, where $V_H=0.25V_{DD}$.

REFERENCES

[1] Dokić, B., "*CMOS NAND and NOR Schmitt circuits*", Microelectronics Journal 27, Vo. 8, November 1996, pp. 757-766.

[2] Zhige Zou, Xuecheng Zou, Dingbin Liao, Fan Guo, Jianming Lei and Xiaofei Chen: '*A Novel Schmitt Trigger with Low Temperature Coefficient*', Circuits and Systems, 2008. APCCAS 2008. IEEE Asia Pacific Conference on, Nov. 30 2008-Dec. 3 2008 pp. 1398 – 1401.

[3] Fei Yuan: '*A High-speed Differential CMOS Schmitt trigger with regenerative current feedback and adjustable hysteresis*', Journal Analog Integrated Circuits and Signal Processing, Volume 63 Issue 1, April 2010, pp. 121-127.

[4] V. A. Pedroni: '*Low-voltage high-speed Schmitt trigger and compact window comparator*', Electronic Letters, 27th October, 2005, Vol. 41, No. 22.

[5] B. Dokić: '*CMOS Schmitt triggers*', IEE Proceedings, Part G, Circuits, Devices, Systems, 131 (5), 1984, pp. 197-202.

[6] *BSIM MOSFET Model User's Manual*, Tanvir Hasan Morshed, Darsen D. Lu, Wenwei (Morgan) Yang, Mohan V. Dunga, Xuemei (Jane) Xi, Jin He, Weidong Liu, Kanyu, M. Cao, Xiaodong Jin, Jeff J. Ou, Mansun Chan, Ali M. Niknejad, Chenming Hu, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720.

Pattern-Based Approach to Current Density Verification

Vazgen Melikyan, Eduard Babayan, Ashot Harutyunyan

Abstract - Methodology of static verification of current density based on layout patterns common in IC designs proposed. The methodology is based on pre-calculation of current density distribution for common layout patterns to use the obtained data to calculate current densities of large circuits partitioning them by selected patterns. The presented experimental results show the effectiveness of the approach.

Keywords – Current density, electromigration, verification, patterns.

I. INTRODUCTION

With increasing technology scaling, physical effects consideration and their impact priorities have changed. In particular, impact of electromigration (EM) increases [1-4]. EM is the mass transport in a conductor due to the momentum transfer between conducting electrons and diffusing metal atoms [1]. Electromigration damages interconnects as amounts of matter leaving and entering a given volume are unequal, the associated accumulation or loss of material results in damage [1]. When atomic flux into a region is greater than the flux leaving it, the matter accumulates in the form of a hillock. If the flux leaving the region is greater than the flux entering, the depletion of matter ultimately leads to a void (Fig. 1) [2].

Obviously, EM results in failure of IC which can be result not only of break or short-circuit, but also a significant increase in the interconnect resistance.
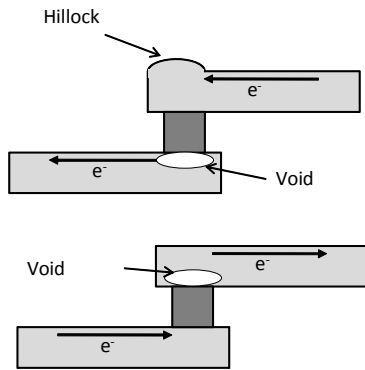
Fig. 1. Hillocks and voids

Vazgen Melikyan and Eduard Babayan are with the Educational Department of Synopsys Armenia CJSC, 41 Arshakunyants ave., Yerevan, Armenia, E-mail: {vazgenm,edbab}@synopsys.com.

Ashot Harutyunyan is with Microelectronics Circuits and Systems of State Engineering University of Armenia, 105 Teryan st., Yerevan, Armenia, E-mail: harash@seua.am.

EM is defined as [3]:

$$ J = -\frac{N_A}{kT} D_0 e^{-\frac{Q}{kT}} eZ^* \rho j \qquad (1) $$

where N_A – density of atoms in the crystal lattice; D_0 – diffusion coefficient; Q – activation energy; eZ^* - resulting charge; ρ - resistivity; k – Boltzmann constant; T – absolute temperature; j – current density.

During IC design it is required to check design against EM possibilities. As it is seen from Eq. (1), EM possibility check can be done by checking current density against maximum allowable current densities. Currently there are different current density verification EDA tools by different vendors. These tools have common disadvantages: they work only on chip level, require additional extraction and simulation steps and large amount of background information, lack error correction, etc. [4,5]

This paper presents methodology of creation of current density verification tool based on common layout patterns which enables high verification performance without need of additional design steps.

II. METHODOLOGY

It is proposed to select common layout patterns (LP), taking into account the frequency of their use in real ICs and relative areas covered by them statistically (Table 1). According to these criteria, the following LPs, shown in Fig. 2 and Table 1 were chosen for modelling.
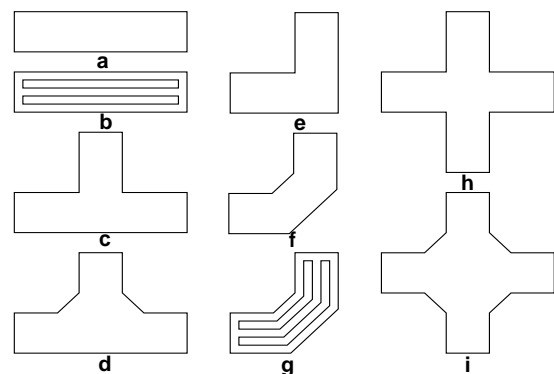
Fig. 2. Common layout patterns selected for modelling

Simulation of these patterns enables automatic estimation of maximum current density in these patterns depending on their geometrical parameters.

TABLE I
STATISTIC DATA FOR PATTERNS SELECTION

LP	IC A		IC B		IC C		IC D	
	Area, %	Count	Area, %	Count	Area, %	Count	Area, %	Count
a.	39	16643910	36	20582714	40	7374407	31	6583941
b.	10	9283	14	16700	0	0	0	0
c.	4	56859	2	13269	9	170152	5	8195
d.	0	0	4	2582	0	0	0	0
e.	16	341139	4	840437	11	166584	15	225007
f.	0	0	12	38474	2	39033	0	5
g.	1	28	1	56	0	0	0	64
h.	3	65401	3	26780	6	95161	2	16339
i.	0	0	5	1341	0	0	0	0
Total	73		81		68		53	

For LP selection current density values in the direction of the normal were taken as boundary conditions. The dependence of the maximum current density on the boundary conditions and geometric parameters of the model was calculated.
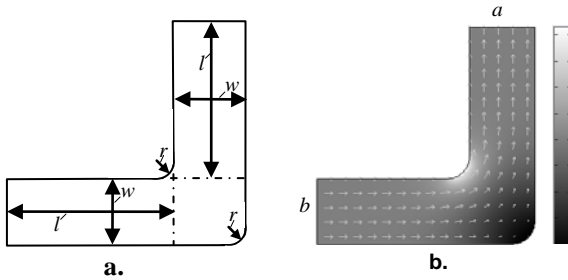
Fig. 3. For modelled LP: a – parameters;
b – current density distribution

The essence of the method is demonstrated below for the example LP in Fig. 3 is shown. In this case currents distribution is uniform in the direction of normal, equal to j_n and $-j_n$ for edges a and b respectively, and 0 for the rest. Current density distribution map shows that in the inner corner of the LP current is thickened, and on the outside, on the contrary, is diluted. Simulation was performed to identify patterns of current distribution for non-uniform boundary conditions.

For edge a of LP in Fig. 4.a, a boundary condition of uniform current distribution $j_n=1$ was set, and current distribution for edge b is shown on Fig.5b and it is mostly concentrated in upper corner. Current density reduces near upper corner and increases near bottom at a distance from edge b. In the middle of the straight segment the densities of these currents are most close to each other (Fig. 5c).

It was found out that with the increase of length of LP branches, the largest and smallest values in Fig.4c tend to 1. Consequently, it can be assumed that when length of
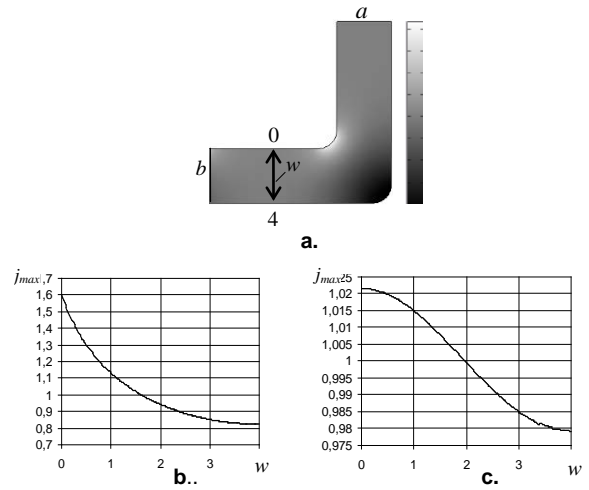
Fig. 4. Uneven distribution of boundary currents

branches l decreases in the considered model, the impact boundary conditions distribution on the largest value of current density (LVCD) decreases. This enables to neglect boundary conditions distribution and its impact on current density distribution. The calculation of the boundary currents distribution leads to solution of differential equations. It is required to find a minimum length of branches l_{min} such that for the lengths of the branches above it, the relative difference between the maximum values of current density model does not exceed the specified error ε at all possible r and w.

The length l_{min} should be found for boundary conditions, which can be assumed the worst from a practical point of view, i.e. for other boundary conditions for the same values of ε smaller values of l_{min} are obtained. As a result of investigations for the considered LP structure shown in Fig. 5 was chosen, which provides the worst boundary conditions.

With the help of the chosen structure, the dependence of l_{min} on the radius r and width w was found (Fig. 5).

An experiment has been made to find the significance of changes of maximum values of current density, depending on the lengths l_1 and l_2 larger than l.

Given that with decrease of length l impact of boundary conditions on LVCD increases, the value of l was chosen as small as possible $l = 3 \cdot r$ for experiment (this is the smallest, because at $l = 2 \cdot r$ the interior edges are equal to zero (Fig. 5))

The value of w was selected equal to l. This value of w can be viewed as worst practically, because with increase of w the impact of boundary conditions on LVCD increases and its error ε is practically unacceptably large. Obviously, the experimental results do not depend on the value of r. Taking $r=1$, $l=3$ and $w=3$ values can be obtained. Due to imposition of branches of values of l_1 and l_2 cannot exceed l. Thus, the value of one of them is fixed and only the value of another changes. Based on the dependence of LVCD on l_1 obtained through the experiment, it can be

concluded that for l_1 nearly equal l an LVCD is obtained which for larger l_1 is less than LVCD by no more than 1% (Fig. 6).
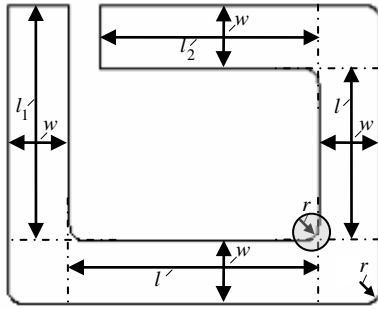
Fig. 5. The selected structure to obtain l_{min}

To find the dependence of LVCD from simultaneous change of l_1 and l_2, first l_1 changed in the range less than l (Fig. 7a), then both changed (Fig.7b). Thus it can be stated that the values found for values larger than l will not change with increase of l_2. Values of l_1 and l_2 can be taken equal to l during calculation of LVCD dependence on considered model parameters.

In the result of experiments it was found out that for worst selected values of $l/w=1.5$, the relative difference of obtained LVCD values is *0.5...0.6%* compared to values obtained for values larger than l (Fig. 8).

For considered LP, with the condition of $l/w≥1.5$ experiments were implemented to find the dependence of current densities on parameters r and w.

In the result of previous experiment it was obtained that j_{max} does not depend on w and r and vice versa; thus it can be expressed as:

$$j_{max}=f(r) \cdot \varphi(w) \cdot j, \qquad (2)$$

where

$$j=I/w \qquad (3)$$

representing current density in uniform area. Thus it is the boundary condition for those edges of the considered LP, which have nonzero current flowing in the direction of normal.
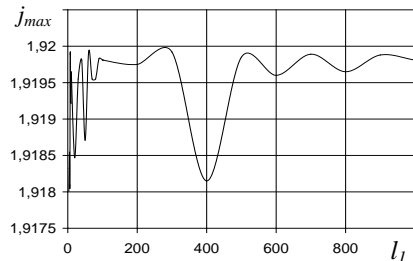
Fig. 6. Dependence of j_{max} on l_1, for $l_1 \geq l$
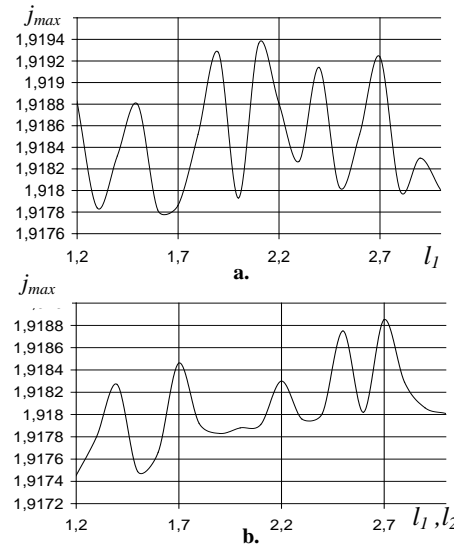
Fig. 7. Dependence of j_{max} a – on l_1, for $l_1 \leq l$, b – on l_1 and l_2 ($l_1 \leq l$, $l_2 \leq l$)
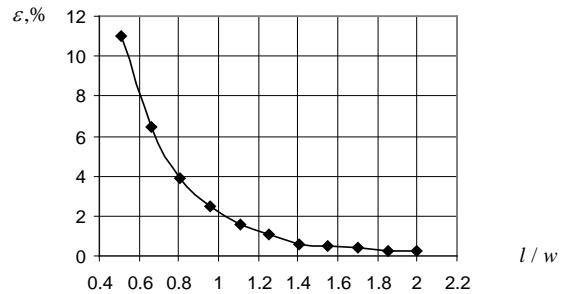
Fig. 8. Dependence of error ε on l/w

To obtain functions f and φ two experiments were implemented resulting in dependencies of jmax on w (Fig. 9a) and r (Fig. 9b) with fixed value of another variable.
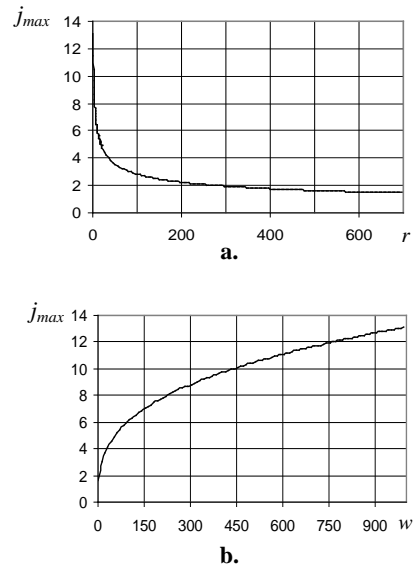
Fig. 9. Dependence of j_{max} on a – w, b – r

In the result of approximation of dependence function, the following was obtained for considered LP:

$$j_{max} = (13,2 \cdot r^{-0.33} - 0,06368) \cdot$$
$$\cdot (0,09743 \cdot w^{0.3365} + 0,0005986) \cdot j \qquad (4)$$

Using expressions Eqs. (2) and (3), for w this is obtained:

$$\frac{0,09743 \cdot w^{0.3365} + 0,0005986}{w} = \frac{j_{max}}{I \cdot (13,2 \cdot r^{-0.33} - 0,06368)} \qquad (5)$$

The general flow of developed method of current density verification is presented on Fig. 10.

Layout Decomposition
decomposition of electrical elements (EE), layout partition ofinterconnects

Maximum Current Density Calculation
Current density calculation in EEs and interconnects

Coarse Verification
Net width checks, via count checks

Fixing Violations
Net widening, redundant via placement, etc.

Partition to LPs
Partition oflayout to a preselected patters

Calculation of LVCD in LPs
Calculation of LVCD using found dependencies

Providing Errors
Reporting errors to user

Fig. 10. General current density verification flow

Experimental software implementing the proposed method was developed. Unlike industrial software, it does not need additional extraction and simulation steps. Experimental results are shown in Table 2.

TABLE II.
COMPUTER TIME AND MEMORY REQUIRED
TO OBTAIN THE CURRENT DENSITY DISTRIBUTION

Parameters	Circuit 1	Circuit 2	Circuit 3	Circuit4
Time, s	0.125	0.391	1.734	7.984
Memory, kB	1.3	8.7	28.9	97.4

For a circuit with 50000 LPs, 104 minutes were required for calculation with conventional software, whereas with proposed method it took only ~10 minutes.

III. CONCLUSION

The developed method of current density verification in ICs and the experimental software package have indisputable advantages over existing similar tools and meet practical requirements of modern IC design.

ACKNOWLEDGMENTS

REFERENCES

[2] H. Ceric, S. Selberherr, "Electromigration in submicron interconnect features of integrated circuits", Materials Science and Engineering, 2011, pp. 53-86

[2] L. Xiaoyu, S. Jiang, W. Yun, Z. Chenhui, "Research on failure modes and mechanisms of integrated circuits", Prognostics and System Health Management Conference (PHM-Shenzhen), 2011, 2011, pp. 1-3.

[3] M. Shao, et al., "Current calculation on VLSI signal interconnects", Sixth International Symposium on Quality of Electronic Design, (ISQED), 2005, pp. 580-585.

[4] B. Li, et al. "Statistical Evaluation of Electromigration Reliability at Chip Level", IEEE Transactions on Device and Materials Reliability, 2011, p. 1-11

[5] J.P. Gambino, T.C. Lee, F. Chen, T.D. Sullivan, "Reliability challenges for advanced copper interconnects: Electromigration and time-dependent dielectric breakdown (TDDB)", 16th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits, 2009, pp. 677-684.

TBT Signal Model for Improved Accuracy of High-level Dynamic Power Estimation Procedure

Bojan Jovanović, Ružica Jevtić, and Carlos Carreras

Abstract - When estimating the dynamic power consumption of DSP datapaths, it is crucial to accurately calculate switching activity produced inside the design. For accurate switching activity calculation the existence of an appropriate data signal model is essential. This paper presents a triple-bit type (TBT) signal model which is used to represent bit-level switching activity at the output of multipliers. The model depends on word-level signal statistics and the number of multiplied input signals. For the sake of comparison with the standard dual-bit type (DBT) signal model, both models (TBT and DBT) are applied to the high-level power estimation of three reference designs implemented in FPGA. Calculated with respect to the measured power, the relative errors of here presented TBT model are four to five times smaller than the errors of the DBT model.

Keywords - Power estimation, TBT signal model, FPGA.

I. INTRODUCTION

Due to increased density of ICs, when the number of transistors per unit area reached a critical point, heat dissipation and consequently power consumption became another parameter (beside speed and area) that VLSI designers must be aware of. If it is not properly optimized during the design phase, power consumption could cause excessive heat demanding increasingly expensive packaging and cooling strategies which might, either add significant cost to the system, or provide a limit on the amount of functionality. In the process of power optimization it is extremely important to have the tools for fast and accurate estimation of power consumption. With such tools, expensive and time consuming iterative physical implementations of the system could be avoided. Furthermore, applying power estimation techniques we could explore a large number of different system architectures to find (after a few iterations) the one with the lowest power consumption. Having in mind that higher levels of design abstraction have the largest power reduction opportunities as well as the shortest power analysis iteration times (between seconds and minutes) [1], we present the TBT signal model which is used for the calculation of bit-level switching activity and applied on a high-level power estimation procedure. The advantages of this signal model were for the fist time briefly reported in

Bojan Jovanović is with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail:bojan@elfak.ni.ac.rs.

Ruzica Jevtić and Carlos Carreras are with the Dept. of Electronics Engineering, ETSIT, Technical Univ. of Madrid, 28040 Madrid, Spain, E-mail:{ruzica,carreras}@die.upm.es.

[2].

The paper is organized as follows. In Section II some power estimation approaches are discussed. The TBT signal model is introduced in Section III, while experimental results are reported in Section IV. Section V summarizes the conclusions.

II. POWER ESTIMATION: STATE OF THE ART

The first EDA software packages for the automation of the IC design process were equipped with various tools intended for the simulation (prediction of the IC behaviour) and analysis of circuit performance: speed, occupied area, detectability of faults etc. As power consumption has become a more and more important issue, many EDA packages are now including tools for its estimation.

The models for power estimation differ in the nature of the power they are trying to estimate (static, dynamic, short-circuit or total power consumption), as well as in the level of abstraction of the target designs. The higher the level of abstraction, the faster the estimations. However, short estimation times at higher leves usually imply less accuracy of the estimates. Roughly, there are two different approaches to address the problem of power estimation: statistical and probabilistic [3].

The statistical approaches simulate the circuit with input vectors and collect statistical data for each node in the circuit. The simplest statistical techniques for power estimation are presented in [4, 5, 6, 7]. They are accurate but memory and time consuming (especially for large circuits) as well as pattern-dependent. In order to cope with the pattern dependence problem, some statistical approaches based on Monte Carlo simulation are presented in [8, 9, 10]. Under the assumption that the power consumed by the circuit over a long period T has a normal distribution, the technique applies randomly generated input patterns to the circuit primary inputs and monitorizes the power dissipation per time interval T.

On the other hand, probabilistic methods analyze the circuit and generate the expressions for the signal probabilities propagated through the circuit [3, 11, 12, 13]. Hence, they do not depend on the number of input data vectors but only on their statistics. These methods also have problems when analyzing large circuits, as the complexity of the analytical expressions depends on the number of inputs and the logic depth of a circuit.

Some unconventional approaches to estimate power consumption can be found in [14]. For a low-level

estimation technique the author uses an in-house program language (AleC++) and a simulator (Alecsis) to extract the total switching capacitances of the circuit. For each logic element inside the design it is necessary to have a low-level model with the informations about its capacitances. The low level nature of the model makes it slow for the analysis of large designs. Another approach, also presented in [14], is based on the integration of the supply current waveform. Since it is the most accurate technique for total power consumption estimation, the author proposed using a three-layer neural network to model the area of the supply current impulse.

Power estimation techniques for static and short-circuit power consumption are described in [15] and [16], respectively.

In the FPGA arena, existing power estimation techniques aim to represent power consumption in the form of an equation. Variable parameters in the equation depend on the various factors (input and output signal statistics, operand word-lengths, circuit fanout, component structure etc.). Some approaches for FPGA power estimation are presented in [17, 18, 19, 20, 21]. The reported power estimation errors are in the range between 10% and over 30%. While some of them are not compared with the real measured power values [17], the other are extremely time consuming (up to the 12 hours) [18, 21] or require long calibration procedures [19, 20].

Finally, there are a few tools designed for commercial FPGAs. The most widely used are XPower from Xilinx [22] and PowerPlay from Altera [23]. These tools provide a detailed power breakdown of a design based on the resource capacitance and utilization as well as data switching activity. In their early versions the tools had limited accuracy. Large errors were detected when the estimates were compared to physical measurements [19]. Later versions are becoming more and more sophisticated and accurate. Additional problems are encountered when complex designs with many signals are to be modelled, as these tools require large amounts of memory and long execution times.

In the rest of this paper we will focus on the high-level dynamic power estimation based on a probabilistic approach.

III. TBT VS DBT SIGNAL MODEL

For the estimation of dynamic power consumption we use the general approach described in [24] as well as widely known expression for CMOS gate dynamic power:

$$P = Vdd^2 \cdot f \cdot C_l \cdot SW = a \cdot SW \qquad (1)$$

where *SW* is the total switching activity produced inside the the design and constant *a* represents the product of three power terms: squared power supply (known for a specific FPGA architecture), clock frequency (fixed for a specific design), and load capacitance *Cl* which is, assumed to be

constant due to regular FPGA structure as in [17]. The constant *a* is obtained empirically in the process of calibration, through a small number of low-level power measurements. The switching activity is computed analytically as it will be explained below.

The switching activity is determined by the present and immediately-past value of a signal. If they are different the switching activity has occurred. In order to calculate the total switching activity of a design we need to start from its inputs and determine the switching activity of its input bits. For this purpose, the dual-bit type (DBT) model is presented in [25]. Under the assumption that DSP input signals are stationary and with a Gaussian distribution, the DBT model calculates bit-level switching activities as functions of the input bit-widths and signal statistics: autocorrelation, variance and mean value. In Fig. 1 we have plotted the bit switching activity in a Gaussian signal word versus the bit position in the word for different autocorrelations. All the signals have a zero mean and the same variance.

There are three switching activity regions that can be clearly distinguished: the LSB region with a fixed switching activity of 0.5, the MSB region with strongly correlated data bits, and the so-called linear region that lies between the two previously mentioned ones.
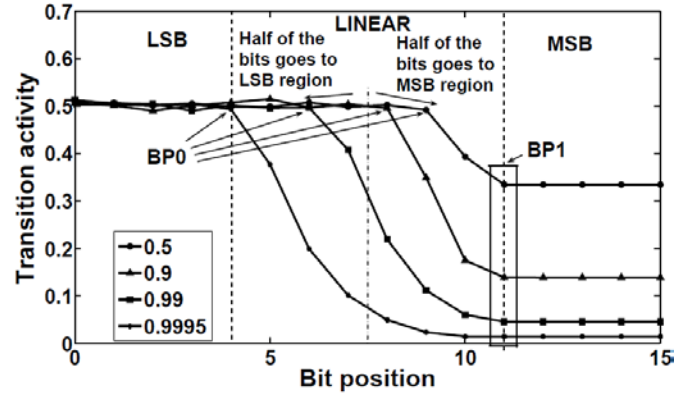
Fig. 1. Bit switching activity vs. bit position in an input word

The breakpoints that divide the regions can be obtained as:

$$BP0 = \left\lceil \log_2(\sqrt{1-\rho^2} \cdot \sigma) \right\rceil$$
$$BP1 = \left\lceil \log_2(6 \cdot \sigma) \right\rceil \qquad (2)$$

The switching activity of the MSB bits (sw_{MSB}) is calculated by knowing its dependency on the probability of the MSB bit being '1' (p_{MSB}), as introduced in [26]:

$$sw_{MSB} = 2 \cdot p_{MSB} \cdot (1 - p_{MSB}) \cdot (1 - \rho) \qquad (3)$$

Once the bit-level input switching activities are known , the switching activity generated inside the component can be easily obtained. For this purpose, the probability method

presented in [24] is used. The approach takes the input bit switching activities and computes the switching parameters of the output and carry bits of the design's components through probabilistic formulas obtained from truth tables of the component's basic cells. Multiplying the estimated switching activity (obtained as the sum of switching activities of all nodes inside the design) by the previously determined constant *a* we obtain the estimated value for the design's dynamic power consumption.

The DBT signal model, however, has proven to be inefficient in modelling the bit-level switching activity at the output of some non-linear DSP designs. The binary multiplier is the typical example of such a design. It has been noted that the output of the multiplier has a distribution that is symmetrical around the mean value but it is not a Gaussian one [27]. The LSB bit of the product exhibits less switching activity than 0.5 because only the product of odd numbers is odd. This is confirmed in Fig. 2 where the bit-level switching activity at the multiplier output is plotted. A new LSB1 signal region containing the LSB bits is clearly noticeable. This region tends to grow as the number of chained multiplications grows. The number of bits affected by the multiplication (breakpoint BPm) is equal to *2 x nm*, where *nm* is the number of multiplied Gaussian processes.
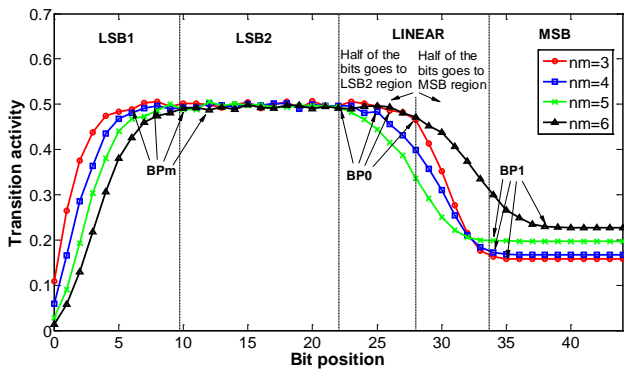
Fig. 2. Bit-level switching activity at the multiplier output

As the multiplier is a common data-path operator in the hardware implementation of many modern DSP designs (exponential, logarithmic, square root, reciprocal functions, FIR and IIR filters, FFTs etc.), power estimates of these designs would be inaccurate if the DBT signal model is used. Consequently, in this paper we present a new TBT signal model which takes into account the LSB1 signal region approximating its exponential-like dependence with the following equation:

$$sw(i) = 0.5 - (0.5 - sw(0)) \cdot e^{-(0.25 + 2^{(-nm+1.25)}) \cdot i} \qquad (4)$$

where *sw(0)* is the switching activity of the LSB bit, which is obtained according to the formulas given in [27], *i* is the bit position, and *nm* is the number of Gaussian processes that have been multiplied up to this point. The switching activity of the rest of the bits, as well as the breakpoints

BP0 and BP1 are obtained according to the DBT method. Fig. 3 shows that the proposed approximations match well with the actual switching activities.
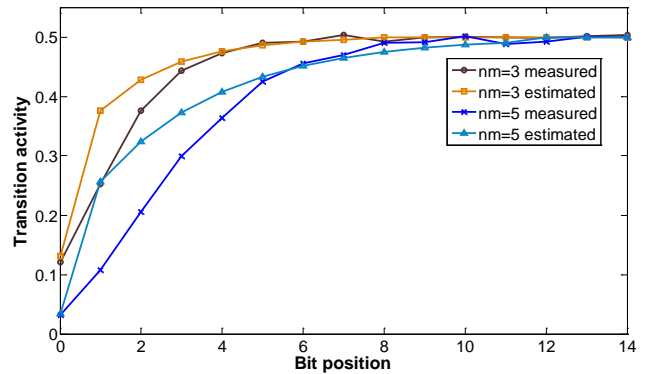
Fig. 3. Actual (BLT) vs. estimated (TBT) switching activities in the LSB1 region

For the evaluation of the TBT signal model, the switching activities produced inside four reference designs with different number of multipliers have been measured and compared with the switching activities obtained when applying TBT and DBT signal models. The results are reported in Table I.

TABLE I
RELAVITE ERRORS BETWEEN THE SUM OF MEASURED AND ESTIMATED SWITCHING ACTIVITIES

nm	3	4	5	6
ErrDBT[%]	4.20	6.29	6.78	10.48
ErrTBT[%]	0.87	1.07	-0.17	2.08

It is obvious that a greater number of the multipliers in the design increases the relative error of the DBT signal model making it more inefficient.

The impact of the TBT signal model on power estimation will be the subject of the next section.

IV. EXPERIMENTAL RESULTS

Several DSP designs have been used in the experimental set. On the one hand, each DSP design is implemented in a Virtex-II Pro XC2VP30 FPGA chip and the design power is measured as described in [28]. In brief, the on-board power measurement system consists of two boards: one with a Xilinx Virtex-II Pro FPGA device, used for measuring the power, and another with an Altera Strarix FPGA device used for loading the input vectors to the first one. In this way, designs implemented in the Virtex-II device are stimulated externally so there is no additional power caused by vector generation that can influence the measured power value. As a result, the measured power corresponds to the static power plus the dynamic power of logic, interconnections and clock. To extract the measured dynamic power consumption value we repeat the following procedure. First, we measure the static power when no input stimuli nor clock are applied. Then, we measure the

clock power together with the static power (all zeroes are applied to the inputs). Finally, we measure the power when various inputs with Gaussian distributions are applied. When we subtract the clock and static power from the total power the result is the power of logic and signals. From this power value we subtract the power of global connections using an in-house C++ program (MARWEL) [29]. This program extracts the lengths of the interconnects from the Xilinx design files and allows for the computation of the power consumption of global connections. The result is the measured dynamic power consumption of the design.

On the other hand, for the same DSP design, we apply the model for dynamic power estimation described in [24]. Switching activities produced inside the design (see Eq. 1) are calculated using the TBT and DBT signal models as well as using the actual bit-level switching activities of the component inputs (BLT). Measured and estimated dynamic power values are then compared to obtain the relative error of the applied model for power estimation. The estimated power obtained from XPower tool (ISE 10.1) is also included in the comparison.

The evaluation set consists of three different DSP designs. The first two are relatively small and correspond to the implementation of the following logical functions:

$$DSP_1 = (x_2 \cdot x_3) \cdot x_2 + (x_1 + x_3) \cdot x_2$$
$$DSP_2 = ((x_1 + x_2) \cdot (x_3 + x_4) + x_1 \cdot x_2) \cdot x_2 \cdot (x_3 + x_4) \quad (3)$$

The DSP_3 design is quite larger and has a structure that reminds one of a 16-tap digital FIR filter implemented as a cascade realisation of eight second-order sections like the one presented in Fig. 4. All three DSP designs are synchronous. The clock frequency for the first two designs is 50MHz. Keeping in mind the complexity of the DSP_3 design, the clock frequency for it is set to only 16MHz in order to keep the static power constant. Table II shows the results for each design when data with different autocorrelation coefficients, ρ, are applied to its inputs.
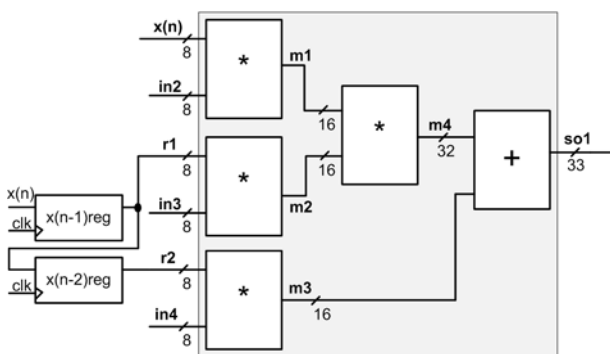
Fig. 4. Second-order section of DSP_3 design

The first two columns show the number of occupied slices for each DSP design as well as the number of embedded multipliers used in the design. The computation times for each DSP design are listed in the next column, followed by the autocorrelation coefficients and the relative errors obtained for each model.

TABLE II
RELATIVE POWER ESTIMATION ERRORS FOR THREE SIGNAL MODELS
(BLT, TBT, DBT) AND XPOWER (XPw)

Bench-mark	Slices	Emb. mult.	Comp. time[s]	ρ	Er(BLT) [%]	Er(TBT) [%]	Er(DBT) [%]	Er(XPw) [%]
DSP_1	212	2	0.92	0	10.3	7.6	17.48	328.79
				0.9	6.94	1.49	6.62	316.48
				0.99	9.33	7.6	9.59	281.70
				0.9995	9.49	8.71	13.68	246.91
DSP_2	192	2	1.1	0	7.92	4.5	11.58	258.45
				0.9	7.51	1.03	7.6	233.50
				0.99	12.18	7.81	11.54	216.23
				0.9995	22.24	21.24	30.99	245.27
DSP_3	2977	8	91.95	0	-0.38	9.91	38.3	455.08
				0.9	-1.55	8.99	37.38	455.06
				0.99	-1.56	10.05	38.24	437.46
				0.9995	-0.78	15.25	41.27	442.62

The greater complexity of the DSP_3 design is confirmed by the number of occupied slices as well as by the computation time needed for its power estimation. Considering relative errors, we can conclude that the TBT model gives far better power estimations (four to five times) than the DBT model, for all DSP designs and for all autocorrelation coefficients. Mean relative errors in power estimation of all three DSP designs for the TBT and DBT models are equal to 8.68% and 22.05%, respectively. Furthermore, the TBT model achieves the biggest improvements with respect to the DBT model in the case of the DSP_3 design. This can be explained by the fact that designs DSP_1 and DSP_2 consist of several adders and multipliers and have just a few bits in the LSB1 zone, so the effect of using TBT instead of DBT is barely noticeable. The number of adders and multipliers in the DSP_3 design is greater, which contributes to the increase in the number of LSB1 bits, so the impact of using a more accurate signal model is more obvious. When analyzing the XPower tool relative errors we can confirm the claims reported in [19] about the large estimation errors of such a tool for small designs in comparison with physical power consumption measurements.

V. CONCLUSION

We have presented the TBT signal model intended for the bit-level switching activity calculation as well as for the integration in high-level probabilistic approaches in dynamic power estimation. Unlike the previously used DBT signal model, it takes into account non-linearities produced at the output of some DSP circuits and introduces a new switching activity region for the LSB bits. The proposed model is not pattern-dependent. It depends only on input signal statistics and bit-widths as well as on the number of prior multiplications inside the design. The validity of the TBT signal model has been confirmed

through on-board dynamic power consumption measurements. Furthermore, in comparison with the DBT model, relative errors of the estimations are quite lower (four to five times), especially when estimating larger designs with more non-linear DSP circuits.

REFERENCES

[1] Raghunathan, A., Jha, N., Dey, S., *"High-level Power Analysis and Optimization"*, Kluwer Academic Publishers, Massachussets, 1998.

[2] Jovanovic, B., Jevtic, R., Carreras, C., *"Triple-bit method for power estimation of nonlinear digital circuits in FPGAs"*, Electronics Letters, Vol. 46, No. 13, June 2010, pp. 903-905.

[3] Machado, F., *"Switching Activity Analysis of Digital Electronic Circuits described at RTL using Probabilistic Techiques. Proposal of an Estimation Method"*, PhD Thesis, Technical University of Madrid, 2008.

[4] Deng, C., *"Power analysis of CMOS/BiCMOS circuits"* In. Proc. of the Int. Workshop on Low Power Design, Apr. 1994, pp. 3-8.

[5] Landman, P., *"Low-Power Architectural Design Methodologies"*, PhD Thesis, Electronic Research Laboratory, Univ. of California, Berkley, Aug. 1994.

[6] Schneider, P., *"PAPSAS: A Fast Switching Activity Simulator"*, PATMOS'95, pp. 351-360.

[7] George, B., *"Power Analysis for Semi-Custom Design"*, CICC, New York 1994, pp. 249-252.

[8] Burch, R., Najm, F., Yang, P., Trick, T., *"A Monte Carlo approach for power estimation"*, IEEE Trans. on VLSI Systems, No. 1, Vol. 1, Mar. 1993, pp. 63-71.

[9] Todorovich, E. et al., *"A Tool for Activity Estimation in FPGAs"*, LNCS, June 2002, pp. 340-349.

[10] Todorovich, E., Boemo, E., *"Statistical Power Estimation for FPGAs"*, LNCS, June 2005, pp. 515-518

[11] Cirit, M., *"Estimating Dynamic Power Consumption of CMOS Circuits"*, Proc. ICCAD, Nov. 1987, pp.534-537

[12] Chou, T., Roy, K., Prasad, S., *"Estimation of Circuit Activity Considering Signal Correlations and Simultaneous Switching"*, Proc. of the IEEE/ACM Int. Conf. on CAD, June 1994, pp. 300-303.

[13] Machado, F., Riesgo, T., Torroja, Y., *"Disjoint Region Partitioning for Probabilistic Switching Activity Estimation at Register Transfer Level"*, PATMOS, Sep. 2008, pp. 1145-1148.

[14] Maksimovic, D., *"Logical Simulation – An estimation of limit properties of designed digital circuit"*, PhD Thesis, University of Nis, June 2000.

[15] Hussam, H., Dhamin, K., Come, R., *"Static Power Estimation of CMOS Logic Blocks in a Library Free Design Environment"*, Int. Journ. of Design, Analysis and Tools for Circuits and Systems, Vol. 1, No. 1, June 2011, pp. 41-52

[16] Nose, K., Sakurai, T., *"Analysis and Future Trends on Short-Circuit Power"*, IEEE Trans. on CAD of ICs and Systems, Vol. 19, No. 9, Sep, 2000, pp. 1023-1030.

[17] Chen, D., Cong, J., Fan, Y., *"Low-Power High-Level Synthesis for FPGA Architectures"*, In Proc. of ISLPED, Aug. 2003, pp. 134-139.

[18] Choi, S., Jang, J., Mohanty, S., Prasanna, V., *"Domain-Specific Modeling for Rapid Energy Estimation of Reconfigurable Architectures"*, The Journ. of Supercomputing, Vol. 26, No. 3, Nov. 2003, pp. 259-281.

[19] Elleouet, D., Savary, Y., Julien, N., *"An FPGA Power Aware Design Flow"*, PATMOS, Sep.2006, pp. 415-424

[20] Abdelli, N., Fouilliart, A., Julien, N., Senn, E., *"High-Level Power Estimation on FPGA"*, IEEE Symp. on Industrial Electronics, June 2007, pp. 925-930.

[21] Anderson, J., Najm, F., *"Power Estimation Techniques for FPGAs"*, IEEE Trans. on VLSI Systems, Vol. 10, No. 12, Oct. 2004, pp. 1015-1027.

[22] ftp://ftp.xilinx.com/pub/documentation/tutorials/xpowerfpgatutorial.pdf

[23] http://www.altera.com/literature/hb/qts/qts_qii53013.pdf

[24] Jevtic, R., Carreras, C., Caffarena, G., *"Fast and Accurate Power Estimation of FPGA DSP Components Based on High-level Switching Activity Models"*, Int. Journ. of Elec. Vol. 95, No. 7, July 2008, pp. 653-668.

[25] Landman, P., Rabaey, J., *"Architectural Power Analysis: The Dual Bit Type Method"*, IEEE Trans. on VLSI Systems, Vol. 3, No. 2, Mar. 19995, pp. 173-187.

[26] Ramprasad, S., Shanbhag, N., Hajj, I., *"Analytical Estimation of Signal Transition Activity from Word-level Statistics"*, IEEE Trans. on CAD of ICs and Systems, Vol. 16, No. 7, 1997, pp. 718-733.

[27] Bitzaros, D., Nikolaidis, S., *"Estimation of bit-level transition activity in data-path based on word-level statistics and conditional entropy"*, IEE Proc. Circuits Devices Syst., Vol. 149, 2002, pp. 234-240.

[28] Jevtic, R., Carreras, C., *"Power measurement methodology for FPGA devices"*, IEEE Trans. Instrum. Meas., Vol. 59, No. 9, June 2010, pp. 237-247.

[29] Jevtic, R., *"High-Level Power Estimation of DSP Circuits Implemented in FPGAs"*, PhD Thesis, Technical Unversity of Madrid, 2009.

Analysis and design of a two-stage CMOS operational amplifier in 150 nm technology

B.Sc. Nikola Ivanišević, Ph.D. Mirjana Videnović-Mišić, M.Sc. Alena Đugova

Abstract - This paper describes the topology of the two-stage CMOS operational amplifier, optimum performance design guidelines and the results of simulations done in Cadence Spectre simulator using the 150 nm technology. The designed operational amplifier is for low-voltage applications and achieves performance in the range of commercially available op amps such as: open loop gain of 75 dB, unity frequency of 50 MHz, 65° phase margin, 0.3 mW power dissipation for V_{DD}=1.8V, slew rate of 10V/μs, maximum settling time 150ns.

Keywords - CMOS, two-stage, operational amplifier.

I. INTRODUCTION

The current market demands for all-in-one devices have led to the development of mixed signal circuits, combining their digital processing power and analogue interface with the real world. General aim for digital circuits is making them smaller, faster and more power efficient. The power supply decreases as a direct result of CMOS scaling and at the same time this result downgrades the performance of analogue circuits. Sub-micrometer technologies, dominantly used for digital designs, create transistors in analogue designs with low output resistance. As a consequence those transistors have poor amplifying performances. Moreover, the power supply voltage decrease results in lower gate-source voltage which in turn decreases mosfet transconductance. These bottlenecks force analogue designers to run many circuit simulations to achieve targeted performance. From this came the motivation to design a low voltage two-stage operational amplifier (op amp for short) in a 150 nm (mixed signal circuit) technology and to achieve performance in the range of commercially available op amp ICs.

The two-stage CMOS op amp is a classical topology which, when carefully designed, can achieve desired performances and is often used as a benchmark for modern designs. It uses two voltage amplifiers connected in cascade to achieve high differential voltage gain. There is also an optional third stage (output buffer) but is often left out of the design if the op amp drives capacitive load. This type of op amps fall into a subcategory called operational transconductance amplifiers (OTA) [1].

Nikola Ivanišević, Mirjana Videnović-Mišić and Alena Đugova are with the Department of Electronics, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia, E-mails: nikolaivanisevic@gmail.com, {mirjam, alenad}@uns.ac.rs.

II. TOPOLOGY ANALYSIS

The first stage of a two-stage op amp is the differential amplifier while the second one is the common source amplifier, shown in Fig. 1. The differential pair is pmos type which defines second stage as nmos common source amplifier. A complementary design is also possible but the one given in Fig. 1 shows higher slew rate and lower flicker noise. However, its disadvantage is higher thermal noise. The overall voltage gains for pmos-nmos or nmos-pmos op amp stage topologies are similar and proportional to the product of the individual gain of each pmos and nmos stage (cascade connection).

There are certain geometric ratios that exist in the design to ensure minimal systematic offset and immunity to variations (supply voltage, temperature and process variations). The systematic (inherent) voltage offset can be minimized if you maintain geometric ratios shown in Eq. (1) [2].

$$\frac{(W/L)_7}{(W/L)_4} = 2\frac{(W/L)_6}{(W/L)_5}. \qquad (1)$$

The transistors Q_3 and Q_4 make a current mirror. When the differential signal is zero (but there is a common mode signal to insure that Q_1 and Q_2 are turned on) the current of the transistor Q_5 splits equally into transistors Q_1 and Q_2. The current I_{D5} is also mirrored into I_{D6} and therefore there is a hidden relationship between I_{D6} and $I_{D1,2}$ that needs to be maintained. If that is not the case, due to the offset voltage and high open loop gain, the output voltage of an op amp will reach one of the supply rails.

To understand how the impact of variations is minimized it is necessary to understand how frequency compensation is gained. Small signal scheme of the op amp is shown in Fig. 2. The voltage controlled current sources represent the amplifying stages while additional passive components represent their parallel resistive and capacitive loads, expressed in Eqs. (2) to (5) [2]. The variable r_{dsi} (i=1, 2, 3…) represents the output resistance of the transistor in the small signal analysis.
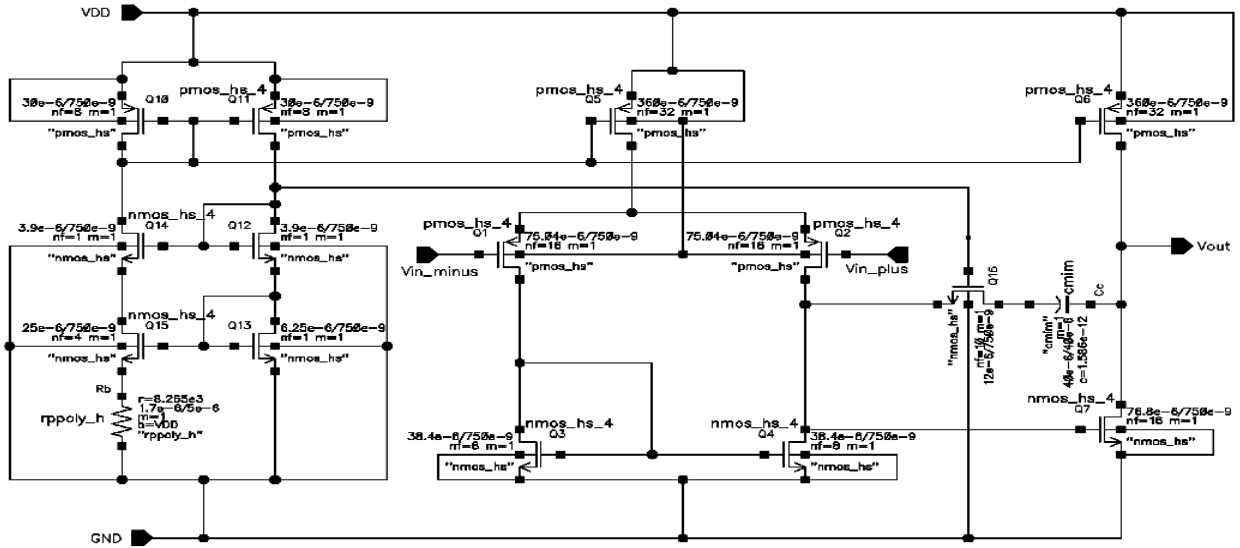
Fig. 1. Schematic of the two-stage op amp circuit.

$$a = (C_2 + C_c)R_2 + (C_1 + C_c)R_1 + g_{m7}R_1R_2C_c. \quad (7)$$

$$b = (C_2C_c + C_1C_c + C_1C_2)R_1R_2. \quad (8)$$

Fig. 2. A small signal model of the two stage op amp used for compensation analysis [2].

$$\omega_{p1} \approx \frac{1}{R_1[C_1 + C_c(1 + g_{m7}R_2)] + R_2(C_2 + C_c)}. \quad (9)$$

In the initial frequency compensation analysis, resistor R_C is neglected. If we solve this circuit and find the voltage gain (its transfer function) the Eq. (6) [2] is obtained. From it, the expressions for the first and second pole can be extracted, if the poles are real and the second pole is on a much higher frequency than the first pole, as show in Eqs. (9) and (10) [2].

$$\omega_{p2} \approx \frac{g_{m7}}{C_1 + C_2 + \dfrac{C_1C_2}{C_c}}. \quad (10)$$

It can be seen from the Eqs. (9) and (10) [2] that the first pole depends on $1/g_{m7}$ while the second is proportional to g_{m7}. With g_{m7} increase poles are further apart from one another which is possible due to capacitor C_C presence. The approach is commonly known as the pole splitting technique while C_C is called Miler's capacitance. There is also a zero in the right half of the complex plane which in terms of a phase shift acts like a left sided pole. This means that the zero inserts a negative 90° into the phase and makes the circuit unstable.

When R_C is introduced into the scheme (in Fig.1 it is Q_{16} polarized in deep triode region), the fundamental difference is in the position of the zero, now expressed with Eq. (11) [2]. A designer can now influence the zero position giving him a couple of options to achieve better op amp stability. The first option is to choose the value of R_C such that the zero is canceled which means its somewhere in the infinity. The second option is to put the zero around the second pole so that they can cancel each other out. This is rarely used because you can't tell in advance the value of C_2 especially in OTAs. The third option is to put the zero

$$R_1 = r_{ds4} \parallel r_{ds2}. \quad (2)$$

$$C_1 = C_{db2} + C_{db4} + C_{gs7}. \quad (3)$$

$$R_2 = r_{ds6} \parallel r_{ds7}. \quad (4)$$

$$C_2 = C_{db7} + C_{db6} + C_L, \quad (5)$$

$$\frac{V_{out}}{V_{in}} = \frac{g_{m1}g_{m7}R_1R_2(1 - \dfrac{sC_c}{g_{m7}})}{1 + a \cdot s + b \cdot s^2}, \quad (6)$$

where a and b are equal to:

slightly higher than the unity gain frequency when there is no resistor. The R_C value is then calculated with Eq. (12) [2].

$$\omega_z = -\frac{1}{C_c(1/g_{m7} - R_c)} . \tag{11}$$

$$R_C \approx \frac{1}{1.2 \cdot \omega_u C_C} . \tag{12}$$

This way we can increase the unity gain frequency by 20% and also the phase margin by approximately 40°. If R_C value is too high the gain and phase characteristics do not steadily decrease above unity-gain frequency which can distort the signal.

From the third option we can describe the design methodology for achieving a desired phase margin (PM). First, chose an initial value for the capacitor C_C e.g, 1 pF. Then write down the frequency and gain for which the phase is equal to 180° + 40° – PM (40° due to moved zero). The third step is to increase the value of C_C by multiplying it with the previously written gain value. After that the value of an actual resistor R_C can be calculated using Eq. (12) [2]. This step will require a series of iterations.

When you find adequate R_C value the final step is to replace the resistor with a mosfet in the deep triode region. The capacitor doesn't allow dc current through Q_{16} so V_{DS16} is always zero. Transistor Q_{16} dimension ratio can be approximately calculated with Eq. (13), which is result of circuit current mirrors and topology. Note that I_{D5} corresponds to the dc current of transistor Q_5 which is copied from the polarization circuit.

$$\left(\frac{W}{L}\right)_{16} \approx \frac{1}{R_C\sqrt{2I_{D5}\mu_n C_{OX}}\sqrt{\dfrac{(W/L)_{10}}{(W/L)_5(W/L)_{12}}}} . \tag{13}$$

The immunity to variations is achieved with the polarization circuit and by maintaining certain geometrical ratios. This means that the position of the zero, given in Eq. (11) [2], doesn't change due to variations. If $1/g_{m7}$ decreases then R_C needs to decrease its value by the same amount so that their difference remains the same. To achieve this, the expression shown in Eq. (14) [2] has to be fulfilled.

$$\frac{(W/L)_6}{(W/L)_7} = \frac{(W/L)_{11}}{(W/L)_{13}} . \tag{14}$$

The frequency stability of the op amp and stabilized values of transconductances are achieved as a result of the polarization circuit topology and resistor R_b, shown in Fig. 1. If the dimensional ratios of Q_{10} and Q_{11} are equal the

currents I_{D10} and I_{D11} will be the same. From the II KVL for the loop which consists of R_b, Q_{13} and Q_{15} dependence of g_{m13} on R_b and well defined geometry is achieved, Eq. (15) [2].

$$g_{m13} = \frac{2}{R_b}\left(1 - \sqrt{\frac{(W/L)_{13}}{(W/L)_{15}}}\right) . \tag{15}$$

All the currents in the op amp are mirrored from the polarization circuit. If the mirroring is ideal (channel length modulation is not taken into account) and body effect is neglected, the op amp transconductances values are calculated from Eqs. (16) and (17) [2].

$$g_{miPMOS} = g_{m13} \cdot \sqrt{\frac{\mu_p}{\mu_n} \frac{(W/L)_i I_{Di}}{(W/L)_{13} I_{D13}}} . \tag{16}$$

$$g_{miNMOS} = g_{m13} \cdot \sqrt{\frac{(W/L)_i I_{Di}}{(W/L)_{13} I_{D13}}} . \tag{17}$$

Where i = 1, 2, 3... corresponds to the index of the transistor in Fig. 1.

III. SIMULATION RESULTS

Simulations were done using the same length for all of the transistors to make current mirroring precise, by minimizing errors due to the side diffusion of the source and drain areas [3]. The length value was chosen so that the transistors behave like long channel devices with high output resistance. As a result, the impact of channel length modulation is decreased resulting in higher differential voltage gain. The optimum length was 750 nm or five times larger then the minimum value (150nm).

Transistor dimensions for the initial simulation were hand calculated to fulfill the targeted performances. Op amp design parameters (diff. gain, unity frequency, phase margin etc.), were chosen in the range of currently available commercial low-voltage op amps values. Targeted performance parameters are shown in Table I while the transistor widths, capacitor and transistor values of the final iteration are shown in Table II.

TABLE I
TARGETED PERFORMANCE PARAMETERS

Differential gain [dB]	≥75
Unity frequency [MHz]	≥50
Phase margin [°]	≥60
Capacitive load [pF]	=3
Slewrate [V/µs]	≥2.7
Power dissipation [mW]	≤0.6
CMRR and PSRR [dB]	≥60

TABLE II
DESIGN PARAMETERS OF THE FINAL ITERATION

$W_{1,2}$	75.04 μm, Nfingers = 16
$W_{3,4}$	38.4 μm, Nfingers = 8
$W_{5,6}$	360 μm, Nfingers = 32
W_7	76.8 μm, Nfingers= 16
$W_{10,11}$	30 μm, Nfingers = 8
$W_{12,14}$	3.9 μm, Nfingers = 1
W_{13}	6.25 μm, Nfingers = 1
W_{15}	25 μm, Nfingers = 4
W_{16}	12 μm, Nfingers = 10
C_C	1.586 pF, A =40x40 μm^2
R_C	8256 Ω

Unless otherwise stated, simulation results are done for room temperature and power supply voltage of 1.8V. Analyses were conducted in the Spectre simulator. For conducting the DC and AC analysis a dc-point trick [4] was needed to insure that transistors were correctly biased. This was done with a switch whose position depends on the type of analysis that is currently running. Fig. 2 illustrates how to connect the switch. The dc generator in the feedback branch provides a different input/output dc-point by changing its dc voltage. The results of input and output common mode voltage ranges are shown in Table III and represent the range in which performances do not fall bellow 95% of the set specifications.
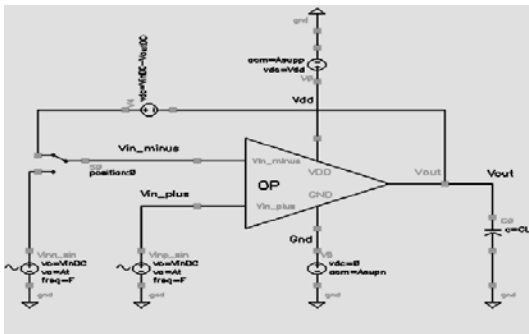
Fig. 2. Test bench circuit for the op amp.

TABLE III
COMMON MODE VOLTAGE RANGES

Parameters	From [V]	To [V]
Diff. gain > 71.25 dB (75dB – 5%)	0	1.32
Unity freq. > 47.5 MHz (50 MHZ -5%)	0	1.01
Phase margin > 57° (60° – 5%)	0	1.439
Input common mode range	0	1.01
--------------------------------------	---------	-------
Diff. gain > 71.25 dB (75dB – 5%)	0.32	1.6
Unity freq. > 47.5 MHz (50 MHZ -5%)	0.10	1.34
Phase margin > 57° (60° – 5%)	0.09	1.8
Output common mode range	0.32	1.34

It is interesting to illustrate how the compensating branch works. In Fig. 3 are shown results without the capacitor C_C and transistor Q_{16}. When C_C is inserted the first pole frequency is lowered and the phase margin is increased (Fig 4). After Q_{16} is introduced into the circuit, the unity frequency and phase margin is increased as shown on Fig. 5. Both behaviors are within conclusions given in Section II.
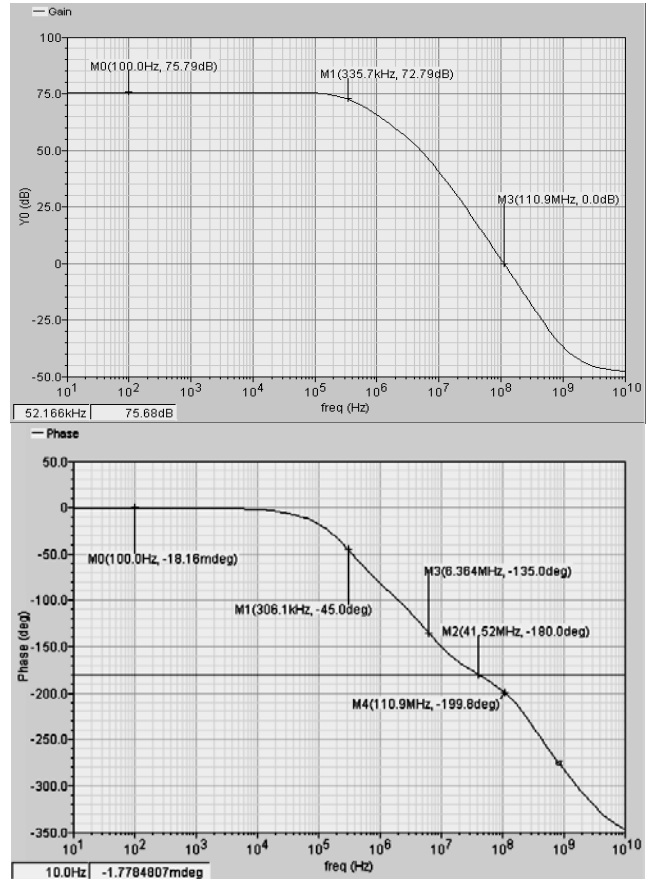
Fig. 3. Simulated transfer function of the op amp without compensation branch.

As op amp presented in this paper is intended for usage in mixed signal circuits, it needs to be able to suppress the interferences. For an example, the interferences that occur in the power supply voltage due to fast switching rate of the digital components. Parameters that describe how well the op amp is immune to these problems are CMRR and PSRR parameters. Simulation results for these parameters are shown indirectly with the gain of the common mode and the power supply signal gain in Fig. 6 and 7 respectively.

The rest of the simulation results are presented in Table IV. The results are achieved for an integrated resistor with temperature coefficient of -5500 ppm/°C. Manufactured circuits usually use an off-chip resistor which provides smaller temperature variations and better performance.
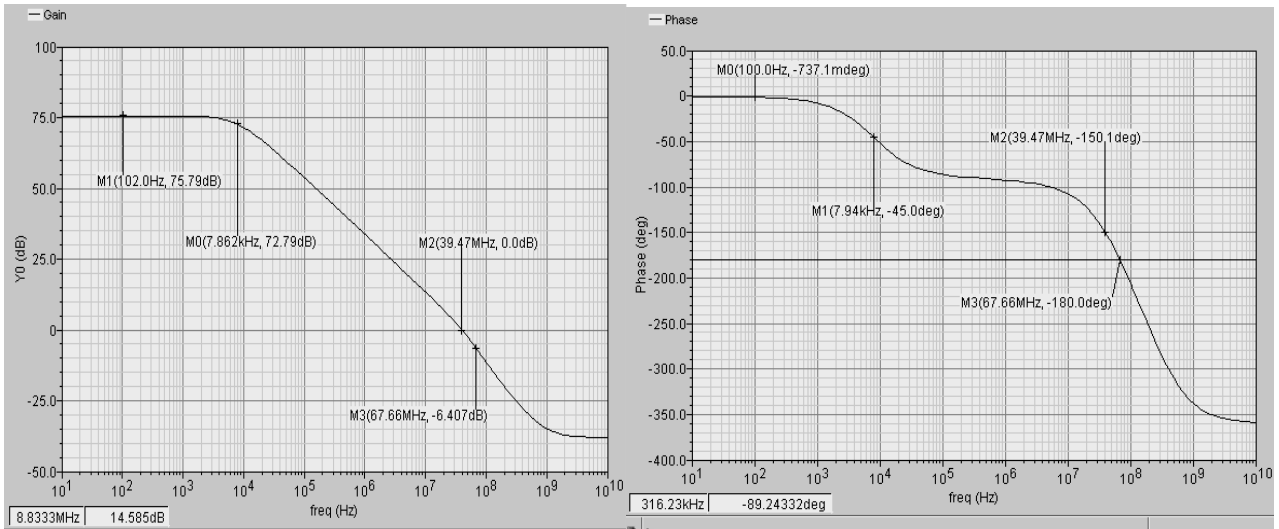
Fig. 4. Simulated transfer function of the op amp with just the Miler's capacitor C_C = 1.586 pF.
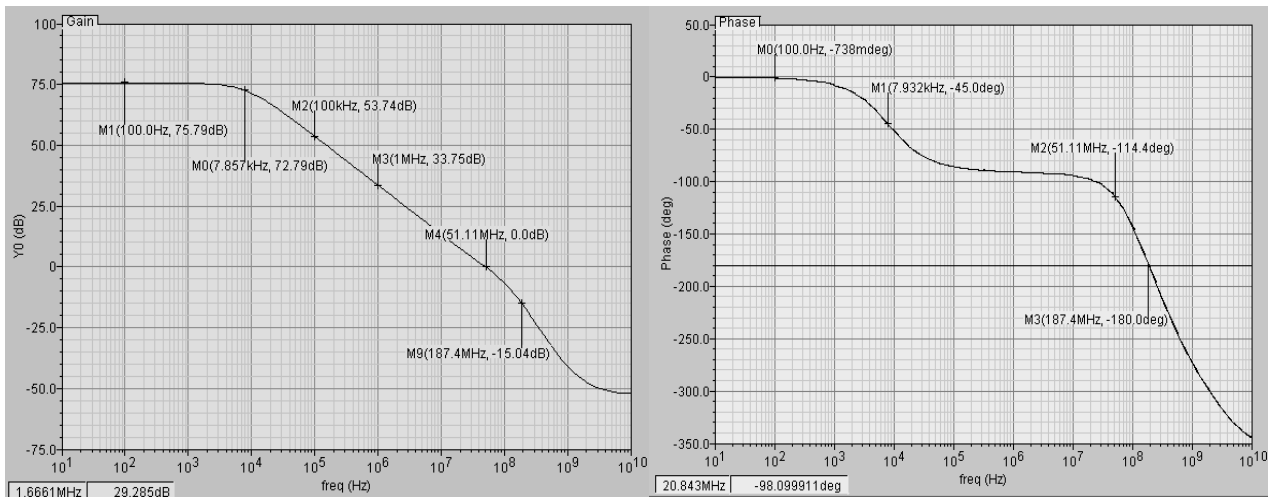
Fig. 5. Simulated transfer function of the op amp with the complete compensation branch.
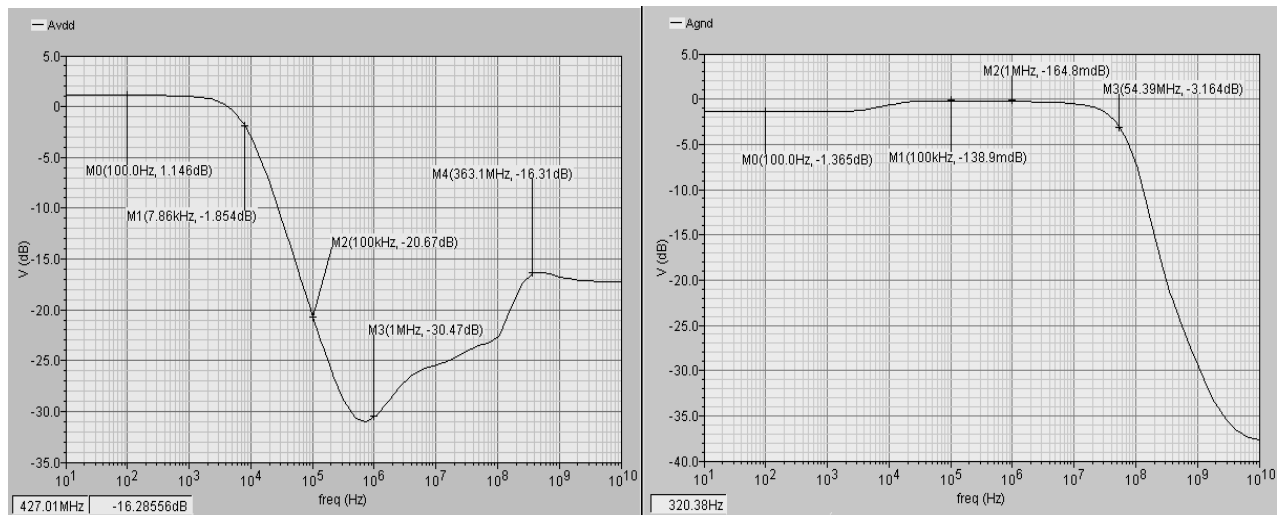
Fig. 6. Simulated results for the gain of a signal coming from VDD (left) and GND (right).

Fig. 7. Simulated results of the gain for a common mode signal at the input.

TABLE IV
SUMMARY OF ALL THE SIMULATION RESULTS

Parameter	Conditions	Min	Typ[1]	Max	Unit
Input CM	T=27°C,V_{DD}=1.8V	0	-	1	V
Output CM	T=27°C,V_{DD}=1.8V	0.32	-	1.34	V
Offset voltage	-40<T<125°C, 0.32<V_{INCM}<1V	-	0.06	0.11	mV
Phase margin	-40<T<125 °C, C_L=3pF	60	67	94[2]	°
Unity frequency	-40<T<125°C, C_L = 3pF,	42	52	117[2]	MHz
-3 dB bandwidth	T=27°C, C_L=3pF,	-	7.86	-	kHz
Slew rate	T=27°C,C_L=3 pF,	-	10	-	V/µs
Settling time	T=27°C,C_L=3 pF,	17[4]	-	150[5]	ns
Overshot	T=27°C,C_L=3 pF,	1.1[4]	-	13.7[5]	%
Linear range	T =27°C, C_L=3pF, V_{inCM}=550mV	-	-	8	MHz
	THD≤1%	2.2	-	43.45	mV
Diff. gain	-40<T<125°C	71	75	77	dB
CMRR	-40<T<125°C	63	73	94[3]	dB
PSRR	-40<T<125°C	63	74	78	dB
Supply current	-40< T<125°C	113	172	312	µA

(1) Typical values are for $V_{CM}=V_{DD}/2$, V_{DD} =1.8 V and T=27 °C.
(2) PM for the case when C_L = 0.
(3) CMRR for the case when V_{DD} + 10%.
(4) Linear settling time.
(5) Linear and nonlinear settling time.

IV. DISCUSSION

Several relationships between op amp parameters and its performance were noticed in the simulation results. First, the total gain and speed (unity frequency and slew rate) of the op amp show different behavior with the change of the supply current I_{DD}. With I_{DD} increase the output resistance decreases $(\sim 1/I_{DD})$ while the mosfet transconductance raises slower $(\sim \sqrt{I_{DD}})$ [3]. Increase in the overall gain, achieved with wider transistors, boosts parasitic capacitances. Consequently, the phase margin of the op amp is lowered. As its speed is directly proportional to the I_{DD}, change of supply current results in opposite gain and speed tendencies. It can be seen from the simulation

results that the input and output common mode ranges have different tendency compared to the CMRR and PSRR parameters. High rejection parameters values require that transistors operate in deep saturation region. Therefore, higher drain-source voltages are needed which results in smaller common voltage range.

There are also a few limitations to the polarization circuit. For start, all transistors in the bias circuit need to be in saturation. In Fig. 1 there are, in either branch, two diode connected mosfets in series and as such their drain-source voltages will be at least equal to the threshold voltage. This doesn't leave much room for the source-gate voltages of pmos transistors and as a result current that they generate will be relatively small. Although negligible dissipation in the bias circuitry is one of the goals, extremely wide transistors for current generators (Q_5 and Q_6) are not an option. Additionally, the body effect of the nmos transistors consumes the already little voltage that you have at your disposal. Moreover, the drain potentials for Q_{10} and Q_{11} are not the same, which due to channel length modulation, results in different currents I_{D10} and I_{D11}.

The maximum capacitive load is 3 pF, which is too small if op amp is used as a discrete component. But if this op amp is used as a small part of a complex chip then this value is sufficient for driving the capacitive load of the next stage. Higher load at the output lowers the phase margin and makes the circuit more unstable because the second pole shifts to lower frequency Eq. (10) [2].

V. CONCLUSION

Design guidelines and simulation results for a two-stage CMOS operational amplifier were presented in this paper. Due to a delicate balance between parameters and the fact that op amp was designed in 150nm technology some compromise had to be made during the design. Targeted performance was accomplished for the specifications set and op amp can be used for low-voltage applications, consuming around 0.3 mW (V_{DD}=1.8V). Better temperature independency and wider operating temperature range (-40°C to 125°C) can be attained with a single off-chip resistor.

REFERENCES

[1] Franco Maloberti, *Analog Design for CMOS VLSI Systems,* Kluwer Academic Press, 2001.
[2] David Johns, Ken Martin, *Analog integrated circuit design*, John Wiley & Sons, New York, 1997.
[3] Behzad Razavi, *Design of Analog CMOS Integrated circuits*, McGraw – Hill Company, New York, 2001.
[4] Ali Fazli, Dai Zhang, Daniel Svärd, *TSEK37 Analog CMOS Integrated Circuits – Analog Lab*, LINKÖPING University 2010.

Resistive Feedback Influence on Ring Oscillator Performance for IR-UWB Pulse Generator in 0.13μm CMOS technology

Jelena Radic, Alena Djugova, Laslo Nadj and Mirjana Videnovic-Misic

Abstract - A CMOS standard three-stage ring oscillator is examined in UMC 0.13μm technologies. As the ring oscillator is a part of IR-UWB (Impulse Radio Ultra Wide Band) pulse generator, its oscillating frequency determines the central frequency of the pulse spectrum and has significant effect on spectrum fitting within UWB FCC mask. The influence of inverter feedback resistors on the ring oscillator frequency and the peak-to-peak amplitude are investigated. Furthermore, as the ring oscillator usually drives a buffer in pulse generator/transmitter chain, dependence of its Figures of Merit on the buffer resistive feedback is presented in the paper.

Keywords – CMOS technology, resistive feedback, impulse radio (IR), pulse generator, ring oscillator, ultra-wideband (UWB).

I. INTRODUCTION

Impulse Radio Ultra-Wide-Band (IR-UWB) technology has emerged as a potential solution for very high data rate short-range communication, and low data rate communication related to localization, targeting both low cost and low power consumption [1-3]. It transmits extremely short pulses, on the order of a nanosecond or less, which occupy a bandwidth up to several GHz. Additionally, IR-UWB technology offers high fading margin for communication systems in multipath environments [3].

The American Federal Communications Commission (FCC) defines a signal as ultra-wideband if it occupies more than 500 MHz of radio frequency spectrum or exhibits a fractional bandwidth of at least 25% [4]. Since the FCC allocated frequency spectrum for UWB technology is 3.1−10.6 GHz, the power level from the UWB transmitter should be small enough not to interfere with the already existing communication systems such as WiMax, Bluetooth and GSM. This requirement limits output power level of UWB TXs at −41.3 dBm/MHz [4]. In the GPS band (0.96–1.61 GHz), there is even more stringent regulation: less than −75.3 dBm/MHz is needed to

Jelena Radic, Alena Djugova, Laslo Nadj and Mirjana Videnovic-Misic are with the Department of Power, Electronics, and Communications, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovica 7, 21000 Novi Sad, Serbia, E-mail: {jelenar_, alenad, lnadj, mirjam}@uns.ac.rs.

avoid interference problems. The PSD (Power Spectral Density) in frequency interval from 1.61 GHz to 3.1 GHz depends on the type of application (indoor, outdoor, GPS, wall & medical imaging, through-wall imaging & surveillance system). In spite of these regulations, there have been many reports of interferences with wireless local area network (WLAN) systems operated in the 5–6 GHz band. Therefore, due to practical reasons, the UWB bandwidth is subdivided into two bands: 3−5 GHz (low-band) and 6−10.6 GHz (high-band).

One of the most critical components of an UWB system is the pulse generator, which has to be low power and low complexity. Additionally, spectrum of generated pulse train has to satisfy the FCC spectral mask, making pulse generator design very challenging. There are several typical techniques for designing it which usually follow all-digital [5], [6], analogue/digital [7] or all-analogue [7], [8] design approach. Digital solutions offer higher integration, lower consumption and better controllability while all-analogue techniques demonstrate circuit simplicity.

As an essential part of analogue/digital pulse generator [5], a ring oscillator is studied in this work. Dependence of its performance on inverter and buffer resistive feedbacks is examined in 0.13μm UMC CMOS technology.

II. STANDARD THREE-STAGE RING OSCILLATOR DESIGN

The pulse generator represents a key block in impulse UWB communications. As pulse shape determines the spectrum characteristic of the UWB signal and effectively dictates specific system requirements, its generation is one of the essential considerations in the UWB design. Fig. 1
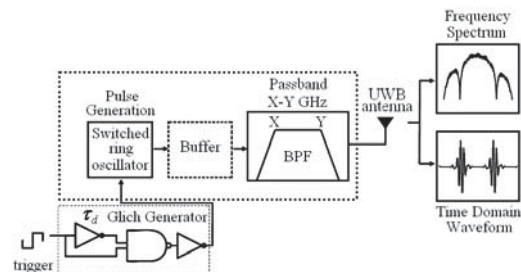
Fig. 1. A IR-UWB transmitter based on ring oscillator as a part of pulse generator.
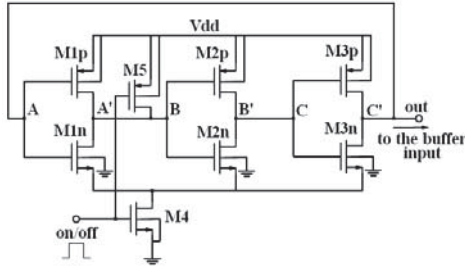
Fig. 2. The three-stage ring oscillator architecture.

shows the basic topology of an IR-UWB transmitter based on ring-oscillator as a part of the pulse generator. It consists of a glitch generator, a switched ring oscillator, a buffer stage and a pulse shaping (band-pass) filter [7]. The glitch generator turns the ring oscillator on/off. It defines ring oscillation time and the width of the pulse. Since the time domain pulse width decides its frequency spectrum [8], it is important to design a pulse, which makes optimal usage of the available spectrum, within the limits imposed by FCC. The switched ring oscillator should generate signal in glitch-defined interval. Its oscillating frequency defines position of the transmitted pulse spectrum within the FCC mask [8]. The buffer isolates the ring oscillator from the pulse shaping filter loading and improves the pulse generator current driving capability. The band-pass filter additionally accommodates the pulse in the allowed spectral mask of FCC.

The switched ring oscillator topology is shown in Fig. 2. It is composed of the three-stage ring oscillator (M_1–M_3) and a pair of oscillation-enabling switches (M_4 and M_5). Due to its simplicity and short start-up time, the ring oscillator is the mostly used architecture in IR-UWB transmitter applications. It has small resistance at each feedback nod which allows fast transient response.

The oscillation-enabling switches, as their name says, control the oscillation process. When the on-off signal (produced by glitch generator) is high, M_4 is turned on (M_5 is turned off), the inverters M_1–M_3 outputs have voltage values determined by the size ratio of the corresponding pMOS and nMOS transistors. Due to the small inverter reactance, the oscillation can start immediately. M_5 turns on (M_4 turns off) at the on-off signal low edge, connecting the M_1 transistor output and the M_2 transistor input to V_{dd}, and effectively shutting down the oscillation.

III. RING OSCILLATOR PERFORMANCE

The design proposed have been simulated in mixed mode/RF UMC 0.13μm CMOS eight-metal technology using SpectreRF Simulator from Cadence Design System. Supply voltage V_{dd} of this technology is 1.2 V.

The ring working frequency depends directly on transistors sizes. If the transistors are larger, the period of the oscillation T rises proportionally, while the oscillating frequency decreases ($f_0=1/T$), and vice versa. For the

NMOS and PMOS ring transistors size, channel width/length: W/L=0.36μm/0.12μm (and identical sizes of all NMOS and PMOS transistors gates), the oscillation frequency of 7.65 GHz has been obtained. It should be noted that this is the smallest NMOS transistor size. The PMOS transistor width (W) could be decreased to 0.32μm, but in inverter topology PMOS transistor is usually equal or two times larger than NMOS transistor. The latter ratio of transistors sizes has been used in buffer stages. To utilize the UWB high-band more effectively, the center frequency of at least 8GHz is required. Higher ring oscillator frequency could be achieved without PMOS transistor M_5. However, this transistor provides start of the oscillation from the same initial state by connecting A' (B) node to V_{dd}. This happens at falling edge the on-off signal, when the oscillation-enabling switch (transistor M_4) turns off effectively shutting down the oscillations.

A. Influence of the inverter resistive feedback

To increase the oscillation frequency at a given DC current, the resistive feedback is used in each inverter stage as proposed in Ref [7]. Resistors are connected between nodes A−A', B−B', and C−C', shown in Fig. 3. Dependence of the ring oscillator performance on the resistor R value is shown in Fig. 4. It can be noticed that with resistor value decrease, oscillation frequency rises,
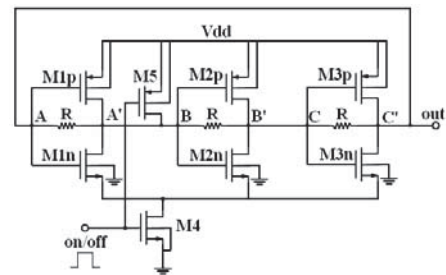
Fig. 3. The three-stage ring oscillator architecture.
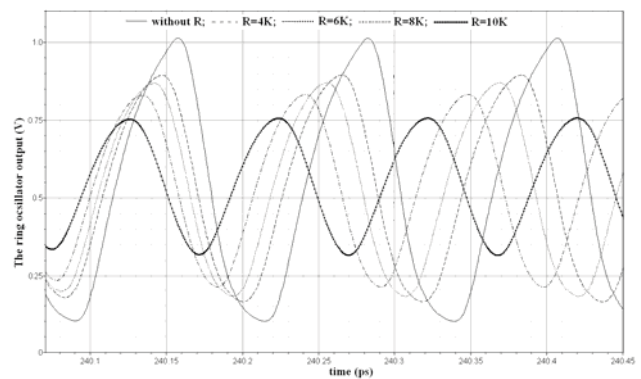
Fig. 4. Dependence of the ring oscillator performance on the inverter feedback resistor value.
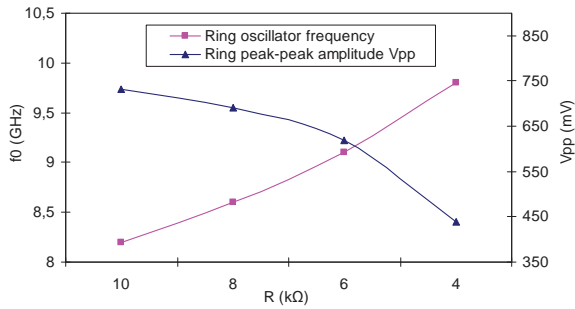
Fig. 5. Dependence of the ring oscillator frequency and the peak-to-peak amplitude V_{pp} on the feedback resistor value.

TABLE I
RING OSCILLATOR PERFORMANCE DEPENDENCE ON THE INVERTER RESISTIVE FEEDBACK

R (kΩ)	f_0 (GHz)	Vpp (mV)
10	8.2	733
8	8.6	690
6	9.1	618
4	9.8	440

while the peak-to-peak amplitude V_{pp} of the ring output signal reduces as expected, Fig. 5. This is caused by the reduction of the inverter gain with decreasing the feedback resistor value. Furthermore, the V_{pp} parameter and the time for reaching the peak signal values reduce resulting in the ring oscillator frequency increase. Simulated V_{pp} and f_0 values of the ring oscillator in 0.18μm technology are presented in Tables I. For R changes from 10 kΩ to 4 kΩ, the oscillating frequency is in the range from 8.1 GHz to 9.05 GHz, while V_{pp} decreases from 733 mV to 440 mV. It can be noticed that the oscillation frequency was changed in the considerably wide frequency range of 1.6 GHz, and the lowest value of 8.2 GHz (obtained for the highest resistor value) is by far larger than 7.65 GHz achieved with standard topology shown in Fig.2. However, the oscillator output amplitude for the lowest resistor value is not large enough to drive properly the subsequent stage (usually a buffer) in the pulse generator chain. Therefore the resistor value should be selected considering the trade-off between the oscillation frequency and the required value for V_{pp}.

B. Influence of the buffer resistive feedback

The method of increasing the ring oscillator frequency by using the inverter resistive feedback is already known in literature, Ref [7]. To the best of our knowledge, the influence of the buffer feedback resistor on the ring performance is still not examined. Fig. 6a shows schematic of the buffer with resistive feedback. To provide better isolation between the ring oscillator and the pulse shaping filter, a two stage buffer is proposed, Fig. 6b. The first stage role is to increase the ring oscillator frequency, while
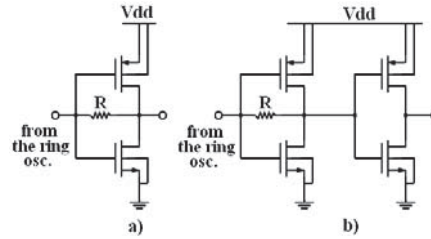
Fig. 6. Buffer topologies: a) Buffer stage with the resistive feedback b) Two-stage buffer.
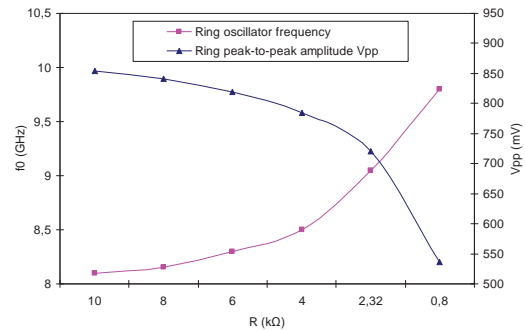
Fig. 7. Dependence of the ring oscillator performance on the buffer feedback resistor value.

TABLE II
RING OSCILLATOR PERFORMANCE DEPENDENCE ON THE BUFFER RESISTIVE FEEDBACK

R (kΩ)	f (GHz)	Vpp (mV)
10	8.1	854
8	8.15	841
6	8.3	820
4	8.5	785
2.32[*]	9.05	720
0.8[**]	9.8	536

*Minimal resistance value for the RNHR resistor model.
*Different kind of the resistor model, RNPP0.

the second buffer stage prevents this effect to change the output bandpass filter shaping. Dependence of the ring oscillator performance on the (first) buffer resistive feedback is shown in Fig. 7. Obtained simulation results are summarized in Tables II. Varying R from 10 kΩ to 800 Ω, the oscillation frequency is increased from 8.1 GHz to 9.8 GHz, followed with decrease in V_{pp} from 854 mV to 536 mV. As the smallest value of the RNHR high-sheet resistor model is 2.32 kΩ, the RNPPO UMC resistor model had to be used for 800 Ω. This value was chosen as it gives the highest frequency value obtained by previous technique. It can be noticed that the available frequency range are increased to 1.7 GHz at the same time with higher V_{pp} parameter values and larger interval of resistance change. Influence of the buffer resistive

feedback on the ring oscillator can be explained by load changing. The three-stage ring oscillator period T is determined by the propagation time of a signal transition through the complete oscillator chain and is defined as:

$$T = 6 \cdot t_P, \qquad (1)$$

where t_P is propagation delay of the gate. An inverter signal propagation time is largely determined by the strength of the driving gate, and the load presented by the output node itself, which sums the contributions of the connecting gates and the wiring parasitic. Change in the buffer feedback resistor modifies the buffer input impedance (based on Miller's theorem) changing simultaneously the ring oscillator load. This leads to propagation time, and consequently the ring oscillator period/frequency change.

III. Discussion

From the results presented in previous section can be seen that the ring oscillator frequency is strongly dependent on the resistive feedback. The change of the oscillating frequency is 20.9 % for the inverter feedback resistor in the range from 10 kΩ to 4 kΩ, and 22.2 % for the buffer feedback and R in the range from 10 kΩ to 0.8 kΩ. Additionally, the maximum frequency (9.8 GHz) obtained by using feedback is remarkably (28.1 %) higher than value (7.65 GHz) achieved without any feedback. Furthermore, simulation results indicate that the peak-to-peak amplitude values of the second method are notably higher comparing to the V_{pp} values of the first technique obtained for the same oscillating frequency. This can be attributed to the fact that the resistive feedback of the inverter changes directly its gain (and thus the total ring oscillator gain) and currents available to charge/discharge the load capacitance resulting in significant change of voltage peak values. Taking into account the main aim of IR-UWB transmitter design, which includes satisfying the FCC mask requirements with as higher as possible the peak-to-peak amplitude, it could be concluded that the second method gives better ring oscillator performance. Moreover, the maximum frequency of 9.8 GHz and higher V_{pp} parameter value has been obtained by using only one resistor of 800 Ω compared with three additional resistors of 4 kΩ used in the ring resistive feedback. In the field of IC design, it is well known that passive components require considerable die area, increasing the cost and causing problem for area constrained applications. Therefore, from the fabrication cost point of view, the design with buffer feedback will be cheaper.

IV. Conclusion

Standard three-stage ring oscillator topology has been analyzed in 0.13μm UMC CMOS technology. Dependence of its performance on the resistive feedback of the ring inverters and the buffer stage has been investigated. Simulations confirmed strong dependency of the ring oscillator frequency and the peak-to-peak amplitude (especially in the case of the ring inverters feedback) on the feedback resistor value. Additionally, the maximum oscillating frequency obtained in both topologies, is significantly higher (28.1 %) compared to the standard ring working frequency. Likewise, simulation results showed better Figures of Merits of the ring oscillator as a part of UWB-IR pulse generator, in case the buffer resistive feedback has been used. In the latter architecture, a two-stage buffer should be used to provide better isolation between the ring oscillator and the pulse shaping filter.

Acknowledgement

References

[1] M. Ghavami, L. B. Michael, and R. Kohno, "Ultra Wideband Signals and Systems in Communications Engineering," John Wiley&Sons Ltd, 2004.

[2] K. Siwiak and D. McKeown, "Ultra-Wideband Radio Technology," John Wiley&Sons Ltd, 2004.

[3] J. R. Fernandes and D. Wentzloff, "Recent Andvances in IR-UWB Transceivers: An Overview ," *IEEE Int. Conf. on Circuits and Systems,* pp. 3284−3287, 2010.

[4] First Report and Order: Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems Federal Communications Commission (FCC), ET Docket 98-153, Adopted February 14, 2002, Released Apr. 22, 2002.

[5] V. V. Kulkarni, M. Muqsith, K. Niitsu, H. Ishikuro, T. Kuroda, "A 750 Mb/s, 12 pJ/b, 6-to-10 GHz CMOS IR-UWB transmitter with embedded on-chip antenna", *IEEE Jour. of Solid. State Circuits*, vol. 44, no. 2, pp. 394-403, Feb. 2009.

[6] H. Kim, Y. Joo, S. Jung, "A tunable CMOS UWB pulse generator", *The 2006 IEEE International Conference on Ultra-Wideband*, Waltham, MA, pp. 109-112, 24-27 Sept. 2006.

[7] S. Sim, D.W. Kim, S. Hong "A CMOS UWB pulse generator for 6–10 GHz applications", *IEEE Microwave and wireless components letters*, vol. 19, no. 2, pp. 83-85, Feb. 2009.

[8] O. Novak, C. Charles, "Low-power UWB pulse generators for biomedical implants", *IEEE International Conference on Ultra-Wideband*, Vancouver, BC, pp. 778-782, 9-11 Sept. 2009.

GNSS Signal Simulation and a Multipath Delay Estimation

Marko S. Djogatović, Milorad J. Stanojević

Abstract - In the satellite navigation systems, distortion of a received signal correlation function, due to the multipath propagation, can gravely degrade position estimation. The positioning accuracy is strongly affected by the quality of the received signal time-delay estimations. In the paper, signal and channel models for the L1 channel GPS C/A signal and the Galileo BOC(1,1) signal will be presented and the multipath mitigation problem analyzed. In addition, the MEDLL algorithm and the particle filter will be presented in detail and mutually compared for different simulated signals and different correlation times.

Keywords - Multipath, Signals, Simulation, Estimation, GNSS.

I. INTRODUCTION

In modern Global Satellite Navigation Systems (GNSS), the most significant source of error that affects the navigational signals during their propagation is multipath [1]. Multipath degrades the correlation function in such a way that it is not possible to determine accurately the signal delay. Therefore, removal of multipath is very important for applications where high precision measurements are required (geodesy and surveying, instrument landing systems, atmospheric sensing) or for indoor positioning where line-of-sight (LOS) signal is highly deteriorated by multipath. In some applications, like remote sensing, removing influence of multipath replicas is not enough. Hence, it is necessary to determine amplitude and time-delay of these replicas.

So far, a various multipath mitigation methods have been developed. Fig. 1 represents hierarchy of commonly used multipath mitigation techniques. Many of these methods are using correlation of early and late signal replicas with a received signal in order to find value of the time-delay that corresponds to the maximum power of correlation. Narrow early-minus-late (EML) correlation technique is derived from standard EML by narrowing 0.5 chip space between early and late correlators to 0.1-0.2 chip space. [2,3]. Other important correlator-based multipath mitigation techniques are: Double Delta (ΔΔ), also known as High Resolution Correlator (HRC) [4], then Strobe and Enhanced Strobe Correlator (ESC) [5], E1/E2 Tracker [6], Multipath Elimination Technique (MET) known as Early Late Slope (ELS) with Pulse Aperture Correlator (PAC) as a simple hardware implementation [7,8] and efficient technique based on Teager-Kaiser

Marko S. Djogatovic and Milorad J. Stanojevic are with the Faculty of Traffic and Transport Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, E-mail: {m.djogatovic, milorad}@sf.bg.ac.rs.

operator (TK) [9].

Maximum Likelihood (ML) estimation is based on the maximum likelihood principle and is very popular approach in signal processing. Once a signal model is specified with its parameters, and data have been collected, the maximum likelihood estimator is used to find the value of parameter that maximizes likelihood function. So far, several maximum likelihood estimation techniques have been used for multipath mitigation: Newton methods with analytical [10,11] and numerical (FIMLA, RML) [12,13] expressions for gradient and Hessian term, then, the most famous ML method, Maximum Estimating Delay Lock Loop (MEDLL) [14,15] with its modifications (ML2 and Reduced Search Space ML - RSSML) [16,17], Multipath Mitigation Technique (MMT) integrated into Novatel's Vision Correlator [18] and iterative methods based on expectation-maximization algorithms (Space Alternating Generalized Expectation-Maximization – SAGE) [19].
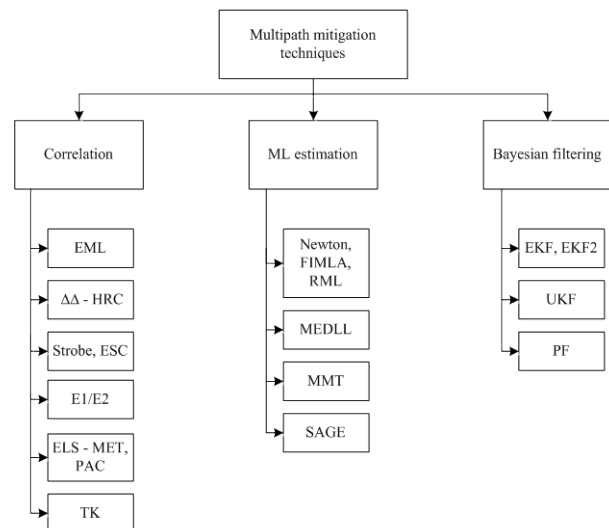
Fig. 1. Hierarchy of multipath mitigation techniques

Bayesian filtering is concerned with the estimation of the underlying probability distribution of a random signal in order to extract the original signal from noisy measurements. In order to mitigate influence of the reflected signal components on LOS signal delay estimation following Bayesian filters have been used: Extended Kalman Filter (EKF) [20,21] and Second-order Extended Kalman Filter (EKF2) [21], Unscented Kalman Filter (UKF) [22] and Particle Filter (PF) [23,24].

In the paper, an emitted and a received signal models for the GPS L1 C/A (coarse/acquisition) signal and the Galileo BOC(1,1) (Binary Offset Carrier) signal will be presented and the multipath mitigation problem

investigated. Moreover, two estimation algorithms will be presented in detail and analyzed: the well-established and efficient MEDLL algorithm in contrast to the newly developed particle filter method.

II. SIGNAL AND CHANNEL MODEL

A. Transmitted signal

The signal $s(t)$ transmitted from one satellite can be written as [3,17]

$$s(t) = \sqrt{E_b} \cdot q(t) \cdot \cos 2\pi f_1 t , \qquad (1)$$

where E_b is the bit energy, $q(t)$ is the navigation data after spreading and f_1 is the L1 carrier wave frequency. Spreading of navigation data bits, $\{d(n)\}$, is done as

$$q(t) = \sum_{n=-\infty}^{\infty} d(n) p(t - nT_b), \qquad (2)$$

where $p(t)$ is the spreading waveform and T_b is the period of one data bit of the navigation message. The spreading waveform can be written as follows [17]

$$p(t) = g(t) \star \sum_{k=0}^{N_c-1} c(k)\delta(t - kT_c). \qquad (3)$$

Above, $g(t)$ is the modulation waveform (GPS L1 C/A signal or composite BOC(−) for Galileo E1-C signal), $\delta(\cdot)$ is the Dirac delta function, $\{c(k)\}$ is the spreading, pseudorandom (PRN) sequence of length N_c and the star sign ($\star$) denotes convolution. T_c is duration of one chip in the code sequence.

The modulation waveform, $g(t)$, can be written as

$$g(t) = g_P(t) \star \sum_{i=0}^{N_{sw}-1} \delta(t - iT_{sw}), \qquad (4)$$

where N_{sw} is modulation order (the number of periods of the square wave within one chip), $T_{sw} = T_c / N_{sw}$ is the period of square wave, and $g_P(t)$ is the shaping pulse. For GPS C/A signal is true that $N_{sw} = 1$ ($g(t) = g_B(t)$) and for BOC modulation that $N_{sw} = 2f_{sc} / f_c$, where $2f_{sc}$ is the is the square wave frequency and f_c is the chip frequency [17].

The shaping pulse $g_P(t)$ can be defined as filtered rectangular pulse using following equation

$$g_P(t) = \frac{1}{\pi T_{sw}} \Big[Si(2\pi bt / T_{sw}) - Si(2\pi b(t / T_{sw} - 1)) \Big], \quad (5)$$

where b describes the location of the cut-off frequency and it is related to the bandwidth, B_w, through the relation

$b = B_w T_{sw} / (2\pi)$. $Si(\cdot)$ is the sine integral. On Fig. 1 a) and b) are shown the GPS C/A pulse and the Galileo BOC(1,1) pulse, respectively, for infinite bandwidth and in the band-limited case ($B_w = 6$ MHz).
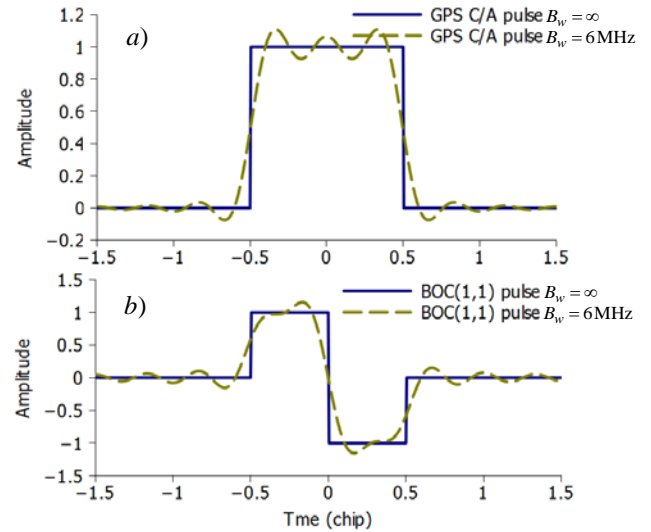
Fig. 1. a) GPS C/A and b) BOC(1,1) pulse in infinite-bandwidth and band-limited case

An expression for the power spectrum of GPS C/A periodic PRN can be written as

$$S_{\text{GPSC/A}}(f) = \frac{1}{N_c^2} \Big(\delta(2\pi f) +$$
$$\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} (N_c + 1)\text{sinc}^2\left(\frac{m\pi}{N_c}\right)\delta\left(2\pi f + m\frac{2\pi f_c}{N_c}\right)\Big), \quad (6)$$

for $N_c = 1023$ [1,3]. The power spectrum density of the BOC($f_{sc}/f_{ref}, f_c/f_{ref}$) centered at the origin can be written as

$$S_{\text{BOC(-)}}(f) = \begin{cases} f_c \left(\dfrac{\tan\left(\dfrac{\pi f}{2f_{sc}}\right)\sin\left(\dfrac{\pi f}{f_c}\right)}{\pi f}\right)^2 , & \text{if } \dfrac{2f_{sc}}{f_c} \text{ is even} \\[4ex] f_c \left(\dfrac{\tan\left(\dfrac{\pi f}{2f_{sc}}\right)\cos\left(\dfrac{\pi f}{f_c}\right)}{\pi f}\right)^2 , & \text{if } \dfrac{2f_{sc}}{f_c} \text{ is odd} \end{cases} \quad (7)$$

where $f_{ref} = 1.023$ MHz [3].

Fig. 2 shows power spectrum density for GPS C/A and Galileo BOC(1,1) spreading signal. From Fig. 2 it can be seen that the BOC(1,1) signal spectrum is symetric split spectrum with two main lobes shifted from the carrier frequency by the amount equal to the subcarrier frequency.
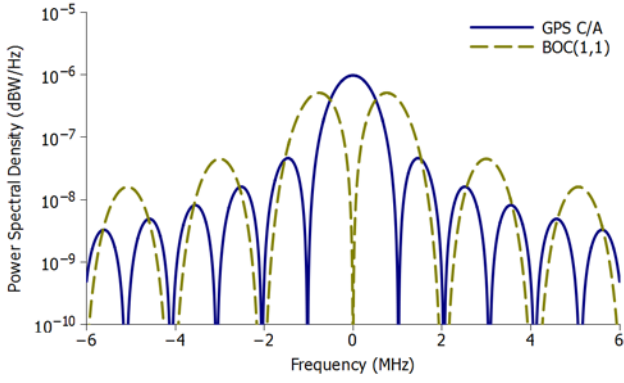
Fig. 2. GPS C/A and Galileo BOC(1,1) spectrum

B. Received signal

The received signal $r(t)$ from one satellite, in multipath environment, is composed of M paths, where one is the LOS signal and the others are reflected rays of the LOS signal. All additional sources of interference are set into a single additive Gaussian noise term, $v(t)$. After carrier removal and filtering, received signal $r(t)$ can be written as [17,23,25]

$$r(t) = a_0 q(t - \tau_0) e^{j\phi_0} + \sum_{m=1}^{M-1} a_m q(t - \tau_m) e^{j\phi_m} + v(t), \quad (8)$$

where a_m is the amplitude of the m-th path, ϕ_m is the phase of the m-th path and τ_m is the channel delay introduced by the m-th path.

C. Problem formulation

Here, we assume that the parameters of the received signal (a_m, τ_m, ϕ_m) are slowly varying, almost constant, during selected observation period. Let us define $\mathbf{a}(t) \in C^{M \times 1}$ and $\boldsymbol{\tau}(t) \in R^{M \times 1}$ as vectors containing complex amplitudes and time delays of the LOS signal and the multipath signals, respectively. The complex vector $\mathbf{a}(t) = \left[a_1(t) e^{j\phi_1(t)} \quad \dots \quad a_M(t) e^{j\phi_M(t)} \right]^T$ is defined. A vector of the delayed signal components is defined as, $\mathbf{q}(t, \boldsymbol{\tau}) = \left[q(t - \tau_1) \quad \dots \quad q(t - \tau_M) \right]$, where $\mathbf{q}(t, \boldsymbol{\tau}) \in C^{1 \times M}$. So, the multipath signal model that is given in equation (8) can be expressed in the vector form as [23,25]

$$r(t) = \mathbf{q}(t, \boldsymbol{\tau}) \mathbf{a}(t) + v(t). \quad (9)$$

Suppose that L samples of the signal are taken with a sampling interval T_s satisfying the Nyquists criterion. Then the sampled data in the k-th correlation period (period of waveform sampling and stacking) can be expressed as

$$\mathbf{z}(k) = \mathbf{Q}(k, \boldsymbol{\tau}) \mathbf{a}(k) + \mathbf{v}(k), \quad (10)$$

where matrix $\mathbf{Q}(k, \boldsymbol{\tau}) \in C^{L \times M}$ is matrix containing L samples of delayed narrowband envelopes of LOS and multipath signals. The received signal and the white noise are expressed as $\mathbf{z}(k)$, $\mathbf{v}(k) \in C^{L \times 1}$, respectively [23,25].

D. ML estimation

According to maximum likelihood estimation theory, when the noise is white, the best estimates of parameters are those values that maximize following likelihood function [25]

$$p(\mathbf{z} | \boldsymbol{\tau}, \mathbf{a}) = (2\pi)^{-\frac{L}{2}} |\mathbf{S}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \left(\mathbf{z}(k) - \mathbf{Q}(k, \boldsymbol{\tau}) \mathbf{a}(k)\right)^H \right.$$
$$\left. \mathbf{S}^{-1} \left(\mathbf{z}(k) - \mathbf{Q}(k, \boldsymbol{\tau}) \mathbf{a}(k)\right)\right), \quad (11)$$

where $\mathbf{S}$ is noise covariance. The minimum of log-likelihood function $\ell(\boldsymbol{\tau}, \mathbf{a}) = \ln p(\mathbf{z} | \boldsymbol{\tau}, \mathbf{a})$ can be found by setting the derivatives $\partial\ell \backslash \partial\mathbf{a}$, $\partial\ell \backslash \partial\boldsymbol{\tau}$ to zero. It is easy to prove that, for the fixed $\boldsymbol{\tau}$, global minimum is attained at

$$\hat{\mathbf{a}}(k) = \left(\mathbf{z}(k)^H \mathbf{Q}(k, \boldsymbol{\tau}) \left(\mathbf{Q}(k, \boldsymbol{\tau})^H \mathbf{Q}(k, \boldsymbol{\tau}) \right)^{-1} \right)^T. \quad (12)$$

Hence, the log-likelihood function, with $\boldsymbol{\tau}$ as parameter, can be written as

$$\ell(k, \boldsymbol{\tau}) = r_{\tilde{z}z}(k) - \mathbf{R}_{z\tilde{Q}}(k, \boldsymbol{\tau}) \mathbf{R}_{\tilde{Q}Q}^{-1}(k, \boldsymbol{\tau}) \mathbf{R}_{z\tilde{Q}}^H(k, \boldsymbol{\tau}), \quad (13)$$

where cross-correlation and auto-correlation matrices are defined as

$$\hat{r}_{\tilde{z}z}(k) = \frac{1}{L} \mathbf{z}(k)^H \mathbf{z}(k); \quad \mathbf{R}_{\tilde{z}Q}(k, \boldsymbol{\tau}) = \frac{1}{L} \mathbf{z}^H(k) \mathbf{Q}(k, \boldsymbol{\tau})$$
$$\mathbf{R}_{\tilde{Q}z}(k, \boldsymbol{\tau}) = \mathbf{R}_{\tilde{z}Q}^H(k, \boldsymbol{\tau}); \quad \mathbf{R}_{\tilde{Q}Q}(k, \boldsymbol{\tau}) = \frac{1}{L} \mathbf{Q}^H(k, \boldsymbol{\tau}) \mathbf{Q}(k, \boldsymbol{\tau}) \quad (14)$$

The superscript H refers to Hermitian transpose or conjugate transpose of complex matrices.

ML estimates of the time-delay and amplitude vector from equation (10) are obtained using following equations [23,25]

$$\hat{\boldsymbol{\tau}}(k) = \min_{\boldsymbol{\tau}(k)} \left\{ \ell(k, \boldsymbol{\tau}) \right\}$$
$$\hat{\mathbf{a}}(k) = \left(\mathbf{R}_{\tilde{z}Q}(k, \boldsymbol{\tau}) \mathbf{R}_{\tilde{Q}Q}^{-1}(k, \boldsymbol{\tau}) \right)^T \Big|_{\boldsymbol{\tau} = \hat{\boldsymbol{\tau}}(k)}. \quad (15)$$

Fig. 3 shows auto-correlation functions for GPS C/A and BOC(1,1) signals in the band-limited case with a selected bandwidth of 6 MHz.
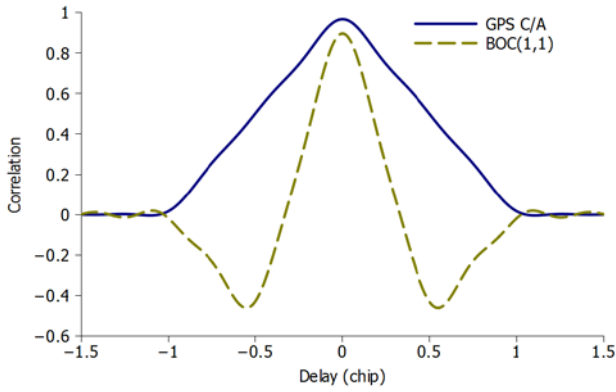
Fig. 3. GPS C/A and BOC(1,1) signal auto-correlation in the band-limited case

III. MULTIPATH ESTIMATING DELAY LOCK LOOP (MEDLL)

When estimating the parameters of log-likelihood function, $\ell(\tau, \mathbf{a})$, in the situation when M-1 multipath components are present in signal, following system of equations can be used [14,15]

$$\left[\hat{\boldsymbol{\tau}}(k)\right]_m = \max_{\tau(k)} \Re \left\{ \left(\mathbf{R}_{\tilde{Q}z}(k,\tau) - \sum_{\substack{i=0 \\ i \neq m}}^{M} \left[\hat{\mathbf{a}}(k)\right]_i \left[\mathbf{R}_{\tilde{Q}Q}\left(k, \tau - [\hat{\boldsymbol{\tau}}(k)]_i\right) \right]_{i,i} \right) e^{-j\arg[\hat{\mathbf{a}}(k)]_m} \right\},$$

$$\left[\hat{\mathbf{a}}(k)\right]_m = \mathbf{R}_{\tilde{Q}z}(k,\tau) - \sum_{\substack{i=0 \\ i \neq m}}^{M} \left[\hat{\mathbf{a}}(k)\right]_i \left[\mathbf{R}_{\tilde{Q}Q}\left(k, \tau - [\hat{\boldsymbol{\tau}}(k)]_i\right) \right]_{i,i}. \quad (16)$$

$$m = 0, \dots, M-1$$

First equation in (16) says that for one signal component, estimated $\hat{\boldsymbol{\tau}}(k)$ is found in the maximum of the cross-correlation function when influence of the other signal components is removed. In the same manner, for the fixed $\hat{\boldsymbol{\tau}}(k)$, complex amplitude is found when influence of the other signal components is removed. When time-delay for one signal component is computed, the time-delays of the other components are not known in advance, so the following iterative algorithm will be used:

MEDLL ALGORITHM
(in the case of LOS signal and one multipath component)

---

1. The correlation function $\mathbf{R}_0(\tau)$ is set to $\mathbf{R}_{\tilde{Q}z}(k,\tau)$.

2. Complex amplitude $\left[\hat{\mathbf{a}}(k)\right]_0$ is found for the largest peak of the correlation $\mathbf{R}_0(\tau)$, while $\left[\hat{\boldsymbol{\tau}}(k)\right]_0$ is calculated as a maximum of the spline-interpolated correlation, $\mathbf{R}_0(\tau)$, using Newton-Raphson method.

3. Using the calculated parameters $\left[\hat{\mathbf{a}}(k)\right]_0$, $\left[\hat{\boldsymbol{\tau}}(k)\right]_0$ correlation is subtracted from $\mathbf{R}_{\tilde{Q}z}(k,\tau)$ to obtain a 2nd correlation peak $\mathbf{R}_1(\tau)$, by the expression $\mathbf{R}_1(\tau) = \mathbf{R}_{\tilde{Q}z}(k,\tau) - \left[\hat{\mathbf{a}}(k)\right]_0 \left[\mathbf{R}_{\tilde{Q}Q}\left(k, \tau - [\hat{\boldsymbol{\tau}}(k)]_0\right) \right]_{0,0}$.

4. Complex amplitude $\left[\hat{\mathbf{a}}(k)\right]_1$ is found for the largest peak of the correlation $\mathbf{R}_1(\tau)$, while $\left[\hat{\boldsymbol{\tau}}(k)\right]_1$ is calculated as a maximum of the spline-interpolated correlation, $\mathbf{R}_1(\tau)$, using Newton-Raphson method.

5. Using the calculated parameters $\left[\hat{\mathbf{a}}(k)\right]_1$, $\left[\hat{\boldsymbol{\tau}}(k)\right]_1$, correlation is subtracted from $\mathbf{R}_{\tilde{Q}z}(k,\tau)$ to obtain a 1st correlation peak $\mathbf{R}_0(\tau)$, by the expression $\mathbf{R}_0(\tau) = \mathbf{R}_{\tilde{Q}z}(k,\tau) - \left[\hat{\mathbf{a}}(k)\right]_1 \left[\mathbf{R}_{\tilde{Q}Q}\left(k, \tau - [\hat{\boldsymbol{\tau}}(k)]_1\right) \right]_{1,1}$.

6. Steps from 2 to 5 are repeated until the predefined stopping criterion is met.

Figs. 4 and 5 are showing multipath error envelopes for GPS C/A signal and Galileo BOC(1,1) signal, respectively, when the narrow EML delay lock loop (with 0.1 chip correlator spacing) and the MEDLL are used. The multipath error envelopes are representing change of the LOS signal ranging error in dependence of the multipath signal delay when multipath component is in the constructive phase ($\phi_1 - \phi_0 = 0°$, solid lines) and in the destructive phase ($\phi_1 - \phi_0 = 180°$, dashed lines). It can be seen that LOS signal ranging error is significantly lower for the MEDLL, regardless of signal used.

Fig. 4. Multipath error envelopes for the GPS C/A signal

Fig. 5. Multipath error envelopes for the Galileo BOC(1,1) signal

IV. PARTICLE FILTER

A. Particle filter theory

The particle filter aims to estimate, recursively in time, state $\mathbf{x}(k) \in C^{2M \times 1}$, based only on the observed data $\mathbf{z}(k) \in C^{L \times 1}$ at time index k. Particle filter is based on Bayesian estimation that follows the posterior density function $p(\mathbf{x}(k)|\mathbf{Z}_k)$ whic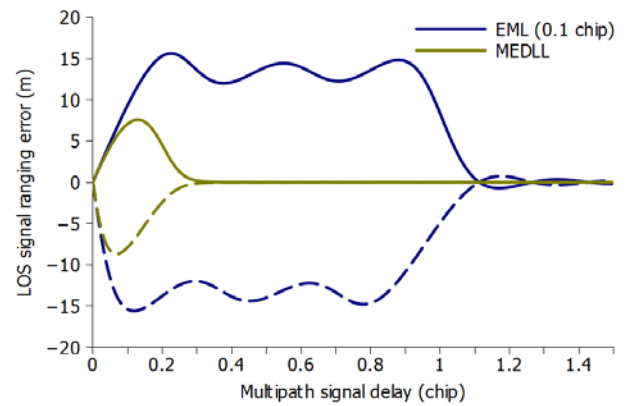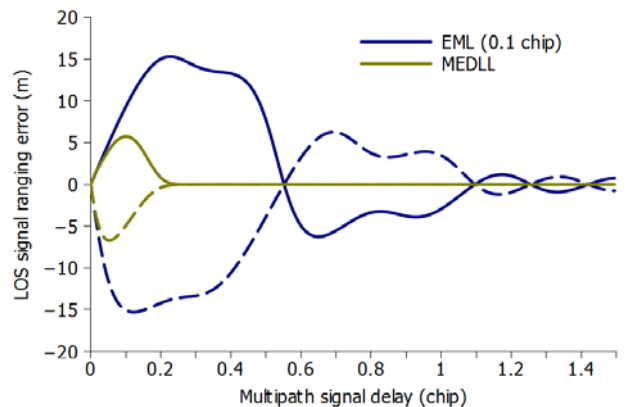h contains information about the state $\mathbf{x}(k)$, where $\mathbf{Z}_k = \{\mathbf{z}(1),...,\mathbf{z}(k)\}$ is set of observations until present time [26].

Nonlinear, non-Gaussian state space model can be written as follows

$$\mathbf{x}(k) = \mathbf{f}(\mathbf{x}(k-1),\mathbf{v}(k)), \qquad (17a)$$

$$\mathbf{z}(k) = \mathbf{h}(\mathbf{x}(k-1)) + \mathbf{w}(k), \qquad (17b)$$

where equation (17a) represents state equation of the discrete-stochastic system defining its dynamical behavior. Second equation is called measurement equation and it returns observed data.

The particle filter approximates the probability density $p(\mathbf{x}(k)|\mathbf{Z}_k)$ by a large set of P particles, $\mathbf{x}^i$, $i = 1,...,P$ where each particle has an assigned relative weight, $w^i(k)$, so that sum of all weights equals one. The location and weight of each particle reflects the value of the density in the region of the state space. The particle filter updates the particle location and the corresponding weights recursively with each new observation. The filtering density, $p(\mathbf{x}(k)|\mathbf{Z}_k)$, and the one step prediction density $p(\mathbf{x}(k+1)|\mathbf{Z}_k)$ are given by a measurement update according to following equations

$$p(\mathbf{x}(k)|\mathbf{Z}_k) = \frac{p(\mathbf{z}(k)|\mathbf{x}(k))\,p(\mathbf{x}(k)|\mathbf{Z}_{k-1})}{p(\mathbf{z}(k)|\mathbf{Z}_{k-1})}, \quad (18)$$

$$p(\mathbf{z}(k)|\mathbf{Z}_{k-1}) = \int p(\mathbf{z}(k)|\mathbf{x}(k))\,p(\mathbf{x}(k)|\mathbf{Z}_{k-1})\,d\mathbf{x}(k), (19)$$

and the time update or prediction according to

$$p(\mathbf{x}(k+1)|\mathbf{Z}_k) = \int p(\mathbf{x}(k+1)|\mathbf{x}(k))\,p(\mathbf{x}(k)|\mathbf{Z}_k)\,d\mathbf{x}(k). (20)$$

The recursion is initiated with known distribution $p(\mathbf{x}(0)|\mathbf{Z}_{-1}) = p(\mathbf{x}(0))$, where $\mathbf{Z}_{-1}$ is set without observations [26].

The likelihood $p(\mathbf{z}(k)|\mathbf{x}(k))$ is calculated from equations (11) and (17b) using the known measurement noise probability density function. This function is used for calculation of importance weights $w_i = p(\mathbf{z}(k)|\mathbf{x}^i(k))$. The aim is to approximate posterior density $p(\mathbf{x}(k+1)|\mathbf{Z}_k)$, with a sum of weighted delta-Dirac functions [26]

$$p(\mathbf{x}(k)|\mathbf{Z}_k) \approx \sum_{i=1}^{P} \tilde{w}_k^i \delta(\mathbf{x}(k) - \mathbf{x}^i(k)), \qquad (21)$$

where the normalized importance weights are defined as

$$\tilde{w}_i(k) = w_i(k) \sum_{j=1}^{P} w_j(k), \qquad i = 1,...,P. \qquad (22)$$

This approach, called Sequential Importance Sampling (SIS) often leads to divergence, where all the weights are tending to zero. Using selection or resampling step this problem can be handled [26]. The main idea behind the resampling step is to discard particles with small weights and to multiply particles with large weights, particles that corresponding to large likelihoods. This is done by drawing a new set of particles, with replacement from the old particles. A suitable measure of degeneracy of the algorithm is the effective sample size N_{eff}. This value cannot be exactly calculated, so an estimate $\hat{N}_{eff}$ is used

$$\hat{N}_{eff} = \frac{1}{\sum_{j=1}^{P} \tilde{w}_j(k)}. \qquad (23)$$

When $\hat{N}_{eff}$ is smaller than a certain user defined threshold, N_{th}, we apply the resampling step in order to decrease the variance of the importance weights [26].

Once we have approximated posterior density we can either determine particle that maximizes it, the so called *maximum a-posteriori* (MAP) estimate, or we can find the expectation, equivalent to the *minimum mean square error* (MMSE) estimate.

B. Particle filter algorithm

The complex-valued state vector $\mathbf{x}(k)$ contains delays for the LOS signal and the multipath signals and their corresponding complex amplitudes. In section II is said that for known delays it is possible to calculate amplitudes, so we are simplifying state vector with vector $\boldsymbol{\tau}(k)$. This vector contains only time delays for the LOS signal and the multipath signals and can be written as $\boldsymbol{\tau}(k) = \begin{bmatrix} \tau_0(k) & \tau_1(k) & \cdots & \tau_{M-1}(k) \end{bmatrix}^T$.

1. **Initialization**. After the acquisition, the LOS signal delay uncertainty is in range $[-T_u, T_u]$, while for the multipath signal, delay is mostly in a range $[\tau_0, 2T_c]$. So,

the particles for the LOS signal and the multipath signal delay can be initialized as

$$\tau_0^i(0) \sim U\left(-T_u, T_u\right)$$
$$\tau_m^i(0) \sim U\left(\tau_0^i, 2T_c\right), \quad i = 1, ..., P, \quad m = 1, ..., M-1 \tag{24}$$

with the weights that are equal.

2. **Importance Sampling.** Since the likelihood function is the Gaussian distribution, it is quite reasonable to propose the Gaussian importance function for particle generation. Here, for the LOS signal, importance function is realized as the Gaussian distribution with mean in the previous MAP estimate for LOS signal delay and with a variance calculated using posterior particles. Similarly, the multipath signal delay is generated using the Gaussian distribution, but in the way that newly generated values for the delay of multipath signals in the particle are larger than the delay of the LOS signal. Thus,

$$\tau_0^i(k) \sim N\left(\hat{\tau}_0^{MAP}(k), \sigma_0^2(k)\right)$$
$$\tau_m^i(k) \sim \tau_0^i(k) + \left| N\left(\hat{\tau}_m^{MAP}(k) - \tau_0^i(k), \sigma_m^2(k)\right)\right|, \tag{25}$$
$$i = 1, ..., P, \quad m = 1, ..., M-1$$

3. **Weight update and estimation**. For every particle, complex-valued signal replica is generated and $\mathbf{Q}(k, \boldsymbol{\tau})$ matrix is formed. Weights are calculated using equations (13) and (14) and then normalized with equation (18). After that, MAP estimate that maximizes posterior density based on equation (15) is found. Also, the a priori error covariance of the delay, $\boldsymbol{\Sigma}(k)$, must be calculated in order to measure estimated accuracy of the time delay. Following equation is used

$$\boldsymbol{\Sigma}(k) \approx \sum_{i=1}^{P} \tilde{w}_i(k) \left(\boldsymbol{\tau}^i(k) - \hat{\boldsymbol{\tau}}^{MAP}(k)\right)\left(\boldsymbol{\tau}^i(k) - \hat{\boldsymbol{\tau}}^{MAP}(k)\right)^T, \tag{26}$$

with $\sigma_m^2 = \left[\boldsymbol{\Sigma}(k)\right]_{m,m}$.

4. **Resampling.** According to the equation (23) value of $\hat{N}_{eff}$ is calculated. If value of $\hat{N}_{eff}$ is less then then the value of threshold N_{th}, multinomial resampling (see [26]) is performed.

V. SIMULATION AND RESULTS

The simulated GPS C/A and BOC(1,1) signals are composed of a LOS signal and one multipath component (*M*=2). Both signals are generated on intermediate frequency of $f_{IF} = 4.092$ MHz with sampling frequency $f_s = 16.368$ MHz. Before carrier removal and

spreading, signal is filtered with 6 MHz bandwidth filter. Accepted relative amplitude between LOS signal and multipath signal is $\alpha = 0.5$ while delay uncertainty T_u is $0.1T_c$. Selected signal-to-noise ratio (SNR) is -20 dB. It is supposed that phase difference between the LOS signal and the multipath signal is 10° while time-delay of multipath component is 0.2 chip.

In case of the GPS C/A signal results are obtained for two different correlation periods. The first period is 1 ms and it corresponds to duration of the GPS C/A spreading sequence, while the second period is prolonged on 4 ms. On the other hand, correlation period for BOC(1,1) signal is 4 ms and it corresponds to duration of the 4092 chips long BOC spreading sequence (f_c=1.023 MHz). The estimation results in the case of the MEDLL and the PF (for *P*=1000 particles) are shown on Figs. from 6 to 11.

Figs. 6 and 7 are showing estimated GPS C/A LOS signal delay and GPS C/A multipath signal delay, respectively, for the MEDLL and the PF filter. On the Fig. 6 can be seen that the MEDLL algorithm has less variance than the PF filter, but the MEDLL is introducing some bias with regard to the LOS signal delay.
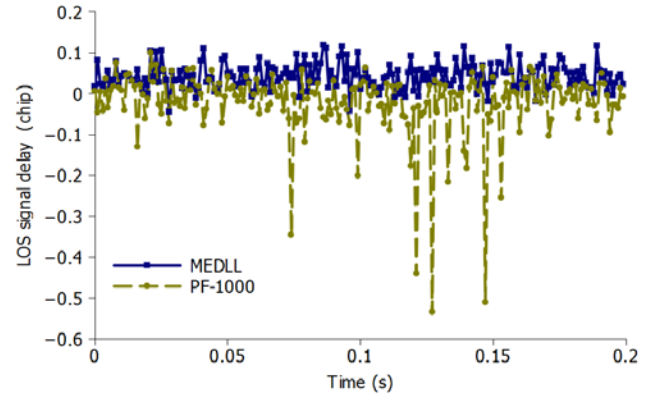
Fig. 6. Estimated GPS C/A LOS signal delay in time with correlation period of 1 ms

As can be seen on Fig. 7, the estimated GPS C/A multipath signal delay in case of the PF is much more precise then the MEDLL algorithm.
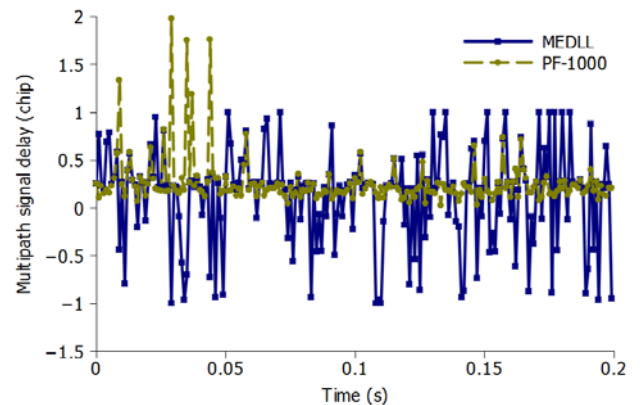
Fig. 7. Estimated GPS C/A multipath signal delay in time with correlation period of 1 ms

As shown on Fig. 8, for the correlation period of 4 ms, the estimated delay of the GPS C/A LOS signal component in the case of the PF is somewhat more precise compared to the estimation obtained using the MEDLL algorithm. Here, like on Fig. 6, the MEDLL algorithm is introducing some bias with regrad to the delay of the LOS signal component.
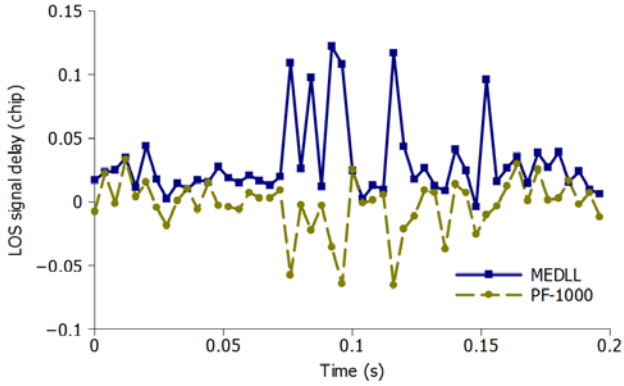
Fig. 8. Estimated GPS C/A LOS signal delay in time with correlation period of 4 ms

On Fig. 9 can be seen, that the estimation of the GPS C/A multipath signal component in the case of the MEDLL and the PF filter is almost the same with the exception of few peaks for the MEDLL algorithm.
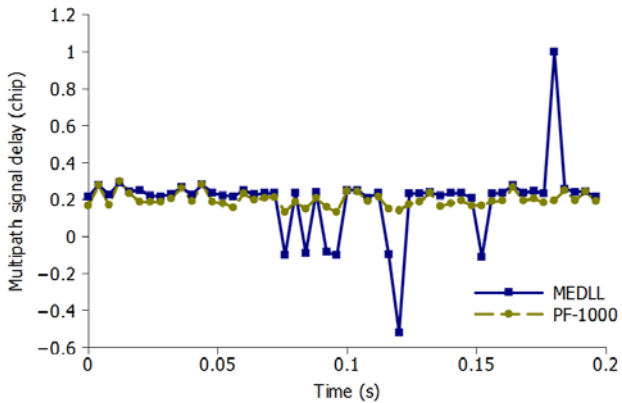
Fig. 9. Estimated GPS C/A multipath signal delay in time with correlation period of 4 ms

Figs. 10 and 11 are showing estimation for the LOS and multipath signal delay when the Galileo BOC(1,1) signal is used. As can be seen on the both figures, estimation does not favour any of the implemented algorithms, and estimation in the case of the MEDLL algorithm is similar to the estimation of the PF filter.

The MEDLL algorithm iterations are stopped when $[\hat{\tau}(k)]_0$ changes 0.1 ns between two successive iteration steps, or after 10 successive steps, whichever occurs earlier. The number of complex samples is 61. These samples are equally spread on the interval from $-2T_c$ to $2T_c$.
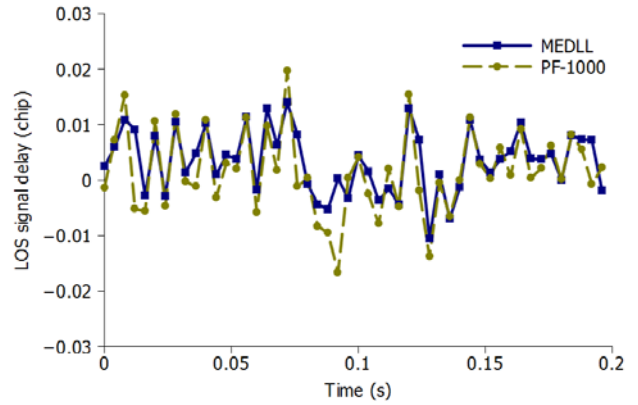
Fig. 10. Estimated BOC(1,1) LOS signal delay
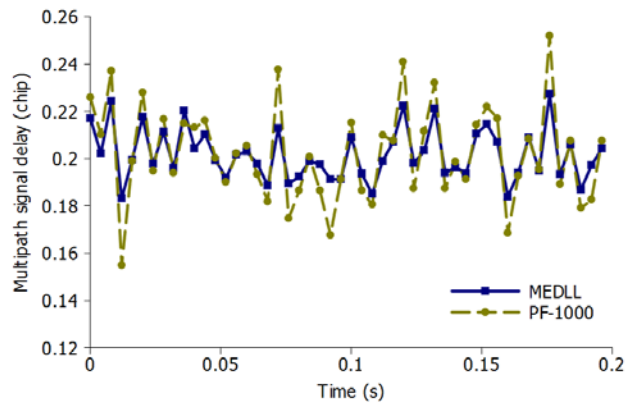
Fig. 11. Estimated BOC(1,1) multipath signal delay

IV. CONCLUSION

In this paper, two algorithms for the multipath mitigation have been presented. The simulation environment is set and composite GPS and Galileo signals are created. Using simulated signals, estimation efficiency of the algorithms is mutually compared. From the analysis it can be concluded that the particle filter, with large number of particles, is more precise then the MEDLL algorithm. This is primarily true when estimating delay of the GPS C/A signal with correlation periods of 1 ms and 4 ms. With larger correlation period MEDLL estimation would be closer to the particle filter estimation.

ACKNOWLEDGEMENT

REFERENCES

[1] Eds. E. D. Kaplan, C. J. Hegarty, "Understanding GPS – Principles and Applications", Second edition, Artech House, London, 2006

[2] A. J. Van Dierendonck, P. Fenton and T. Ford, "Theory

and Performance of Narrow Correlator Spacing in a GPS Receiver", Navigation: Journal of the Institute of Navigation, Vol. 39, No. 3, Fall 1992

[3] K. Borre, D. M. Akos, N. Bertelsen, P. Rinder and S. H. Jensen, "A Software-Defined GPS and Galileo Receiver - A Single-Frequency Approach", Birkhäuser, Berlin, 2007

[4] H. So, G. Kim, T. Lee, S. Jeon and C. Kee "Modified High-Resolution Correlator Technique for Short-Delayed Multipath Mitigation", The Journal of Navigation, Vol. 62, pp. 523-542, 2009

[5] V. A. Veitsel, A. V. Zhdanov and M. I. Zhodzishsky, "The Mitigation of Multipath Errors by Strobe Correlators in GPS/GLONASS Receivers", GPS Solutions, Vol. 2, No, 2, pp. 38-45, 1998

[6] A. Schmid, A. Neubauera, H. Ehmb, R. Weigel, N. Lemke, G. Heinrichs, J. Winkel, J. A. Ávila-Rodríguez, R. Kaniuth, T. Pany, B. Eissfellerd, G. Rohmer and M. Overbeck, "Combined Galileo/GPS architecture for enhanced sensitivity reception", AËU - International Journal of Electronics and Communications, Elsevier, Vol. 51, No. 1, pp. 1-8, 2004

[7] B. R. Townsend, P. C. Fenton. "A Practical Approach to the Reduction of Pseudorange Multipath Errors in a Ll GPS Receiver", in Proceedings of the 7th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1994), Salt Lake City, UT, Sept. 1994, pp. 143-148.

[8] J. Jones, P. Fenton and B. Smith, "Theory and Performance of the Pulse Aperture Correlator", Technical Report., Novatel, Alberta, Canada, Sept. 2004

[9] R. Hamila, E. S. Lohan and M. Renfors, "Subchip Multipath Delay Estimation for Downlink WCDMA System Based on Teager–Kaiser Operator", IEEE Communications Letters, Vol. 7, No. 1, pp. 1-3 2003

[10] J. Selva, "Complexity reduction in the parametric estimation of superimposed signal replicas", Signal Processing, Elsevier, Vol. 84, pp. 2325–2343, 2004 W.

[11] M. Lentmaier, B. Krach, "Maximum Likelihood Multipath Estimation Comparison with Conventional Delay Lock Loop", in Proceedings of 19th International Technical Meeting of the Institute of Navigation Satellite Division (ION GNSS 2006), pp. 1742-1751, Fort Worth, Texas, USA, Sep. 2006

[12] M. Sahmoudi, M. G. Amin, "Fast Iterative Maximum-Likelihood Algorithm (FIMLA) for Multipath Mitigation in Next Generation of GNSS Receivers", in Proceedings of Fortieth Asilomar Conference on Signals, Systems and Computers (ACSSC '06), pp. 579 – 584, 2006

[13] M. Sahmoudi, M. G. Amin, "Robust tracking of weak GPS signals in multipath and jamming environments", Signal Processing, Elsevier, Vol. 84, No. 7, pp. 1320-1333, 2009

[14] R. D. J. van Nee, "The Multipath Estimating Delay Lock Loop", in Proceedings of IEEE Second International Symposium on Spread Spectrum Techniques and Applications (ISSSTA'92), Yokohama, Japan, November 29-December 2, 1992

[15] N. Delgado, F. Nunes, "Theoretical Performance of the MEDLL Algorithm for the New Navigation Signals", in Proceedings of Conference on Telecommunications - ConfTele, Aveiro, Portugal, Vol. 1, pp. 1 - 4, May 2009

[16] J. Soubielle, I. Fijalkow, P. Duvaut, A. Bibaut, "GPS Positioning in a Multipath Environment", IEEE Transactions on signal processing, Vol. 50, No. 1, 2002

[17] M. Z. H. Bhuiyan, E. S. Lohan, and M. Renfors, "Code Tracking Algorithms for Mitigating Multipath Effects in Fading Channels for Satellite-Based Positioning", EURASIP Journal on Advances in Signal Processing, Article ID 863629, 17 pages, 2008

[18] P. C. Fenton, J. Jones, "The Theory and Performance of NovAtel Inc.'s Vision Correlator", in Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005), Long Beach, CA, pp. 2178-2186, Sept. 2005

[19] F. Antreich, J. A. Nossek, "Maximum Likelihood Parameter Estimation in a GNSS Receiver", in Proceedings of Conference on Wave Propagation in Communication, Microwave Systems and Navigation (WFMN07), Chemniz, Germany, July 2007

[20] R. A. Iltis, "Joint Estimation of PN Code Delay and Multipath Using the Extended Kalman Filter", IEEE Transactions on Communications, Vol. 38, No. 10, pp. 1667-1685, 1990

[21] N. I. Ziedan, "GNSS Receivers for Weak Signals", Artech House, London, 2006

[22] G. Yuan, Y. Xie, Y. Song, H. Liang, "Multipath parameters estimation of weak GPS signal based on new colored noise unscented Kalman filter", in Proceedings of 2010 IEEE International Conference on Information and Automation (ICIA), Harbin, China, pp. 1852 – 1856, June 2010

[23] P. Closas, C. Fernandez-Prades, J. A. Fernandez-Rubio, "Bayesian DLL for Multipath Mitigation in Navigation Systems Using Particle Filters", in Proceedings of 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toulouse, France, pp. IV - IV, September 2006

[24] M. Lentmaier, B. Krach and P. Robertson, "Bayesian Time Delay Estimation of GNSS Signals in Dynamic Multipath Environments", International Journal of Navigation and Observation, Article ID 372651, 11 pages, 2008

[25] C. Fernandez-Prades, J. A. Fernandez-Rubi, G. Seco, "Joint Maximum Likelihood Estimation of Time-Delays and Doppler Shifts", in proceedings of Seventh International Symposium on Signal Processing and Its Applications, pp. 523 - 526 vol. 2 , July 2003

[26] J. V. Candy , "Bayesian Signal Processing – Classical, Modern and Partical Filtering Methods", John Wiley & Sons, 2009

Simulation and Modeling of Integrated Hall Sensor Devices

Nebojša Janković, Sanja Aleksić and Dragan Pantić

Abstract - In this paper the reviews of 3D simulation procedure and modelling of Hall sensor realized in standard high-voltage CMOS technologies are given. The complete manufacturing process flow and electrical characteristics of cross-shaped Hall sensor, vertical Hall sensor and MAGFET are simulated by using Silvaco and ISE TCAD software package tools. In addition, the efficient electrical models of these devices are derived and successfully implemented in circuit simulator SPICE.

Keywords **-** TCAD, Hall cross-shaped sensor, vertical Hall sensor, MAGFET, equivalent circuit model[1], SPICE.

I. INTRODUCTION

A variety of integrated magnetic sensitive passive and/or active devices can be designed and fabricated in standard CMOS IC technology without design rule violations [1]. In the family of such devices, the Hall plates and magnetic-field-sensitive MOSFETs (MAGFETs) appears to be the most popular integrated structures for sensor applications. Analysis of the integrated magnetic sensor devices based on the Hall effect or carrier deflection has generally been based hitherto on simple, intuitive analytical models. While some of these models remain valuable heuristic tools for trial device design and analysis in certain limiting cases, they are inappropriate for general device structures and operating conditions. Consequently, a precise physical simulation of magnetic sensors is highly desirable in order to optimize the design and operating conditions of these sensors with respect to high magnetic sensitivity [2].

Unlike the simulation of conventional semiconductor IC devices, the numerical simulations of magnetic microsensors are relatively new. Namely, the vectorial nature of the Lorentz force requires complex analysis in the space that can be properly made only by three-dimensional (3D) simulations. Only recently, the modules for 3D simulations of semiconductor magnetic sensors have become available as a part of commercial software packages such as Silvaco [3,4] or Sentaurus/ISE TCAD [5].

In this work, we present the results of the 3D numerical simulations of integrated Hall plates and MAGFETs using the commercial Technology Compute Aided Design (TCAD) software. Subsequently, based on the 3D magnetic sensor simulations, the efficient electrical models of these devices derived from will be described and their successful implementation in circuit simulator SPICE will be demonstrated.

II. 3D CARRIER TRANSPORT WITH MAGNETIC FIELD IN SEMICONDUCTORS

The effect of the magnetic field $\boldsymbol{B}$ on a carrier travelling with velocity v is to add a term $q \cdot (v \times \boldsymbol{B})$ called the Lorentz force to the force that the carrier already feels. The magnetic field density $\boldsymbol{B}$ is a vector (B_X, B_Y, B_Z) in units of Tesla (T=Vs/m2). The Hall coefficients R_n, R_p characterize the transverse quasi Fermi level gradient caused by the magnetic field acting on the electron (hole) current density vector $\boldsymbol{J}_{n,p}$. Assuming the isothermal condition $\nabla T = 0$, R_n ,R_p are expressed as:

$$R_{n,\,p} = -\frac{\nabla \phi_{n,\,p} \cdot (\boldsymbol{B} \times \boldsymbol{J_{n,p}})}{(\boldsymbol{B} \times \boldsymbol{J_{n,p}})^2} \tag{1}$$

where $\nabla \phi_{n,\,p}$ is the gradient of respective electron and hole Fermi potentials. In the following analysis, only electrons will be considered while the same formalisms also hold for holes.

Based on the solution of Boltzmann's transport equation under a relaxation time approximation, the isothermal magnetic field dependent electron current density $\boldsymbol{J}_n$ in isotropic semiconductor material can be represented in the implicit form as [6]:

$$\nabla \phi_n = -\sigma_n^{-1} \cdot \boldsymbol{J_n} - \boldsymbol{B} \times (R_n \cdot \boldsymbol{J_n}) \tag{2}$$

The deflection that the magnetic field causes on electric currents in semiconductors is reflected by vector products with $\boldsymbol{B}$ appearing on the right-hand side of Eq. (2). The closed form analytical solution of $\boldsymbol{J}_n$ can be obtained from Eq. (2) only under the assumption of low magnetic fields. Then it yields:

$$\boldsymbol{J_n} = -\sigma_n \nabla \phi - \sigma_n \frac{1}{1 + (\mu_n^* \boldsymbol{B})^2} \cdot$$
$$\cdot \left[\mu_n^* \boldsymbol{B} \times \nabla \phi_n + \mu_n^* \boldsymbol{B} \times (\mu_n^* \boldsymbol{B} \times \nabla \phi) \right] \tag{3}$$

where $\mu_n^* = R_n \cdot \mu_n$ denote the Hall mobility of electrons. The electron current density vector $\boldsymbol{J_n}$ described by Eq. (3) can be transformed in a matrix form as:

[1] Nebojša Janković, Sanja Aleksić and Dragan Pantić are with the Department of Microelectronics, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia,
E-mail: {nebojsa.jankovic,sanja.aleksic.dragan.pantic}@elfak.ni.ac.yu.

$$\boldsymbol{J_n} = \boldsymbol{J_{n0}} \cdot M = (-\sigma_n \cdot \nabla \phi_n) \cdot M \qquad (4)$$

with:

$$M = \frac{1}{1+a^2+b^2+c^2} \begin{Bmatrix} 1+a^2 & ab-c & ca+b \\ c+ab & 1+b^2 & bc-a \\ ca-b & a+bc & 1+c^2 \end{Bmatrix} \qquad (5)$$

where $\boldsymbol{J_{n0}}$ is the zero-magnetic field electron current density, $a=\mu*B_X$, $b=\mu*B_Y$, and $c=\mu*B_Z$. The matrix form of $\boldsymbol{J_{n,p}}$ represented by Eqs. (4) and (5) is used for both electrons and holes to include the magnetic field effects in ATLAS [3]. Note that the both Eqs. (3) and (4) are derived from an expansion of Eq.(2) in powers of magnetic field $\boldsymbol{B}$ in the approximation of low magnetic fields. More precisely, they are accurate only if the weak magnetic field condition $\mu^* \cdot \boldsymbol{B} \ll 1$ is satisfied. Thus, for example, in case of a field of 1 Tesla applied to silicon crystal with a typical carrier mobility of 0.1 m^2/(Vs), the product $\mu^* \cdot B$ is 0.1 so that the weak field condition is satisfied. It is important to emphasize that all magnetic device simulations described in this work are performed with $\boldsymbol{B} \leq 1$ Tesla in order to preserve the validity of Eqs. (3) and (4) and obtain realistic results. Hence, the magneto-resistance effects appearing as the change of material electrical resistance at extremely high magnetic fields $\boldsymbol{B} > 3$ Tesla are not possible to simulate with present commercial device simulators.

III. TCAD OF INTEGRATED HALL CROSS-SHAPED AND VERTICAL HALL SENSOR

A. TCAD of cross-shaped Hall sensor in CMOS technology

A cross-shaped Hall sensor fabricated in bulk CMOS technology cannot be properly simulated without the inclusion of 3rd dimension. A complete fabrication process flow of cross-shaped Hall sensor was simulated using parameters of standard AMS 0.8μm CMOS technology. The device is realized in deep N-well region with additional p$^+$layer in the middle of the sensor between contacts C1 and C2, and sensing contacts C3 and C4. Using process simulator DIOS [5], a several 2D doping profiles along sensor's main x- and y-axes were obtained and one of them, in the middle of Hall sensor structure, is shown in Fig.1 [7]. Then, a 3D device structure is generated by data exchange and interpolation between simulated 2D cross sections using program DIP [5]. The obtain 3D Hall sensor structure and discretization grid are shown on Fig. 2.

The Hall sensor electrical characteristics were obtained by device simulator DESSIS [5]. A standard drift diffusion model with doping-dependant Hall mobility, Shockley-Read-Hall and Auger recombination models were used in simulations. Also, two identical voltmeters with 10GΩ input

resistances were assumed to be attached to the Hall sensing contacts. Fig.3 shows the examples of 2D electron current densities along XZ and YZ planes and 3D potential distributions simulated for case of homogenous perpendicular magnetic field B=2T and V_{IN}=4V. The influence of $\boldsymbol{B}$ on current deflection and potential difference between Hall contacts is clearly visible. In addition, result in Figs.3a) and 3b) confirms the beneficial influence of shallow P$^+$ diffused layer (see Fig.1) pushing the peak electron current toward lower doped N-well region with higher Hall coefficients [8]. Finally, 3D potential distribution in the whole simulation domain of Hall cross sensor for B_Z=2T and V_{IN}=2V on C2 contact is shown on Fig. 4.

Fig.4
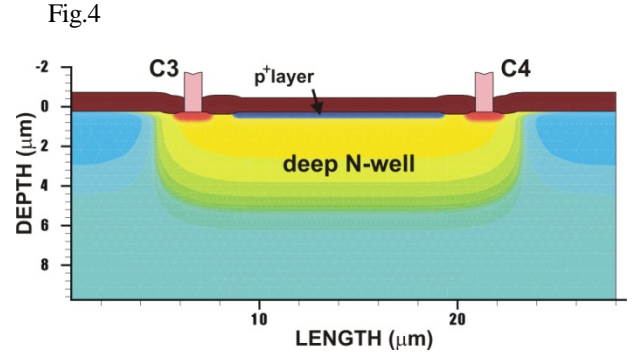
Fig. 1. 2D doping profile in the middle of Hall sensor structure through sensing contacts C3 and C4.
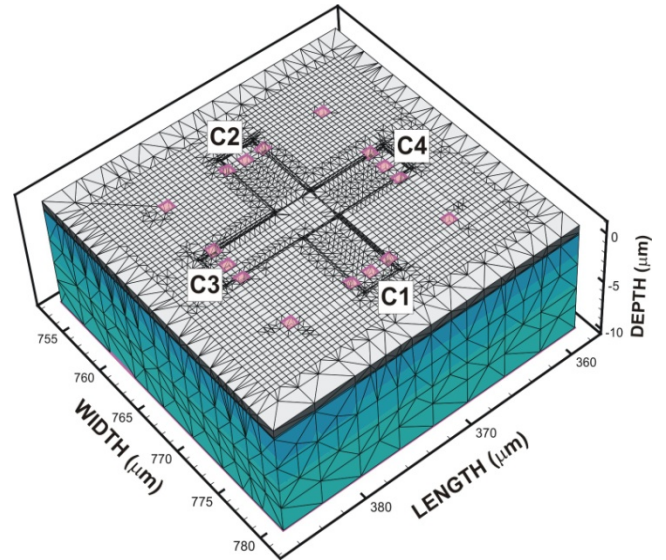
Fig. 2. 3D discretization grid in the Hall sensor simulation domain.

B. TCAD of vertical Hall sensor in CMOS technology

The vertical Hall sensor (VHS) layout with five contacts in a line on top of a low-doped n-diffusion region, surrounded by p-diffusion layer, is based on geometry given in [9,10]. VHS is also realized in high-voltage AMS 0.8μm CMOS technology. Since preconditions for
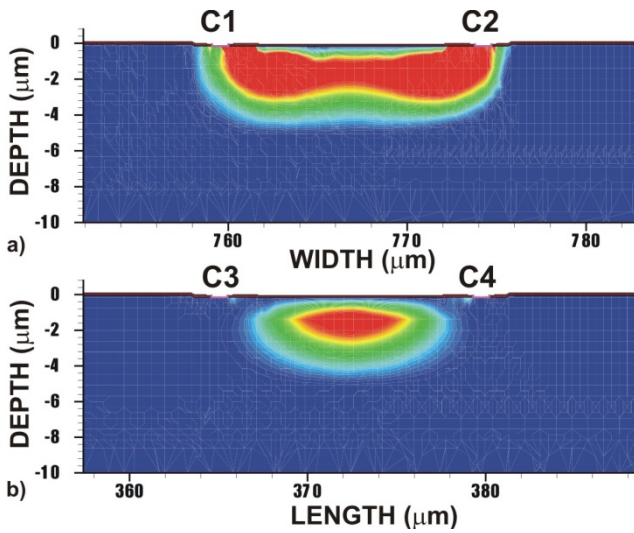
Fig. 3. Electron current density in: a) cross section through C1 and C2 contacts, and b) cross section through sensitive contacts C3 and C4 (B=2T, V_{IN}=4V).
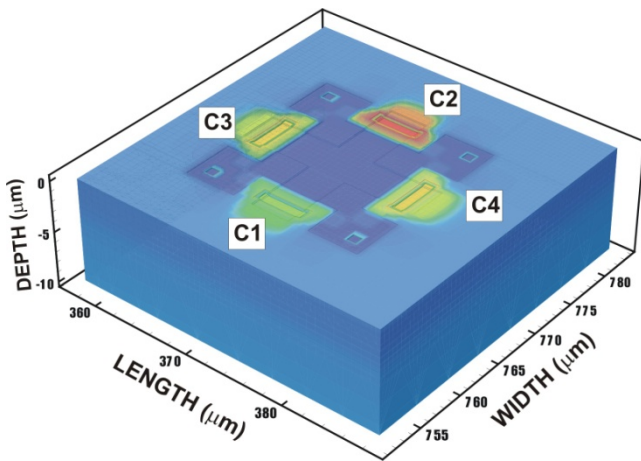
Fig. 4. 3D potential distribution for B=2T and V_{IN}=4V.

achieving high magnetic sensitivity of the VHS are low-doped and deep active area profiles, it is easy to understand why this technology was the first choice, because we realize sensor in deep n-diffusion layer (DNTUB, depth 7μm), on p-substrate. In accordance with the conventional design of the VHS, contacts sizes are 1.5x1.5μm, while the distance between contacts is 10μm.

For the process simulation of the vertical Hall sensor 2D process simulator DIOS and 3D doping profile generator MESH, as a part of the ISE TCAD system, were used. As a result, 3D doping profile of VHS sensor obtained by using the doping reduction method [11] and additional p+ region between contacts is shown in Fig. 5. The electrical characteristics of VHS, for biasing conditions: V_{IN}=5V, V_{OUT}=0V, and magnetic field B=0.5T and 1T) were simulated using 2D/3D device simulator DESSIS. For the modeling of the measurement, the VHS is connected to voltmeters with an input resistance of 10 MΩ

at both sensitive contacts (S_1 and S_2). The obtained simulation results, current sensitivity S_I and voltage sensitivity S_V, are: S_I=348V/(AT) and S_V=0.274V/(VT) for B=0.5T and S_I=318V/(AT) and S_V=0.259V/(VT) for B=1T. The potential and electron current density distributions in the half domain of VHS for B (B_X=0, B_Y=0, B_Z=1T) are shown on Figs. 6 and 7.
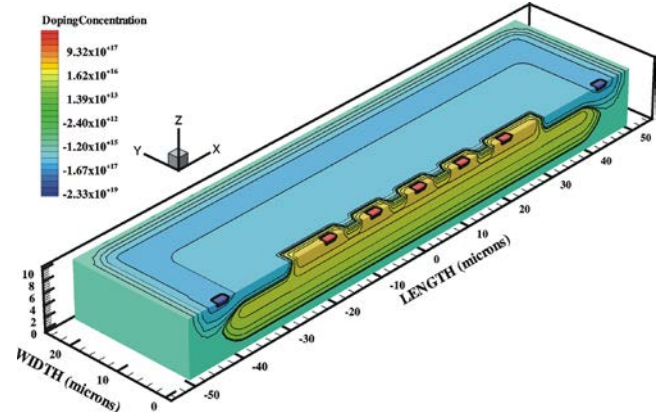
Fig. 5. 3D doping profile in the half domain of VHS realized in 0.8μm HV-CMOS technology
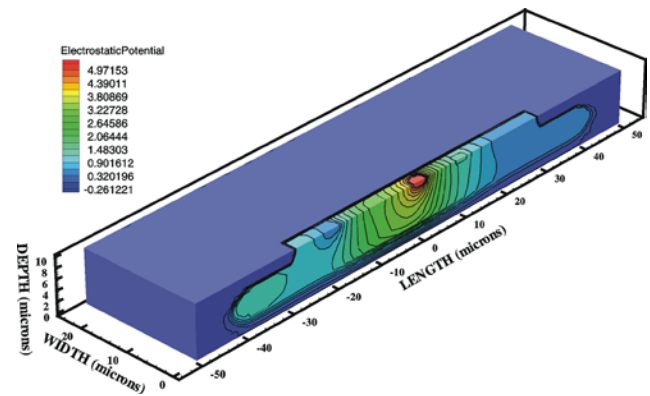
Fig. 6. Potential distribution of VHS for V_{IN}=5V and B=1T.
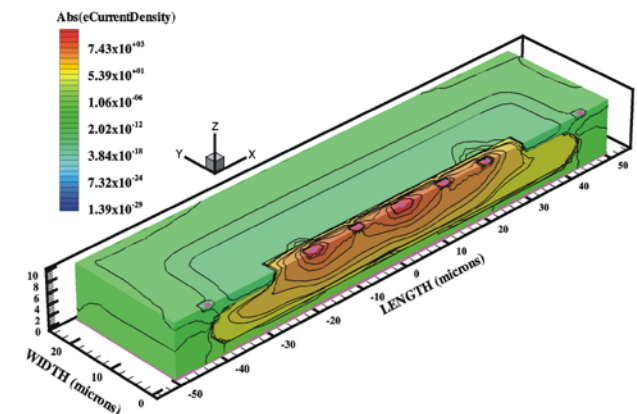
Fig.7. Electron current distribution of VHS for V_{IN}=5V and B=1T.

C. Equivalent circuit model of Hall plates

A unified circuit model developed for both vertical and horizontal Hall sensors is shown in Fig. 8. In contrast to more complex matrix models with junction field effect transistors (JFETs) [6], the circuit model shown in Fig.4 is a symmetrical four-cell lumped element model where JFETs are replaced with the non-linear resistors R_{NT}. The resistivity of R_{NT} is calculated with formula:

$$R_{NT} = \left[a + b \cdot \exp\left(\frac{V_{NT}}{c} \right) \right] \cdot \left(\frac{T}{T_{nom}} \right) \cdot \left(1 + d \cdot B^2 \right) \quad (6)$$

where V_{NT} is the resistor's internal voltage drop and T_{nom} is the referent temperature (room temperature). *a*, *b*, c and *γ* are parameters extracted from fitting Eq.6 with the results of 3D simulations of Hall plates (Figs. 3-7) performed for different supply voltages with *B*=0. The last multiplayer term in Eq.6 represents the magneto-resistance effect which becomes important at high magnetic fields. Constance *d* is extracted from measuring the experimental Hall plates since the magnetoresistance effects are not possible to simulate with present TCAD software as explain in Section II of this paper. The other circuit elements F_{XYI} and F_{YXI} shown in Fig.8 are the current sources controlled by currents trough the respective zero-voltage ampermeters V_{XYI} and V_{YXI} as indicated by arrows. The *X*, *Y* current components when mirrored must be multiplied by magnetic and spatial dependant coefficients from V_{XYI} and V_{YXI} as described in [6]. The D1-D5 elements are the inversely polarised unity diodes used to model the p-n junction distributed capacitances in the sensor. Proper modeling the sensor's time response is important in case of, for example, performing dynamic offset and noise cancellation.

D. Modeling results

The electrical circuit model of Hall plates shown in Fig.8 has been implemented in SPICE circuit simulator. The magnetic field is represented with a separate voltage generator sourcing a voltage equal in magnitude to B_Z. The model parameters defined in previous section has been extracted from fitting the modeling results with measured vertical and horizontal Hall plates characteristics fabricated in 0.8μm CMOS technology. The efficiency of magnetic sensor circuit model to predict the electrical and sensory characteristics of Hall plates are demonstrated in Figs.9-14.
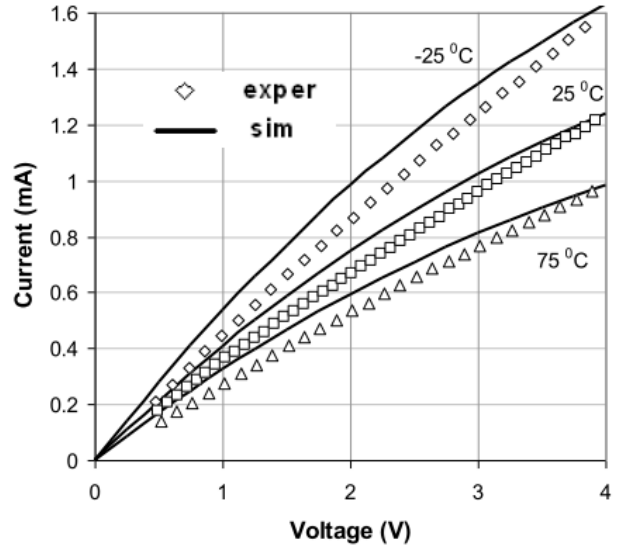
Fig. 9. Modelled and experimental the current–voltage characteristics of a cross-shaped Hall sensor.

Fig. 8. Unified equivalent circuit model of vertical and/or horizontal integrated Hall plates realized in standard CMOS technologies.
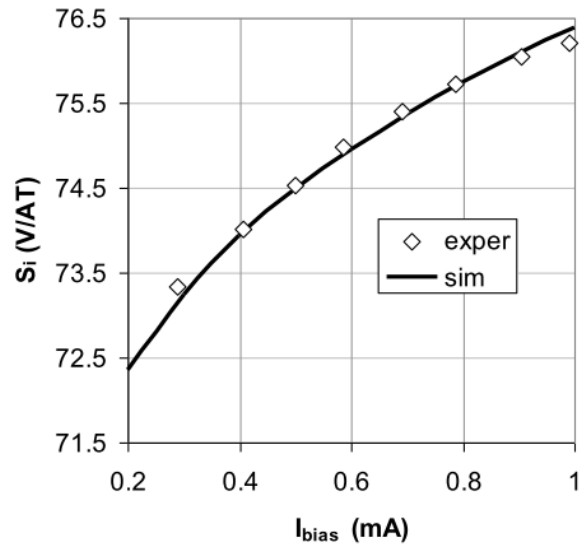
Fig. 10. The supply current-related sensitivity on the bias current in vertical Hall sensor.
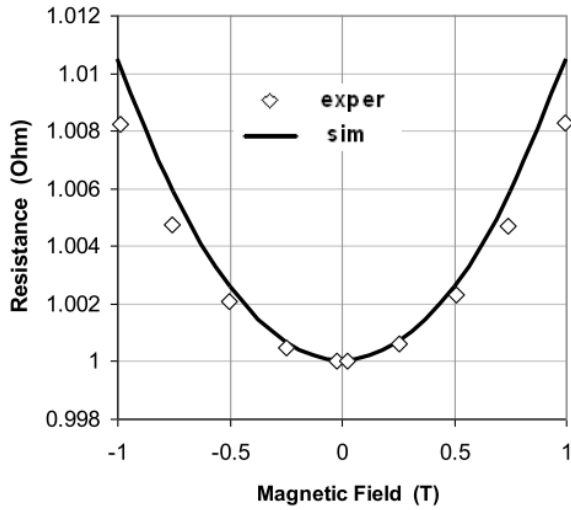
Fig. 11. Modeling and measurements of the magnetoresistivity effects.

Fig. 14. Modelled and experimental Hall voltage response in vertical Hall sensor to the magnetic field.

IV. TCAD OF MAGFETs

The inversion layer of MOSFET can be used as the active region of a magnetic sensor. This active region can exploit the Hall effect for Hall based sensors, or the carrier deflection, if the device has a split-drain. The structure of conventional Split-Drain MOSFET (MAGFET) is identical to a MOSFET but the drain is split in two or more parts as shown in Fig.15. The ability of integrating the bias and control circuitry on the same chip with MAGFET device makes this sensor structure particularly attractive.

Fig. 12. The reduction of Hall voltage at high frequencies, due to the presence of the distributed diode capacitance in a cross-shaped Hall sensor.

A. TCAD study of MAGFET in CMOS technology

A MAGFET with L=125μm, W=100μm, t_{ox}=60nm gate oxide, and substrate doping N_D=10^{15}cm^{-3}, is studied in our case. A concave MAGFET mask layout and standard 0.35μm CMOS technology are adopted for process simulation, yielding 45μm wide drain regions separated by a 10μm oxide gap. The internal potentials and carrier distributions of the MAGFET in presence of the perpendicular magnetic field B_Z were then obtained using the 3D device simulator ISE DESSIS. Fig. 16 shows the electric field distribution in the channel simulated for V_{GS} =5V, V_{DS} =1V and B_Z =100mT, where B_Z was orientated in the z-axis direction. It can be seen that the electric field iso-lines are asymmetrical with respect to the (z,x)-plane at $y = 0$. This asymmetry is caused by the accumulation of electrons in the upper channel region due to the influence of Lorentz force. It also causes the difference in drain currents at D1 and D2 contacts. The later is illustrated by Fig. 17 showing the drain current density distribution in the channel of MAGFET simulated without and with magnetic field B_Z [12].

Fig. 13. Modelled and experimental a current-related magnetic response of cross-shaped Hall sensor.
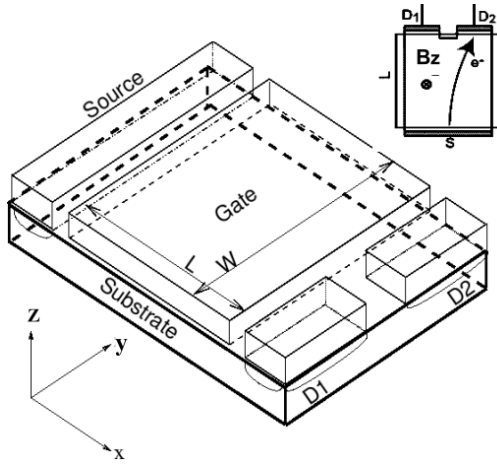
Fig. 15. The 3D structure of MAGFET with carrier deflection shown in the inset.
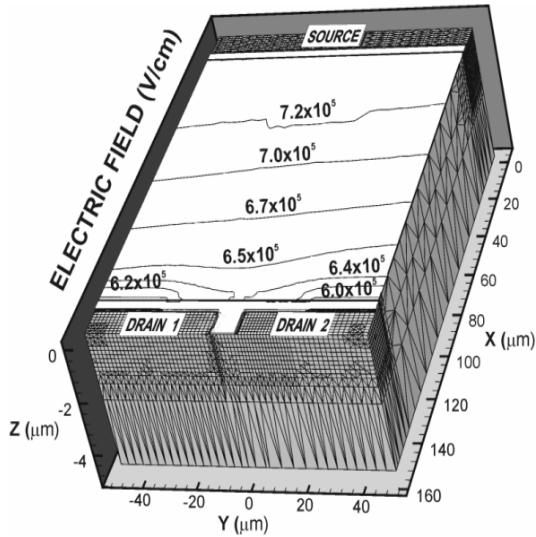
Fig. 16 Electric field distribution in the channel of MAGFET obtained with 3D simulations for B_Z=100mT.

B. Equivalent circuit model of MAGFETs

The MAGFET operation is emulated with two identical NMOSTs operating in parallel. The channel carrier transport has to be represented with two identical RC chains as illustrated in Fig. 3. Depending on the sign (+ or -) of the applied perpendicular magnetic field B_Z, the equivalent resistors R_k in one of channel chains will simultaneously decrease or increase under the action of the Lorentz force due to carrier accumulation or depletion, respectively. In the expressions underlying the distributed MOSFET model [14], the magnitude of R_k is inversely proportional to the square root of the substrate doping

concentration e.g. $\sqrt{N_{beff}}$ (see Eq. (A4) in Ref [13]). Hence, in order to include magnetic effects into the existing MOSFET model, the new effective substrate doping variable N'_{beff} is defined instead of the N_{beff} parameter as:

$$N'_{beff} = N_{beff} \pm \Delta n\left(x, B_Z\right) = N_{beff} \pm a \cdot B_Z \qquad (7)$$

where $+$ and $-$ signs stand for the different directions of carrier deflection in one of the NMOST channels as illustrated in Fig. 18. The Eq. (7) is the key modification to the MOSFET model [14]. It is obtained from TCAD study of MAGFET showing that there is approximately linear relation between B_Z and accumulated Hall charge at the one side of the channel. The empirical constant a appearing in (4) becomes fitting parameter used to calibrate the MAGFET model. When B_Z=0, the MAGFET model reverts to the original MOSFET model [14].

a)

b)
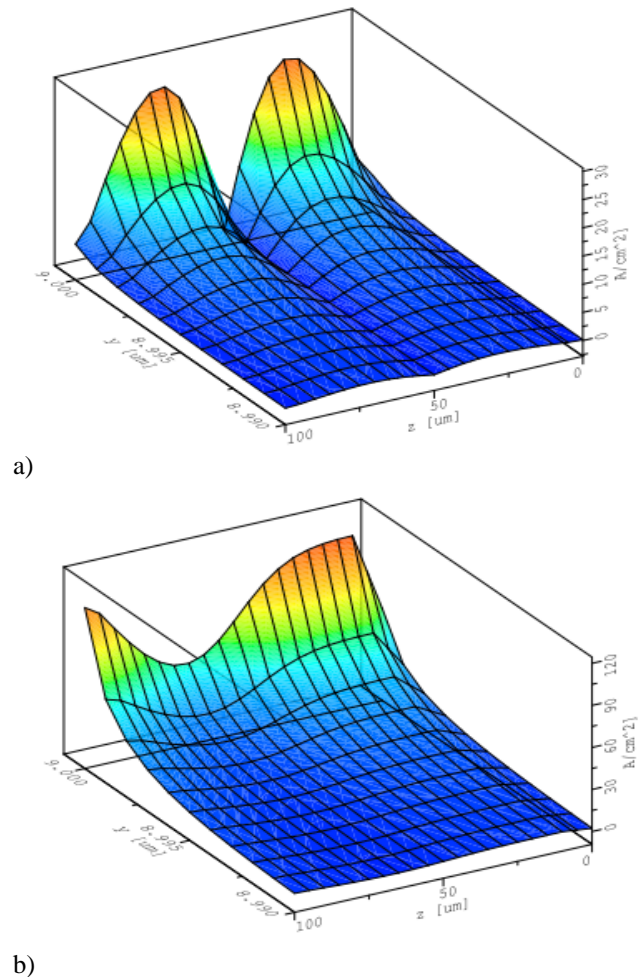
Fig. 17. Drain current density distributions near D1 and D2 MAGFET's contacts simulated without (a) and with magnetic field B_Z=40mT (b) [12].
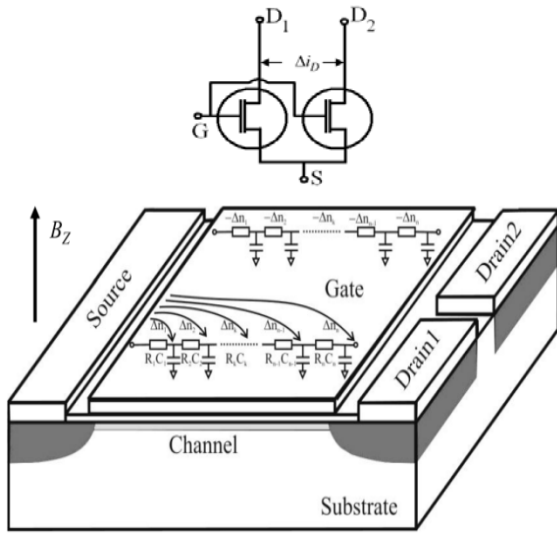
Fig. 18. Split-drain MAGFET represented with the two magnetic sensitive NMOSTs

C. Modeling results

The MAGFET model is implemented in SPICE in the form of a sub-circuit with two NMOSTs as illustrated in Fig.18. As in the case of Hall plates simulation with SPICE, the magnetic field is also represented here with a separate voltage generator sourcing a voltage equal in magnitude to B_Z. This voltage source drives a special "magnetic" node in the MAGFET sub-circuit that connects B_Z with the N'_{beff} variable of the modified MOSFET model following Eq. (7).

Fig.19 shows the comparisons between 3D simulations and modeling results with experimental data taken from [13] of drain current imbalance $\Delta i_D = I_{D1} - I_{D2}$, while the relative sensitivities of MAGFET versus V_{DS} and V_{GS}, for B_Z=0.1T are shown on Figs. 20 and 21.
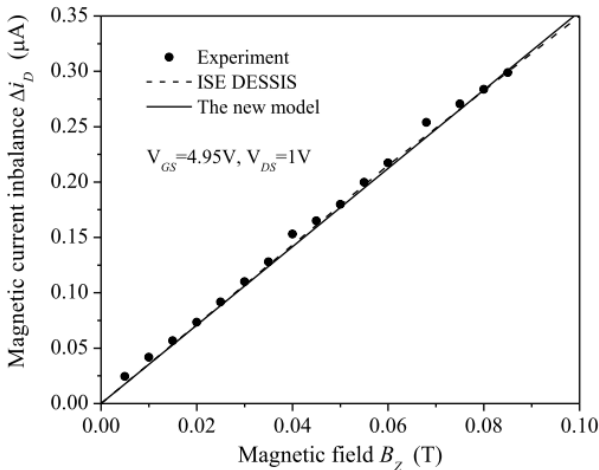
Fig. 19 The simulated, modeled and experimental MAGFET current imbalance Δi_D dependence versus the magnetic field B_Z.
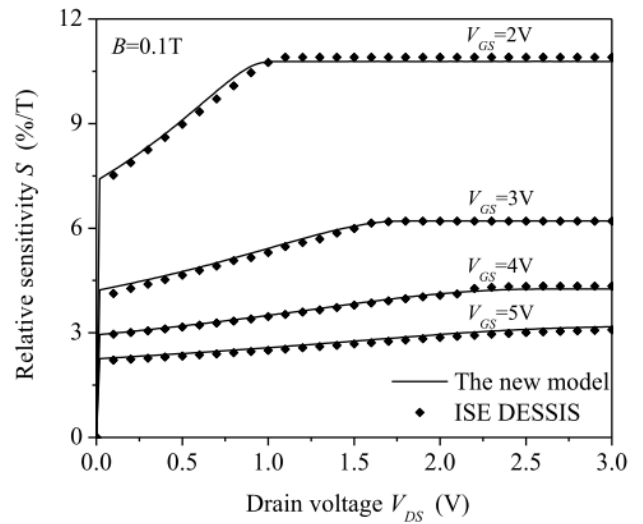
Fig. 20. Relative sensitivity S of MAGFET versus: V_{GS} extracted from 3D device numerical simulations and from the SPICE MAGFET model.
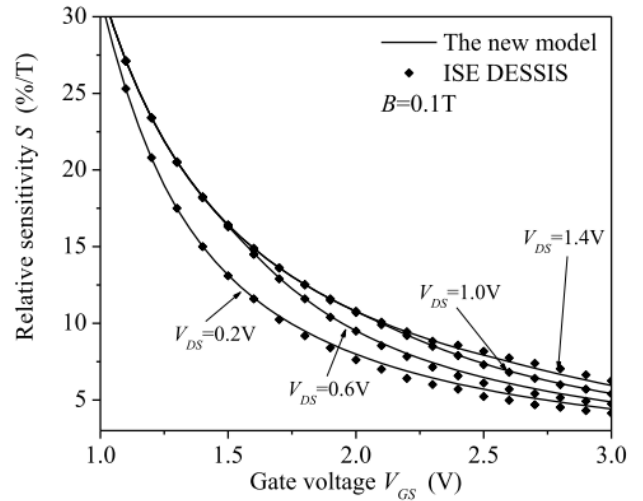
Fig. 21. Relative sensitivity S of MAGFET versus V_{DS} extracted from 3D device numerical simulations and from the SPICE MAGFET model.

V. CONCLUSION

In this work, the results of 3D TCAD of integrated Hall sensor devices manufactured by using the standard AMS 0.8μm and 0.35μm high-voltage CMOS technologies are presented. The complete technology process flow and electrical characteristics of cross-shaped Hall sensor, vertical Hall sensor and MAGFET are simulated by using Silvaco (ATHENA, ATLAS) and ISE (DIOS, DESSIS, MESH, DIP) TCAD software package tools. In addition, based on the 3D numerical simulation of magnetic sensors, the efficient electrical models of these devices are derived and successfully implemented in circuit simulator SPICE.

ACKNOWLEDGEMENT

REFERENCES

[1] Baltes, H.P., Popovic, R.S., *"Integrated Semiconductor Magnetic Field Sensor"*, Proceedings of the IEEE, Vol. 74, No. 8, August, 1986, pp. 1107-1132.

[2] Allegretto, W, Nathan, A., Baltes, H, *"Numerical Analysis of Magnetic Field Sensitive Bipolar Devices"*, IEEE Trans. Computer-Aided Design, Vol. 10, No. 4, Feb., 1991, pp. 501-511.

[3] *ATHENA User's Manual – Process Simulation Software*, SILVACO, Santa Clara, USA, 2009.

[4] *ATLAS User's Manual – Device Simulation Software*, SILVACO, Santa Clara, USA, 2009.

[5] *ISE TCAD – User Manual, Rel. 7.0*, Integrated System Engineering AG, Zurich, Switzerland.

[6] Wachutka, G., *"Unified Framework for Thermal Electrical, Magnetic and Optical SemiconductorDevice Modeling"*, COMPEL, Vol. 10, No. 4, 1991, pp. 311-321.

[7] Jovanovic, E., Pesic, T., Pantic, D., *"3D Simulation of Cross- Shaped Hall Sensor and its Equivalent Circuit Model"*, Proc. of 24th International Conference on Microelectronics (MIEL'04), Vol. 1, Nis, Serbia, May 2004, pp. 235-238.

[8] Popović, R.S., *Hall Effect Devices,* Second edition, IOP Publishing Ltd, Bristol and Philadelphia, 2004.

[9] Popović, R.S., *"The Vertical Hall-effect Device",* IEEE Electron Dev. Lett., EDL-5, No. 9, 1984, pp. 357-358.

[10] Schuring, E., Demierre, M., Schott, C., Popović, R.S., *"A Vertical Hall Device in CMOS High-voltage Technology",* Sensors and Actuators A: Physical, Vol. 97-98, No. 1, April 2002, pp. 47-53.

[11] Jovanovic, E., Pantic, D., Pantic, D., *" Simulation of Vertical Hall Sensor in High-voltage CMOS Technology "*, Proc. 6th International Conference on Telecomunication in Modern Satelite, Cable and Broadcasting Services (TELSIKS'03), Vol. 2, Nis, Serbia, October 2003, pp. 811-814.

[12] Rodrigez-Torres, R., Gutierrez-D., E.A, Klima, R., Selberherr, S., *"Three-Dimensional Simulation Split-Drain MAGFET at 300K and 77K",* Proc. 32rd European Solid-State Research Conference (ESSDERC 2002), September 2003, pp. 151-154.

[13] Torres, R., Klima, R., Selberherr, S., *"Analysis of Split-Drain MAGFETs",* IEEE Electron Devices, Vol. 51, No. 12, 2004, pp. 2237-2245.

[14] Pesic, T., Jankovic, N., *"A Compact Non-Quasi_Static MOSFET Model Based on the Equivalent Non-linear Transmission Line",* IEEE Trans. On Computer-Aided Design of Integrated Circuits and Systems, Vol. 24, No. 10, October 2005, pp. 1550-1561.

Modelling of Printed Circuit Boards in Closed Environment Using TLM Method

Bratislav Milovanović, Nebojša Dončov and Jugoslav Joković

Abstract - In this paper, possibilities and effectiveness of Transmission Line Matrix (TLM) method for modelling of electromagnetic emissions from a printed circuit board (PCB) in closed environment are considered. The method is applied to account for the interactions between the PCB and enclosure by including the basic physical features of the PCB. A basic test board, placed in enclosure, is modelled in configurations where feeding and terminations are realized through TLM wire ports. Also, effects of wire monitoring probe and aperture used in experimental setup are considered. Comparison with reference results based on measurements and Method of Moments (MoM) simulations, confirms the validity of the numerical model.

Keywords – Printed circuit board, wire, enclosure, TLM method.

I. INTRODUCTION

The rapid development and utilization of advanced digital techniques for information processing and transmission in modern communication systems have led to a further evolution of semiconductor technology to nanometre regime. A number of complex components and devices, usually in high-density packaging, can be found in today's communication systems resulting in a very challenging electromagnetic (EM) field environment. Therefore, electromagnetic compatibility (EMC) [1] has become one of the major issues when designing these systems, especially some of their parts such as printed circuit boards (PCBs) and integrated circuits (ICs)

Clock rates that are driving PCBs are now in the GHz frequency range in order to increase dramatically processing speed. Therefore, considering even a few higher harmonics of clock rates takes design of such circuits well into the microwave regime. PCBs are becoming increasingly more complex and as a consequence quantifying their EM presence is more difficult. In the microwave frequency range, PCBs have dimensions of the order of several wavelengths and thus become efficient radiators and receivers of EM energy. In addition to that, high-density packaging, widely applied to PCB design, could cause a significant level of EM interference (EMI) between neighbours PCBs, particularly when they are placed in closed environment. These effects in combination

Bratislav Milovanović, Nebojša Dončov and Jugoslav Joković are with the Department of Telecommunications, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {bata, doncov, jugoslav}@elfak.ni.ac.rs.

with the driving down of device switching voltage levels are making signal quality/integrity and emission/ susceptibility critical EMC issues in next generation high-speed systems.

Differential numerical techniques, such as the finite-difference time-domain (FD-TD) method [2] and the transmission line matrix (TLM) method [3], are common tools for computational analysis of numerous EM and EMC problems. However, a full-wave three dimensional (3D) numerical simulation to accurately reproduce the EM field around a PCB usually requires substantial computing power and simulation run-time. Therefore, one efficient technique based on equivalent principle [4], providing simplified equivalent dipole models to accurately predict the radiated emissions without reference to the exact details of the PCB has been recently proposed in [5]. The model has been deduced from experimental near-field scanning and it includes not only the excitation but also physical features of PCB such as its ground plane and dielectric body, both very important in closed environment. However, such model can be very complex and run-time consuming when incorporating it into conventional calculation algorithms of FD-TD or TLM methods.

For some of the geometrically small but electrically important features (so-called fine features), such as wires, slots and air-vents, few enhancements to the TLM method have been developed [6-8]. These compact models have been implemented either in the form of additional one-dimensional transmission line network running through a tube of regular nodes or in the form of equivalent lumped element circuit, allowing to account for EM presence of fine features without applying a very fine mesh around them. Compared with the conventional approach, these models yield a dramatic improvement in computer resources required.

Similar compact model could be developed for PCB allowing for efficient implementation into TLM algorithm procedure and accurate representation of EM emissions and coupling of PCB. In order to develop such model, an extensive full-wave analysis has to be conducted in order to fully characterize EM presence of PCB either in free space or in closed environment. In this paper, we consider the basic test PCB placed in rectangular enclosure as typical closed environment for PCBs. It consists of L-shaped microstrip track on FR4 substrate [5]. The impact of radiated emission of this simple PCB structure, with wire feed and terminated probes at its ends, on EM field distribution inside the enclosure is investigated. In addition,

a monitoring probe, used in practice to sample the EM field interaction between PCB and enclosure as well as an aperture on one enclosure wall (e.g. used for outgoing or incoming cable penetration from and to PCB) are also taken into account. TLM method, enhanced with the compact wire model is used to carry out this investigation while the simulation results are compared with reference results based on measurements and Method of Moments (MoM) simulations [5].

II. MODELLING PROCEDURE

A. TLM method

In TLM method, a 3D electromagnetic (EM) field distribution in a PCB structure in enclosure is modelled by filling the space with a network of transmission lines and exciting a particular field component in the mesh by voltage source placed on the excitation probe. EM properties of a medium in the substrate and enclosure are modelled by using a network of interconnected nodes. A typical node structure is the symmetrical condensed node (SCN), which is shown in Fig. 1. To operate at a higher time-step, a hybrid symmetrical condensed node (HSCN) [3] is used. An efficient computational algorithm of scattering properties, based on enforcing continuity of the electric and magnetic fields and conservation of charge and magnetic flux, is implemented to speed up the simulation process. For accurate modelling of this problem, a finer mesh within the substrate and cells with arbitrary aspect ratio suitable for modelling of particular geometrical features, such as microstrip track, are applied. External boundaries of arbitrary reflection coefficient of enclosures are modelled in TLM by terminating the link lines at the edge of the problem space with an appropriate load.
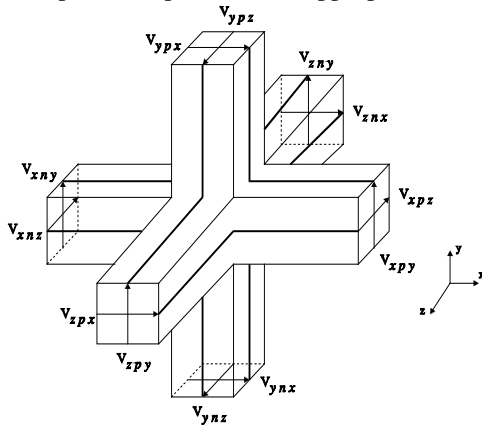
Fig. 1. Symmetrical condensed node

B. Compact wire TLM model

In TLM wire node, wire structures are considered as new elements that increase the capacitance and inductance of the medium in which they are placed. Thus, an appropriate wire network needs to be interposed over the existing TLM network to model the required deficit of electromagnetic parameters of the medium. In order to achieve consistency with the rest of the TLM model, it is most suitable to form wire networks by using TLM link and stub lines (Fig. 2) with characteristic impedances, denoted as Z_{wy} and Z_{wsy}, respectively.

An interface between the wire network and the rest of TLM network must be devised to simulate coupling between the electromagnetic field and the wire.
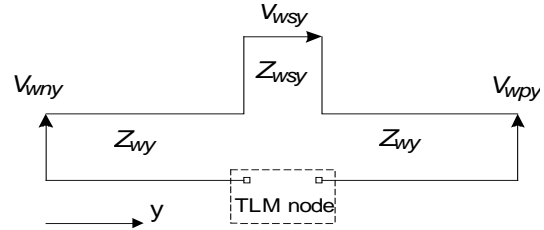
Fig. 2. Wire network

In order to model wire elements, wire network segments pass through the centre of the TLM node. In that case, coupling between the field and wire coincides with the scattering event in the node which makes the scattering matrix calculation, for the nodes containing a segment of wire network, more complex. Because of that, an approach proposed in [6], which solves interfacing between arbitrary complex wire network and arbitrary complex TLM nodes without a modification of the scattering procedure, is applied to the modelling of microstrip structures.

The single column of TLM nodes, through which wire conductor passes, can be used to approximately form the fictitious cylinder which represents capacitance and inductance of wire per unit length. Its effective diameter, different for capacitance and inductance, can be expressed as a product of factors empirically obtained by using known characteristics of TLM network and the mean dimensions of the node cross-section in the direction of wire running [6].

Following the experimental approach that using inner conductor of coaxial guide as a probe, numerical characterisation of EM field inside the cavity can be done by introducing wire ports at the interface between wire probes and enclosure walls.

III. NUMERICAL RESULTS

Results presented here illustrates possibilities of TLM method for determining the emissions from basic PCB structure in form of test board with a microstrip printed on the dielectric substrate [5]. Therefore, for verification purposes, the numerical TLM results from this board are compared with reference results based on MoM, as well as corresponding measurements [5].

The basic test PCB is a 2-mm wide L-shaped microstrip track (l_1=40mm, l_2=20mm) on one side of a $PCB_x \times PCB_y \times PCB_z$=(80×50×1.5)mm^3 board made from FR4 substrate with relative permittivity ε_r =4.5. The geometry of the board is shown in Fig. 3.
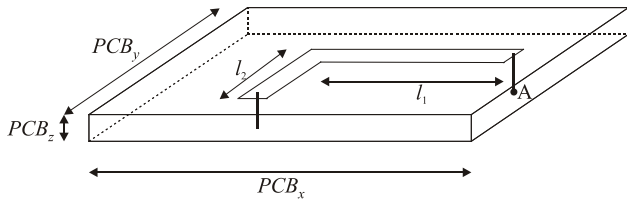
Fig. 3. Geometry of basic test PCB

Resonant frequencies (MHz)	Measured [5]	TLM
PCB in enclosure	900	906
	1290	1284
	1740	1750

The test board PCB is mounted on the bottom of an enclosure in the form on rectangular metallic box with dimensions $a \times b \times c = (284 \times 204 \times 75) mm^3$. PCB is powered by external RF signals via probe (with radius of 0.5mm) placed at one end of microstrip track (point A). This structure allows for an accurate modeling of enclosure through reflection coefficients of boundaries, while feed and terminated probes are modeled through compact wire model applying generator and loads in TLM wire ports at the ends of microstrip track.

When a PCB is inside an enclosure, it is of particular interest to investigate the behavior near the resonant frequencies of the enclosure. Therefore, numerical results of resonant frequencies in modeling closed environment structure are analysed. Fig 4 presents the resonant frequencies obtained from vertical electric field (E_z) sampled at point 35-mm above the PCB.

In Table I values of resonances obtained using TLM simulation are compared with reference values found experimentally by observing the field magnitude inside the enclosure. Comparing with measurements, the simulation results of test boards show that the inclusion of basic features, such as the microstrip track and substrate, in addition to the wires elements for feeding and terminations, permit an accurate prediction of emitted fields to be made in enclosures that have interactions with the PCBs inside. The difference in frequency is less than 10 MHz and may be because the enclosure, used for the measurements, has features not included in the numerical model, such as probe used for EM field monitoring and aperture made on one enclosure wall [5].

Therefore, in Table II values of resonances are compared when, first, the probes used for EM field monitoring, and later the aperture, are incorporated into the TLM model together with the enclosure, to simulate the real closed environment problem (Fig 5). The 30-mm length monitoring probe is placed in vertical direction, mounted on the top wall of enclosure. It is also described by the compact wire model. Aperture with dimensions $a_1 \times b_1 = (60 \times 10) mm^2$ is placed on top wall on enclosure above PCB according to experimental setup.
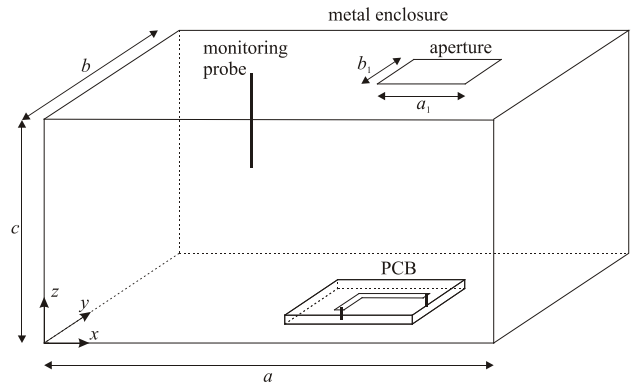
Fig. 5. TLM model of basic test PCB in enclosure with monitoring probe and aperture

Structure	Resonant frequencies (MHz)	Electric field (mV/m)
PCB in enclosure	906	164
	1284	197
	1750	186
PCB in enclosure with probe	897	295
	1280	421
	1738	368
PCB in enclosure with aperture	905	163
	1284	196
	1749	185

Fig. 4. TLM numerical results of vertical component of electrical field in enclosure with basic test PCB

Obtained numerical results illustrate the shift in resonant frequency when additional features are incorporated particularly. It is found, as expected, that

resonances appear at different frequencies if the model includes the monitoring probe. The presence of probe causes difference in frequency about 10 MHz and also leads to a change in the peak field magnitude because it becomes a secondary EM field radiator. On the other hand, impact of aperture is minimal because its dimension is much smaller than the volume of the enclosure so that do not disturb EM field distribution inside enclosure. However, aperture presence could increase the level of EM field radiated outside the enclosure which should be taken into account when emission EMC compliance test of PCB is conducted.

Fig. 6 shows the full patterns of E_z on the plane 35-mm above the bottom of the enclosure given by the TLM simulation at resonant frequencies illustrating EM field distribution of an enclosure due to the physical presence of a PCB and monitoring probe. The TLM results of PCB in enclosure simulation have very good agreement with corresponding results based on MoM simulations [5]. Obtained patterns confirms that modeling the dielectric and track of a PCB as well as wire probes is essential in enclosed environment simulations.

a)   b)
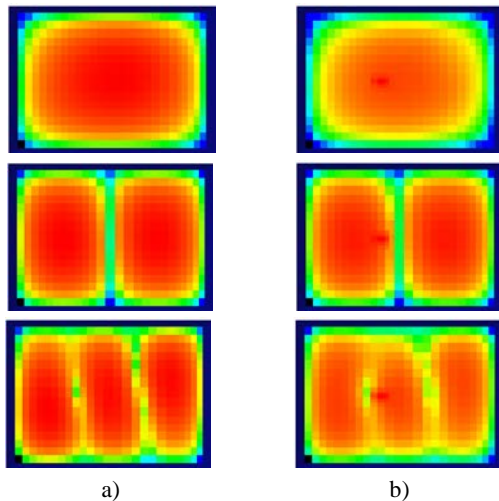
Fig. 6. Patterns of E_z at resonant frequencies, given by the TLM simulation of PCB in enclosure: a) without, b) with monitoring probe

IV. CONCLUSION

Starting that one of the main interests in EMC tests is the intensities and distributions of the radiated fields from equipment under test (EUT), results are presented here of the emissions from basic PCB structure. A method applied to determine radiated emissions from a PCB using a model based on TLM modelling of test board in closed environments, to account for the interactions between the physical presence of the PCB and the enclosure.

According to presented results based on examples of basic test PCB in closed environments, it is found that the TLM method is very convenient for modelling of PCB

structures in form of microstrip track on board made from substrate placed in enclosure. Compact wire TLM model allows modeling of wire conductors used for conections of different layers in PCB, and also probe for monitoring EM field in enclosed space. Good agreement has been achieved between results obtained by using TLM method and those obtained by using the MoM and measurements.

Generally, TLM method could also be appled for more complex multilayer PCBs, but considerations must be given to issues of computational costs, resolution and accuracy. Also, emissions from small elements in PCBs and edges due to enclosure modes will need particularly good characterization. This may require inclusion of additional parameters represented in equivalent models. Nevertheless, here it is demonstrated that the TLM method have the potential to characterize emissions from PCB structures in realistic environments and making it possible to perform system EMC studies.

ACKNOWLEDGEMENT

REFERENCES

[1] Christopoulos, C., "*Principles and Techniques of Electromagnetic Compatibility*", 2nd edition, CRC Press, Boca Raton, FL, 2007.

[2] Kunz, K. S., Luebbers, R. J., "*The Finite Difference Time Domain Method for Electromagnetics*", CRC Press, Boca Raton, FL, 1993.

[3] Christopoulos, C., "*The Transmission-Line Modelling (TLM) Method*", IEEE Press in association with Oxford University Press, Piscataway, NJ, 1995.

[4] Balanis, C. A., "*Antenna Theory - Analysis and Design*", John Wiley and Sons, New Your, 1997.

[5] Tong, X., Thomas, D.W.P., Nothofer, A., Sewell, P., Christopoulos, C., "*Modeling Electromagnetic Emission From Printed Circuit Boards in Closed Environment Using Equivalent Dipoles*", IEEE Transactions on Electromagnetic Compatibility, Vol. 52, No. 2, May 2010, pp. 462-470.

[6] Wlodarczyk, A. J., Trenkic, V., Scaramuzza, R., Christopoulos, C., "*A Fully Integrated Multiconductor Model For TLM*", IEEE Transactions on Microwave Theory and Techniques, Vol. 46, No. 12, December 1998, pp. 2431-2437.

[7] Trenkic, V., Scaramuzza, R., "*Modelling of Arbitrary Slot Structures Using Transmission Line Matrix (TLM) Method*", International Symposium on Electromagnetic Compatibility, Zurich, Switzerland, 2001, pp. 393-396.

[8] Dončov, N., Wlodarczyk, A. J., Scaramuzza, R., Trenkic, V., "*Compact TLM Model of Air-vents*", Electronics Letters, Vol. 38, No. 16, 2002, pp. 887-888.

Advanced DC Motor Drive for Haptic Devices

Miroslav Božić, Darko Todorović, Miloš Petković, Volker Zerbe, Goran S. Đorđević

Abstract - Haptics covers many different forms of mechanical interaction with human senses by engaging, touch, vibrations and forces/torques, established for the purpose of augmenting the feedback structure during human-machine interaction. A haptic device has mechanical part, moved by actuators from one and human hand or fingers from the other, actuators, drives, sensing elements, as well as algorithms designed to control the interaction between human and machine in positioning and motion control tasks. With such a system, motors can be controlled in a way to simulate various environments, defined by their material and dynamics, for example pushing soft ball uphill. Haptic devices are becoming more popular in medical applications after introduction of modern medical robots with many different extensions for minimally invasive surgery or diagnostics based on palpation. This paper discusses one DC motor driver custom designed for the purpose of designing a haptic device for medical applications.

Keywords – Haptics, DC motor drive.

I. INTRODUCTION

Stepping out of industry was a challenging task for ro–botics scientists and engineer. Even now, after more than 25 years after establishing a first medical robot application where PUMA560 robot was placing a needle for brain biopsy using CT guidance, still we have no autonomous robot ready for any medical intervention. In reality, we are witnessing very effective robot installations reaching the level of restrained medical assistants, reliably and passively replicating or augmenting human manual commands presented at the handles, joysticks or specially designed mechatronic interfaces. The most prominent is Intuitive Surgical's da Vinci Surgical System. Simple but effective explanation is that living organisms are so complex in their structure and emerging forms, being healthy tissue or not, it requires another living organism with abundance of sensorimotor skills to perform even the simplest surgical intervention. From engineering point of view, it can be said we can have programmed elemental interaction but we cannot have a complete program that will engage the skills in reliable manner of timing, sequencing or scaling. For that, we still need a surgeon that has all theoretical and practical knowledge while the robot will be only its extension towards better precision in positioning and applying force, less tremor, leading to successful and less invasive surgery. However, the most important achievement of robotic sur-gery is that it might help less trained surgeon to perform a standard surgery in a more reliable way but its overall success is limited within a scope of human skills. There is no surgery or diagnostics that robot can do while human cannot. Even worse for robotics, statistical evaluation of robotic-assisted surgery (RAS) vs. manual surgery shows no obvious benefit to patient's health, as RAS takes longer, it requires skills available only to additionally trained surgeons located at top-notch hospitals, and robotic surgery systems are still not developed enough at the point where human get hands on the robot. The weakest point of today's surgical robot technology is its human-machine interface part as the variety of interactions at that port is huge and so complex no modern technology tools and algorithms can provide its dependability.

The work presented in this paper aims the technology that will enable better HMI based on mechatronic system often called joystick, actuated with high performance DC motors. We developed Advanced Motor Drive for Haptic Devices (AMD–HD) with plenty of interfaces, functions, and processor support that can handle even motor skills of humans based on data after extensive exercising. The drive can be coupled with the other drives towards program-mable bilateral interaction thus leaving a global control effort for the upper level where the strategy of interaction is considered. This paper describes the drive, its structure and purpose, accompanying software for drive programming to meet requirements needed for truly versatile haptic device.

II. AMD–HD DRIVE DESCRIPTION

The AMD–HD motor driver is designed to meet requirements such as precision, dynamics, reliability, connectivity, and scalability of high performance DC motors such as Maxon's RE series motors are. It supersedes generic Maxon drives well known as expensive and not so reliable drives. Among other, more expensive and reliable drives on the market, we did not find any driver that will naturally suit the needs in haptic devices for fast model-based control between the drive interactions due to the need of mechanical cross coupling required for achieving desired mechanical impedance projected at human hand in the whole workspace. The AMD–HD is based on micro-controller system that handles all: velocity and torque estimation, speed and torque control, position control, and even model-based control. Dedicated PC software handles GUI. The Driver and GUI software communicate via RS232/485 serial bus thus enabling multiple drive control within the same supervisory application at PC side.

Miroslav Božić, Darko Todorović, Miloš Petković, Goran S. Đorđević are with Department of Control Engineering, Faculty of Electronic Engineering, University of Niš, Serbia. goran.s.djordjevic@elfak.ni.ac.rs
Volker Zerbe is with University of Applied Sciences, Erfurt, Germany. volker.zerbe@fh-erfurt.de

A. Drive properties

The Drive nominal operating voltage is between 20 VDC and 90 VDC. Maximum output voltage is 72 VDC. Maximum output current 15 A (for less than 30 sec). Continuous output current is 10 A. Pulse Width Modulation frequency is 40 kHz. While the sampling rate is programmable and can be as low as 10kHz. Maximum motor speed is limited by maximum permissible speed (motor) and max. output voltage (controller). The dimensions of the drive are WxLxH: 148x148x40mm. Total weight with cooler is approx. 150grams.

Photo of the AMD–HD driver is shown on Fig. 1, with main modules numbered as:
1. Power Supply Unit
2. Control Unit
3. Motion Feedback Module
4. Analog-Digital IO module
5. Communication module
6. Motor current sensing module
7. H-bridge

Fig. 1. AMD–HD DC motor drive

The Drive has dedicated Power Supply Unit, with standard +5V, +12V, +V DC voltages for H-bridge. Noise is redu-ced by DC-DC converter with +V at input. Additional stabilization of voltage at the converter output supplies control circuitry. Microchip's PIC18F4431, besides its standard peripherals, has 4 independent complementary 14bit PWM modules; motion feedback module for data logging from quadrature signals of incremental optical encoder, 200 ksps 10bit AD conver–ter; all making it almost perfect for the task we are target–ing. Both, TTL and Differential inputs from encoders are handled by using 26LS32 line driver and 74HC157 Quad 2-Input Multi–plexer as data selector. Encoder selection is possible within GUI software. Additional analog input 0 to 5VDC is available for position or velocity commands. Finally, two digital open collector inputs DI1 and DI2 are available for drive configuration from the GUI. The Drive communi-cates with GUI at PC level via RS232 point to point and RS485 multi point serial communication.

B. Current sensing

Current measurement is one of the most important properties of the Drive, as required for mechatronic technologies such as haptics. Precise current sensing can be related to the interaction force exerted between mechanics (linked to the motor) and human hand. For that to happen, it is crucial that mechanics efficiency of all transmissions (primarily, from gearbox, and secondary) is as low as possible and considerable amount of human machine interaction is feed back to the motor shaft for further estimation of current. The nature of such interaction points out sensitivity over bandwidth as hand movements are more soft and slow than strong and fast. Having that in mind we choose LEM HXS10-NP current sensing element. Sensor output is fed into non-inverting CMOS Op-Amp MCP601. The gain is set so that 1VDC corresponds to motor current of 1A. By setting up resistance the maximum gain is 2.5VDC/10A. The amplified signal is brought to analog input of micro–controller. The digital signal of current sensed is sent to GUI at PC.

Having in mind that output voltage of LEM sensor is:

$$V_{LEM} = V_{REF} \pm \frac{0.625 \cdot l_M}{l_{MAX}},$$

and the gain of amplifier is:

$$G = 1 + \frac{R_3}{R_2},$$

then, the voltage at the microcontroller input is

$$V_{MCU} = V_{REF} \pm \frac{0.625 \cdot l_M}{l_{MAX}} \cdot G.$$

C. GUI software

Dedicated GUI software, MCA-1 Monitor, handles all drive settings, PC peripheral configuration, data recording and visualization. Its main widow is shown in Fig. 2.

Fig. 2. Main window of MCA-1 monitor GUI

This application handles:

1. COM port selection on PC where the driver is connected,
2. Address selection to be used by driver, for read and write of new parameters,
3. Free address assignment to non-configured driver
4. Inspection of Drive parameters such as: type, driver supply voltage, motor nominal voltage, maximal pulses of incremental optical encoder, encoder output signal type, motor brake operating voltage, gearbox ratio, digital input function selection.
5. PWM duty cycle selection along with direction of rotation.
6. Motor parameter inspection
7. Current sensor voltage
8. Motor current sensing
9. Communication log.

This application also enables current data logging and grahipical visualization in time-based diagrams, as shown in Fig. 3.
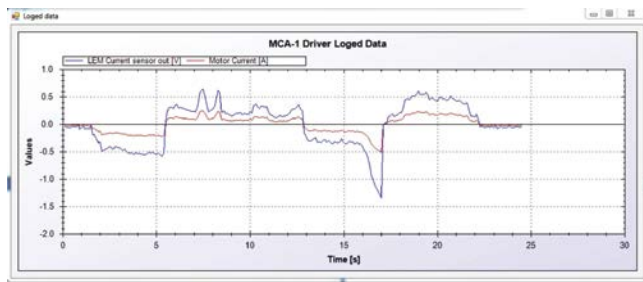
Fig. 3. Recorded voltage signal (at LEM sensor output) and calculated motor current

Arbitrary interaction with gearbox shaft with fingers produces the torques sensed and recorded by the driver. From Fig. 4, it is obvious that calculated motor current, shown with blue line, compared to LEM current sensor output, shown with red line, has very similar dynamics, almost without phase shift, and gain increase of approximately 82. The phase shift and filtering properties of the two signals should be correlated for further elaboration.

D. Communication protocol

With rapid development of embedded systems, and significant price and development time reduction, distributed data gathering and local processing, based on reusable modules is becoming the mainstream model for rapid prototyping and final product development. In order to fulfill all the needs in matters of bandwidth, reliability and information security, appropriate protocol for communication between DC motor driver and PC, as well as between drivers themselves was defined, and communication libraries for .Net and microcontroller were developed.

The AMD–HD driver communicates with PC GUI software via RS232 PTP or RS485 MP connection. Jumpers on the board are used for type selection. Also, two drives can be attached via I2C bus, but only the Master of the two drives can be controlled while the Slave will follow the Master. Message exchange protocol is explained with the timing diagram given in Fig. 4.

At first, we defined the time diagrams that describe the flow of messages and rules of communication between devices on global level shown on Fig. 4. Here we can see two timelines; first timeline is associated with AMD-HD (DC motor driver) and second is associated with PC which acts like a master device in communication. Because we use RS485 serial communication protocol on physical layer, we must have at least one master node at a time. As you can see on figure xx, the CRC is added at the end of each message and checked on each side. If CRC is not equal with calculated one, then the message retransmission is requested from both sides. In some modes, like PC monitoring, retransmission is not requested because the new driver state will be sent in next message again. The message with error will not be considered in that case. If 10 messages in a row come with errors then there is a communication failure and communication is stopped, to be sure not to make some damage.
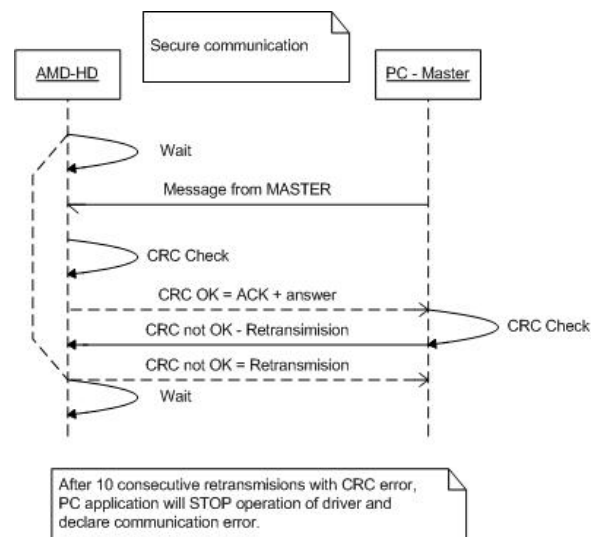
Fig. 4. Secure communication timing diagram

Every message has defined structure and meaning. The structure of message is determined by the MUN (Message Unique Number) field. Based on this field, recipient can determine type of message, length and order of useful bytes. On Fig. 5 the structure of each message is shown.

Byte 0	Byte 1	Byte 2	Byte 2+n	Byte 2+n+1	Byte 2+n+2	Byte 2+n+3
ADR_REC	ADR_SEND	MUN	DATA_BYTES	NBR	#	CRC

Fig. 5. Message format

First and second byte define addresses of receiver and sender respectively. Third byte is already mentioned MUN byte, after which useful, data containing bytes are defined represented with DATA_BYTES. After useful data bytes, message length byte is placed and used for decoding purposes, after which termination byte represented with "#" character. At the end of the message calculated CRC is placed. Decoding on receivers side starts with finding the termination character and comparison between counted and the length of message in NBR byte. At the end CRC is checked and message is declared good or with errors.

E. Simulation

Modeling of H-bridge motor driver is considered trivial as extensive research was published in the last decades such as [1], and [2]. Modeling of Maxon DC Motors in Simulink is already done as well, and one of better models is found in [3] as a freeware. We have used that motor model but with catalogue data for Maxon motor RE40 [4, 5] as parameters. From practical point of view these simulations are not needed in haptic devices. Instead, for safety reasons we do need state transition diagrams of the drive for safe interaction with human operator. This should also include the safety system tailored for such purposes.

III. EXPERIMENTAL SETUP

This study and development has been undertaken for projects related to medical applications where human-machine interface is needed through controllable bilateral mechanical interaction. We chose high performance Maxon DC motor RE40, code 263075, operating on 48VDC, with 987 rpm no load speed, maximum continuous torque of 0.19Nm and torque constant of 461mNm/A, [4]. Such a motor is recognized among the others of the same family as the best transformer of output torque into motor current. Its low speed also makes it very suitable to be used in haptic interfaces. Even alone motor has enough torque to produce sensible force/torque on a standard mechanical device such as joystick. Here we added a high efficient planetary gearbox from the same manufacturerwith 4.3:1 gear ratio and 9.4gcm^2 mass inertia [5]. Such a gearbox is very compliant backwards, meaning that motor can sense the torque applied on the gearbox shaft without considerable loss of information. The motor is also equipped with standard incremental optical encoder HEDS series with 500ppr. AMD-HD motor drive along with gear-motor-sensor application is shown in Fig. 6.
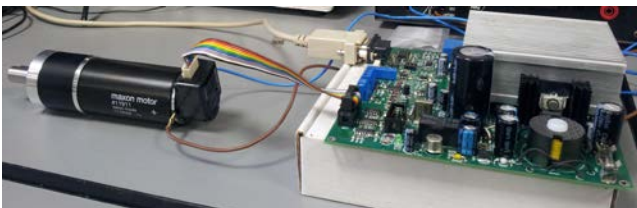
Fig. 6. Drive (to the left) and geared motor RE40.

IV. CONCLUSION

Advanced drive for DC motors is designed with focus on driving haptic devices with such system. It has several advantages as compared to the drives met on the market. First, it has high sensitive current sensor LEM HXS10-NP included meaning we can measure the iteraction with environment (even through a gearbox of modest ratio) by measuring the current. Also, it can operate simultaneously with other drives, exchanging data on current, velocity, position and actual state of the motor. This makes it very useful in integrating it into a hierarchical control system where upper level of control is done at embedded controller while the motion itself along with interaction is left to be handled at the driver side. This is particulary useful in haptic devices where the interaction is complex dynamics and it requires special calculations. Finally, on-board processor can be used for customized model-based control (analytic or data-based) thus making this drive a perfect companion of the motor in the given mechatronic system.

REFERENCES

[1] V. Gupta. "Working and analysis of the H - bridge motor driver circuit designed for wheeled mobile robots." *in Proc. Advanced Computer Control (ICACC), 2010 2nd International Conference*, pp. 441 – 444, ISBN: 978-1-4244-5845-5, 27-29 March 2010

[2] Wai Phyo Aung. "Analysis on Modeling and Simulink of DC Motor and its Driving System Used for Wheeled Mobile Robot" *World Academy of Science, Engineering and Technology*, ISSN 2010-37632, 2007

[3] S. Kozola and D. Doherty. (2007, May). "Using Statistics to Analyze Uncertainty in System Models", *MATLAB Digest*,

[4] Maxon. "Maxon DC motor RE40 catalog page", Internet: http://test.maxonmotor.com/docsx/Download/catalog_2005/Pdf/05_083_e.pdf [Oct. 12, 2011].

[5] Maxon. "Maxon Planetary Gearhead GP 42 C catalog page", Internet: http://www.electromate.com/db_support/downloads/244245.pdf [Oct. 12, 2011].

Energy efficiency and fault tolerance analysis of hard real-time systems

Sandra Đošić and Milun Jevtić

Abstract - In this paper the tradeoff between dynamic voltage and frequency scaling (DVFS) techniques and faults tolerance are considered. We analyzed this tradeoff using one heuristic-based DVFS algorithm which we designed. The proposed algorithm minimizes the energy consumption of one real-time task set when the transient faults are exceeded. It is assumed that the tasks execute on processors with variable frequency and voltage levels. The simulation results show that our proposed algorithm can be used for real-time systems analysis from the perspective of finding compromise between energy efficiency and fault tolerance.

Keywords - Dynamic voltage and frequency scaling, Fault tolerance, Real-time systems.

I. INTRODUCTION

Hard real-time systems (HRTS) play an important role in many areas of daily life: robotics, cosmic research, automotive industry, process control, factory automation… Those systems have been designed in order to be safe and extremely reliable. They are usually realized as real time systems with the ability of tolerating some faults. A fault-tolerant HRTS has to ensure that faults in the system do not lead to a failure.

Dependability of one real-time system can be affected by different kinds of faults, including transient, permanent and intermittent faults. Among these, the transient faults are much more common than faults of other two types. Since transient faults have the feature that they occur and then disappear, fault tolerance can be achieved, running the task affected by a transient fault again (i.e. re-executing the task). It means that time redundancy can be used as fault-tolerance techniques by using free slack time in the system schedule to perform recovery executions, [1].

Beside high level of dependability, energy efficiency is crucial to many real-time systems due to their limited energy supply and severe thermal constraints of the operating environment. Dynamic Voltage and Frequency Scaling (DVFS) is the most popular and widely deployed technique for reducing power and energy consumption of processors [2], [3], [4], [5]. Nowadays, DVFS is a commonly used technique for energy management and is supported by many commercial processors [6].

Fault tolerances through time redundancy as well as energy management through frequency and voltage scaling

Sandra Đošić and Milun Jevtić are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {sandra.djosic, milun.jevtic}@elfak.ni.ac.rs.

have been well studied in the context of real-time systems.

For HRTS that require both fault tolerance and energy efficiency, there is a lack of efficient solutions. Simply applying fault recovery techniques and energy minimization techniques one after the other only results in inferior quality. This is because minimizing energy first may not leave enough slack for fault recovery, and minimizing energy after fault recovery reservation treats normal task executions and re-executions (for fault recovery) equally, which is equivalent to optimizing the worst case that happens rarely. Since, free slack time is a limited resources, it is obvious that more slack time for DVFS technique means less time for fault tolerance, and vice versa. Therefore, there is a tradeoff between low energy consumption and high fault-tolerance.

The tradeoff between DVFS techniques and faults tolerance is focus of this paper. In accordance with that, we designed one heuristic-based DVFS algorithm and used it for energy efficiency and fault tolerance analysis of HRTS.

The rest of the paper is organized as follows. Section II describes real-time system, power, fault and feasibility models. Next section III introduces our proposed heuristic-based DVFS algorithms. Section IV gives the simulation results and finally, Section V presents our conclusions.

II. MODELS DESCRIPTION

A. System model

We consider one uniprocessor real-time system with variable CPU frequency f_j ($j=1,...,m$) where $f_j < f_{j+1}$. The voltage and the operating frequency of the CPU may be switched between m values. This system can be used for one real-time task set execution. We assume a set of n real-time tasks, $\Gamma=\{\tau_1,..., \tau_n\}$ where each tasks are defined by a minimum inter-arrival time T_i, worst case execution time (WCET) C_i and deadline D_i. We assume that $D_i \leq T_i$, for $i = 1, 2, ..., n$. The WCET of real-time tasks corresponds to executing the task at the maximum frequency f_m. For simplicity, we assume that the WCET of a task scales linearly with the processing speed. So, if we scale the operating frequency by a factor α, then WCET must be scaling by factor $1/\alpha$, i.e.

$$C_i (f_j)= C_i (f_m) f_m / f_j.$$

Each task is assigned a unique priority p_i and all of them are periodic, fully preemptive and independent. Algorithm for scheduling real-time tasks could be any priority assignment algorithm, [7].

B. Power model

Power consumption of an active processor can be modeled as

$P_A(f) = P_d(f) + P_{ind}$,

where $P_d(f)$ and P_{ind} are frequency dependent power and frequency independent power respectively [8]. Frequency dependent power is

$P_d(f) = V^2(f) C_{ef} f$

where V is supply voltage and it is a function of operating frequency, C_{ef} is the switch capacitance and f is the frequency. Beside power, for DVFS techniques energy is equally important and it is defined as the integral of power over time.
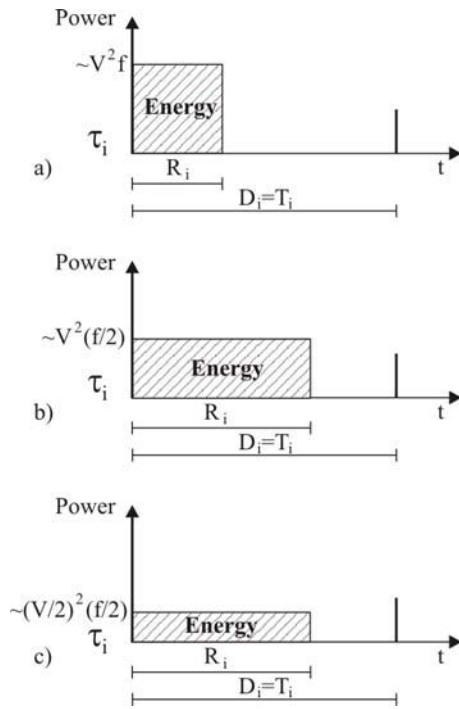
Fig. 1 Power and energy consumption for real-time task τ_i:
a) for frequency f and voltage V
b) for frequency $f/2$ and voltage V
c) for frequency $f/2$ and voltage $V/2$

Fig. 1 a) illustrates energy consumption of one real-time task τ_i on operating frequency f and for supply voltage V. Energy of real-time task τ_i is proportional to the marked rectangle area. Fig. 1 b) presents the situation when the operating frequency f is reduced by half and because of that task needs more time to execute. In that situation processor's power consumption is lower but the energy

consumption remains the same. Fig. 1 c) shows the influence on power and energy consumption when supply voltage V is reduces by half. Lowering the supply voltage V the significant amount of energy could be reduced, because of the quadratic relation between power and V. The maximum energy reduction is obtained by lowering the supply voltage and operating frequency simultaneous.

C. Fault model

We assume that faults can occur during execution of any task. We consider transient faults and assume that the consequences of a fault can be eliminated by simple re-execution of the affected task at its original priority level and at its original CPU frequency. The re-execution of the corrupted task must not violate timing constraints of any task in Γ.

D. Feasibility model

In our approach we use the response time analysis (RTA) to check the feasibility of fault tolerant real-time task set. In the RTA, the fault-tolerance capability of a RTS is represented by a single parameter, T_F, which corresponds to minimum time interval between two consecutive faults that the RTS can tolerate. More about RTA can be found in [9], [10].

The basic equation characterize for RTA is Eq. (1).

$$R_i^{n+1} = C_i + \sum_{j \in hp(i)} \left\lceil \frac{R_i^n}{T_j} \right\rceil C_j + \left\lceil \frac{R_i^n}{T_F} \right\rceil \max_{j \in hp(i) \cup i} (C_j) \quad (1)$$

With Eq. (1) the response time R_i of a task τ_i could be calculated. Eq. (1) has three main addends. The first is WCET C_i for a task τ_i. The second presents interference due to preemption by higher priority tasks. We use hp(i) to denote the set of tasks with higher priorities than i, hp(i)={$\tau_j \in \Gamma | p_j > p_i$}. The third addend refers to possible faults in the system. If we assume that inter-arrival time between faults is T_F then there can be at most $\left\lceil \dfrac{R_i}{T_F} \right\rceil$ faults during the response time R_i of task τ_i. Since these faults could occur during the execution of task τ_i or any higher priority task which has preempted τ_i, each fault may add $\max\limits_{j \in hp(i) \cup i} (C_j)$ to the response time of task τ_i. So, the third addend in Eq. (1) presents an extra time needed tasks recovery due to faults.

Since R_i appears on both sides Eq. (1) is recurrence relations which starts with $R_i^0 = C_i$. The solution is found when $R_i^{n+1} = R_i^n$. If during the iteration process we get that $R_i^{n+1} > D_i$ then task τ_i is infeasible and iteration process must be terminated.

Fig. 2 illustrates RTA applied on one simple RTS with two periodic real-time tasks τ_i and τ_j. Two faults occur in the system and the time between these two consecutive

a)

b)
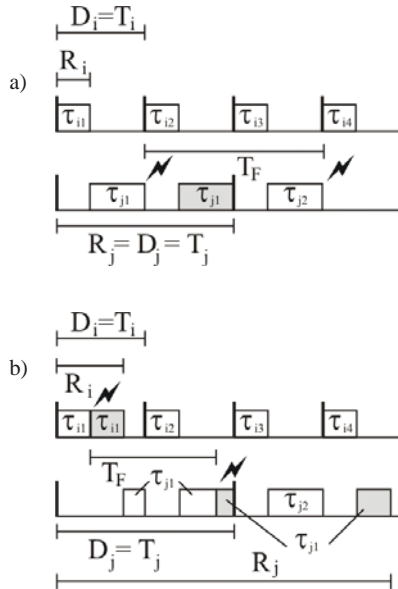
Fig. 2 a) T_F is long enough and RTS is fault tolerant
b) T_F is not long enough that RTS stays fault tolerant

faults is T_F.

Fig. 2 a) presents the situation when first fault occurs just a little bit before the end of tasks τ_{j1}. System overcomes this fault by executing task τ_{j1} again. This is situation when T_F is long enough and real-time system can overcome these faults. System of these two tasks are schedulable i.e. both tasks execute before their deadlines, D_i and D_j. Response time of tasks τ_i and τ_j are the output results of RTA and they are also shown on Fig. 2 a).

Fig. 2 b) presents scheduling of the same real-time tasks τ_i and τ_j when two faults occur in the system. Now, time between two consecutive faults T_F is not long enough and real-time system cannot tolerate these faults. First fault occurs just a little bit before the end of tasks τ_{i1} execution. Real-time system can overcomes this fault by executing task τ_{i1} again. Second fault occurs just a little bit before the end of tasks τ_{j1} execution. Now time redundancy is not enough to tolerate this fault. Systems starts procedure for overcoming fault by executing task τ_{j1} again but timing characteristics of tasks τ_{j1} cannot be satisfied and τ_{j1} missing its deadline i.e. $R_j > D_j$. This is not acceptable in

one hard real-time system, so in this case real-time system is not fault tolerant.

III. PROPOSED DVFS ALGORITHM

The focus of this section is to explain our proposed DVFS policy which fulfils energy efficiency and faults tolerance requirements. For this purpose we created a heuristic-based algorithm to find appropriate execution frequency for each task, from the real-time tasks set, that minimize energy consumption when faults are absent. The RTA is the basic of our proposed algorithm. This analysis is used to guarantee feasibility of real-time tasks set and fault tolerance.

Fig. 3 shows the pseudo code of our proposed algorithm. The input parameters for the algorithm are:
- CPU frequency f_j ($j=1,..., m$) where $f_j < f_{j+1}$ and m is number of frequency levels;
- characteristics for all n real-time tasks from the set: inter-arrival time T_i, worst case execution time C_i, priority p_i and deadline D_i, for $i=1,..., n$;
- minimum time interval between two consecutive faults T_F.

Input: CPU frequency levels f_j ($j=1..m$),
      characteristics for n real time tasks (C_i, D_i, T_i, p_i),
      fault tolerant constraint (T_F)

_____

(1) for each *Task* in *TaskSet* set *Task's_Freq* to f_m and set *Task's_Key* to true;
(2) repeat step (3) to (7) until there are true *Task's_Key* in the *TaskSet*;
(3) for each unlock *Task* in *TaskSet* do
(4)   temporarily set *Task's_Freq* to *Lower_Task's_Freq*;
(5)   *if* new *TaskSet* is not feasible
(6)   then set *Task's_Key* to false;
(7)   else calculate *ΔPower* as Power(*Task's_Freq*) – Power(*Lower_Task's_Freq*);
(8) find *Task* with maximum *ΔPower* and set *Task's_Freq* to *Lower_Task's_Freq*;

_____

Output: *TaskSet* with new frequency assigne to each *Task*

Fig. 3 Heuristic algorithm solution

The algorithm starts with assigning the maximum execution frequency, f_m, to each real-time task, step (1). Also, at the beginning, all tasks are allowed to change the frequency - we say that all tasks are unlocked. An iteration of the algorithm decreases the frequency of one task for one frequency level. The chosen task is one for which the frequency decrement yields maximum power reduction among all unlocked tasks provided that tasks set remains feasible. To find such task, the algorithm checks all

currently unlocked task. For example, frequency index of one unlock task τ_i is temporarily decreased for one frequency level, i.e. from f_j to f_{j-1}, step (4), and feasibility of task-set is tested using Eq. (1), step (5). If task-set is not feasible, τ_i is locked, step (6). Otherwise, if task-set is feasible, the difference between power consumption of τ_i at lower (f_{j-1}) and higher (f_j) frequency is calculated, step (7), as:

$$\Delta P_i = (C_i(f_j) - C_i(f_{j-1}))/T_i.$$

Then, τ_i's frequency is changed back to f_i. After checking all tasks, one that remains unlocked and provides the maximal power reduction is selected, and its frequency index is decremented, step (8). Additionally, the selected task is locked if its new frequency equals 1, i.e. corresponds to the lowest execution frequency, f_1. After that, the algorithm enters the next iteration. The algorithm finishes when there are no more unlocked tasks. The frequency assignment to each task is algorithm's output.

IV. SIMULATION RESULTS

The simulator we realized is based on our proposed heuristic-based algorithm. The input parameters of the simulator are real-time task set characteristics, processor's voltage and frequency levels and fault constraints. On the bases of proposed algorithm, simulator has to find the appropriate execution frequencies for each real-time task that lead to the maximum energy savings for the given fault tolerance constraints. Power consumption of task set are simulator's output result.

TABLE I
TASKS SET FROM GENERIC AVIONICS PLATFORM

τ_i	p_i	C_i (ms)	$T_i = D_i$ (ms)
Nav_Status	1	1000	1
BET_E_Status_Update	2	1000	1
Display_Stat_Update	3	200	3
Display_Keyset	4	200	1
Display_Stores_Update	5	200	1
Nav_Steering_Cmds	6	200	3
Tracking_Target_Upd	7	100	5
Display_Hook_Update	8	80	2

We perform simulations with a number of synthesized real-time tasks sets and few real-world applications. The characteristics of on of the real-world application are summarized in Table I. It is a task set taken from the Generic Avionics Platform (GAP) used in [11].

For the CPU frequency levels we used data from [6] based on the published data of Intel Xscale PXA270. The data sheet of this processor is available online at its manufactures' websites. We used specifications listed in Table II.

TABLE II
FREQUENCY AND VOLTAGE LEVELS OF INTEL XSCALE PXA270

Frequency (MHz)	Voltage (V)	Active power consumption (mW)
624	1.55	925
520	1.45	747
416	1.35	570
312	1.25	390
208	1.15	279
104	0.9	116
13	0.85	44.2

Fig. 4 shows the simulation results for GAP task set and Intel Xscale PXA270 processor. The x-axis represents the ratio of T_{Fmin} to T_F. T_{Fmin} is minimum time interval between two consecutive faults that the task set can tolerate on maximal executing frequency and T_F is input simulation parameter. This axis represents the normalized T_F value which is proportional to fault tolerance of the task set. The y-axis represents the power reduction calculated in percents. This reduction is presented as power saving with respect to the power consumption at maximum frequency.

The simulation was done for three possible scenarios connected with processor. In the first case we used all 7 voltage levels, in the second 4 (0.85V, 1.15V, 1.35V, 1.55V) and in the third just 2 voltage levels (1.15V, 1.55V).

All three scenarios indicate the same fact that power reduction leads to less fault tolerance and vice versa. It can be concluded that power reduction is better when more voltage levels are included. Now, due to simulation results, we can better perceive the tradeoff between power consumptions and fault tolerance. For example, let's suppose that power reduction demands for the given task set are between 16% and 20%. It can be seen, from the Fig. 4, that 7 voltage levels processor fulfill the power reduction demands. Also, fault tolerances vary for the given power reduction interval, so the best is to choose one with maximal tolerances.

Realized simulator offers the possibility to analyze one real-time task set when the main question is to find compromise between power or energy consumption and fault tolerance constraints. Our opinion is that this simulator could be successfully used in the RTS design proces
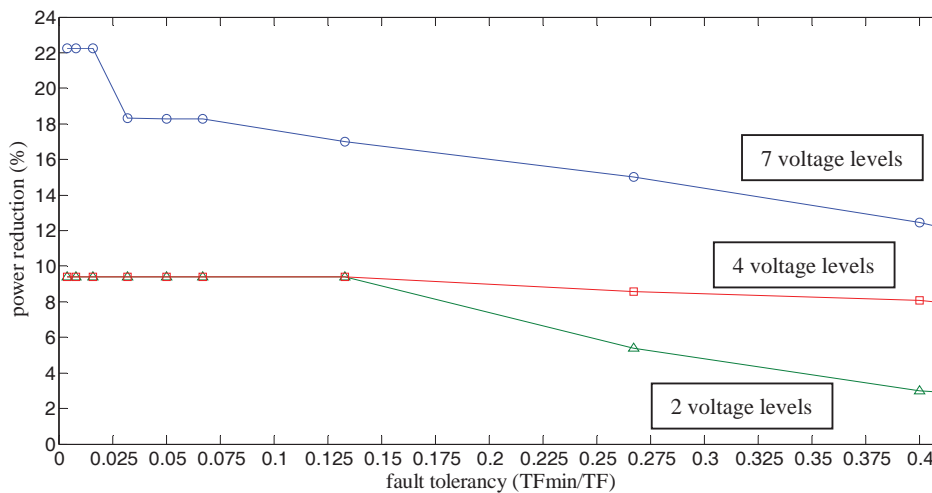
Fig. 4 The simulation results

V. CONCLUSION

This paper studies the trade-off between energy efficiency and fault tolerance for real-time task sets. Recognizing that the problem in discrete systems is NP-hard, we proposed heuristic-based approach which minimizes energy of task set for the given fault tolerant constraints. Our approach is realized for HRTS analysis when is necessary to examine connection between energy consumption and fault tolerance through time redundancy.

We considered only dynamic power in this paper. It should be noted that during the past decades, transistor sizes entered deep submicron regimes where static power consumption is now a non negligible source of power dissipation even in running mode. Thus, the total power consumption (i.e. dynamic plus static power) has to be optimized instead of simply reducing dynamic power.

ACKNOWLEDGEMENT

REFERENCES

[1] Đošić, s., Jevtić, M., "*Scheduling in RTS Using Time Redundancy for System Recovery After Faults*", Proceedings of papers, Indel 2004, Banja Luka, pp. 146-149, November 2004.

[2] Woonseok, K., Dongkun, S., Han-Saem, Y., Jihong, K., Sang, M. L., "*Performance Comparison of Dynamic Voltage Scaling Algorithms for Hard Real-Time Systems*", Proceedings of the Eighth IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS'02)*, pp. 219 – 228, 2002.

[3] Ahmadian, A. S., Hosseingholi, M., Ejlali, A. "*A Control-Theoretic Energy Management for Fault-Tolerant Hard Real-Time Systems*", 2010 IEEE International Conference on Computer Design, pp. 173-178, 2010.

[4] Zhu, P., Yang, F., Tu, G., Luo, W.,"*Fault-Tolerant Scheduling for Periodic Tasks based on DVFS*", Proceedings of the 9th International Conference for Young Computer Scientists, pp. 2186 – 2191, 2008.

[5] Santos, R. M., Santos, J., Orozco, J. D., "*Power saving and fault-tolerance in real-time critical embedded system*", Journal of system Architecture 55, pp. 90-101, 2009.

[6] "*Intel PXA270 Processor Electrical, Mechanical and Thermal Specification Data sheet*", www.phytec.com/pdf/datasheets/PXA270_DS.pdf, 2005.

[7] Cottet, F., Delacroix, J., Mammeri, Z., "*Scheduling in Real-Time Systems*", John Wiley & Sons, 2002.

[8] Dakai, Z., Melhem, R., Mosse, D., "*The Effects of Energy Management on Reliability in Real-Time Embedded Systems*", Proceedings of the 2004 IEEE/ACM International conference on Computer-aided design, pp. 35-40, 2004.

[9] Đošić, S., Jevtić, M., "*Analysis of Real-Time Systems Timing Constrains*", SSSS2010, 3rd Small Systems Simulation Symposium 2010, February, 12-14, Faculty of Electronic Engineering, Niš, Serbia, pp 56-60, 2010.

[10] Lima, G., Burns, A., "*An Optimal Fixed-Priority Assignment Algorithm for Supporting Fault-Tolerant Hard Real-Time Systems*", IEEE Transaction on Computers, Vol. 52, No. 10, pp. 1332-1346, October 2003.

[11] Locke, C. D., Vogel, D. R., Mesler, T. J., "*Building a Predictable Avionics Platform in Ada: A Case Study*", Proceedings of IEEE Real-Time Systems Symposium, pp. 181–189, 1991.

Computer Model for Analysis and Re-design of Crystal Filters

Milorad Paskaš, Miroslav Lutovac, Dragi Dujković, Irini Reljin and Branimir Reljin

Abstract – Quartz crystals can be modelled by LC equivalent sections. By using such presentation crystal filters of desired order can be presented as equivalent LC ladder networks. In this paper a method for automated generation of schematic model of crystal filter as a ladder LC structure is described. From schematic model the analysis of filter transfer function can be derived and filter re-design will be possible in order to satisfy desired filter specifications.

Keywords – Crystal filters, LC filter, Quartz filter, Transfer function.

I. INTRODUCTION

DEVELOPMENT of technological techniques for production of microelectronical components with monocrystal quartz units is actual topic. Crystal quartz units used as resonators in crystal oscillators and filters are implemented in varous areas in telecommunications such as digital networks, cellular networks and protected/encrypted radio communication networks for military and civilian usage. Since these networks have large spatial coverage and millions of clients it is important to provide high level of synchronization of the system, stability of carry frequencies and appropriate selectivity. All these requirements are achieved with quartz based components.

In this paper we present the first results on computer based analysis of LC lossless filters as the idealized model of crystal filters. Manual calculation of transfer function of higher order LC filters is usually higly demanding and error-prone task. Furthermore computer model of filter enables analysis of impact of parameter variation on transfer function of the filter.

The paper is organized as follows. Chapter two describes equivalent circuit of quartz and crystal filters. Chapters three and four give brief description of Matlab functions for drawing scheme and calculation of transfer

Milorad Paskaš is with the Innovation Center of Faculty of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia, e-mail: milorad.paskas@gmail.com

Miroslav Lutovac is with Singidunum University, 11000 Belgrade, Serbia, e-mail: lutovac@etf.rs

Dragi Dujković, Irini Reljin and Branimir Reljin are with Faculty of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11000 Belgrade, Serbia, e-mail: irinitms@gmail.com, {dragi,breljin}@etf.rs
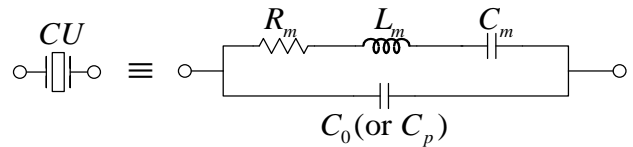
Fig. 1. Equivalent circuit of the crystal unit

function of the LC filter. Applications of these functions are explained as a few examples in chapter five. Finally, chapter six brings conclusions and future work and improvements to be further made.

II. CRYSTAL FILTERS

Crystal unit (CU) is the mechanical vibrating element characterized by *piezoelectricity*: the ability to generate an electric potential in response to applied mechanical distortion, and vice versa. As other mechanical vibrating elements, CU is characterized by its natural frequencies and can be presented by its equivalent electrical scheme as shown in Fig. 1. Series components, R_m, C_m, L_m, are determined by the vibration of the crystal material itself (internal parameters), while the capacitance C_0 models the electrodes on the crystal plate and the stray capacitances from the crystal enclosure (external parameter). Typical plot of reactance, X, vs. frequency, for the crystal unit, is shown in Fig. 2.
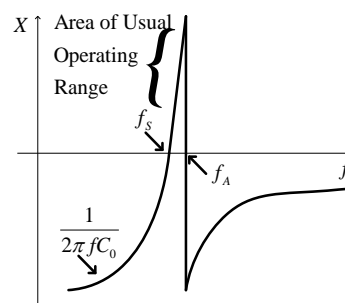
Fig. 2. Reactance X versus frequency for a crystal unit.

The serial (resonance) frequency, f_S , is determined by L_m and C_m, and is given as:

$$f_S = \frac{1}{2\pi\sqrt{L_m \cdot C_m}}$$

(1)

while the parallel (antiresonance) frequency, f_A, is:

$$f_A = \frac{1}{2\pi} \sqrt{\frac{C_m + C_0}{L_m C_m C_0}} \cong f_s (1 + C_m / 2C_0) \Big|_{C_0 \gg C_m} \quad (2)$$

Due to extremely high Q-factor, the parallel resonance frequency is only few kilohertz higher than the serial one. Crystals below 30 MHz generally operate between these two frequencies, where high sensitivity, $\Delta X/\Delta f$, is observed. Since in this range inductive nature prevails ($X>0$), any external capacitance will diminish the frequency.
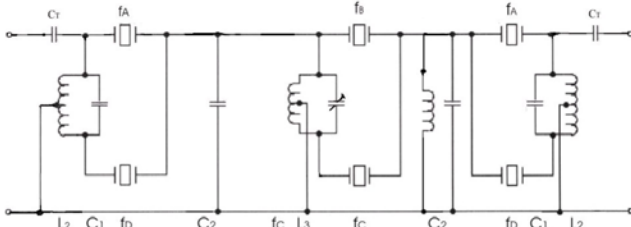
Fig. 3. Electrical scheme of the 50 MHz crystal filter

Manufacturers normally cut and trim crystals to obtain unloaded resonance frequency higher than specified, and by external parallel capacitance the frequency is adjusted to desired value. Crystals above 30 MHz (up to >200 MHz) generally operate at serial resonance, where the impedance is equal to the serial resistance, R_m, which is small (<100 Ω) [1].

Typical crystal filter is shown at Fig. 3. Based on the specifications of the amplitude characteristics in stop and pass band, filter order and corresponding electrical network are determined. The filter is design is based on assumption of real (non ideal) filter elements. Filter design defines the requirements of the crystal specimen. Crystal filters are designed through various nonlinear and graphical methods and sometimes there is no simple method for testing its parameters and amplitude characteristic [2]. By using such equivalent scheme and neglecting resistive loss, crystal filters of desired order are presented with equivalent LC ladder networks.

III. SCHEMATIC REPRESENTATION

The code for schematic representation of lossless LC filters is realised using Schematic Solver [3]. Header and the description of function for LC filter drawing is given:

```
function DrawLC(n,k,x0,y0,dx,s,F)
LC filter drawing.
Inputs:
    n – number of nodes;
    k - indicator: if k=0 this function
        draws Cs and Ls; if k=1 this
        function draws Ys and Zs;
    x0,y0 - origin coordinates;
    dx,s - size of elements in scheme;
    F - font size;
Output:
    Figure with filter scheme.
```

Therefore filter is drawn only by specifying number of nodes. Value k indicates whether scheme is described by capacitors and inductances or by impedances and admittances. Other input variables consider size of elements, their position in the scheme and font size. An example is given in Figs. 4 and 5 in case of n=2.

Function `DrawLC` is initialized by number of nodes and that information is used to set the position of filter elements (grid). Then nodes, lines, ground, voltage source, resistors, capacitors and inductors are placed within the grid using functions: `drawnode`, `drawlhv`, `drawnode` `drawgrnd`, `drawvs`, `drawres`, `drawnod drawcap` and `drawimp` [3], respectively. Input and output sections of the filter are drawn without prior knowledge of number of nodes, while the middle section of the filter is dependent on number of nodes.
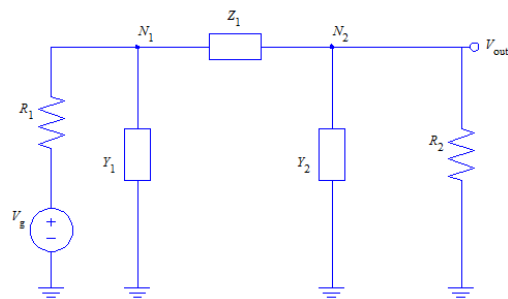
Fig. 4. Representation of LC filter with two nodes: impedance/admittance representation
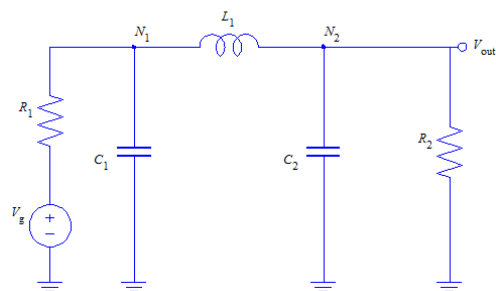
Fig. 5. Representation of LC filter with two nodes: capacitor/inductor representation

IV. TRANSFER FUNCTION CALCULATION

Function `TFcalcLC` calculates transfer function of the filter and generates its graph. The list of input and output parameters for function calculating transfer function of LC filters is given:

```
Function[H,H1,H2]=TFcalcLC(n,w,R10,R20,L0,C)
Transfer function (TF) calculation in LC
filter.
```

```
Inputs:
    n - number of nodes;
    w - frequency range(rad/s);
    R10 - input resistor value;
    R20 - output resistor value;
    L0 - 1x(n-1) array with
         inductors' values;
    C0 - 1xn array with capacitors'
          values;
Outputs:
    H - TF calculated w.r.t.
        R1,R2,Z1..Zn-1,Y1..Yn;
    H1 - TF calculated w.r.t.
        R1,R2,L1..Ln 1,C1..Cn,s;
    H2 - TF calculated w.r.t. s;
    figure with |H(iw)| plot and
         H(s) fraction in Latex.
```

This function is initialized by number of nodes as with the `DrawLC` function while here the frequency range is needed to be precised. Resistors' values, R10 and R20, all inductors' and capacitors' values, L0 and C0, should be given. Similarly as with the `DrawLC` function, all filter elements (impedances and admittances) are initialized depending on number of nodes.

After initialization the system of linear equations in Laplace domain is symbolically solved. Solution with impedances/admittances is given as output H and solution with capacitors/inductors is given as output H1. Numerical solution for transfer function is calculated by substitution of numerical inputs in the H1 and substituting 's' with 'jω'. It is plotted and the transfer function with numerical values is given within the figure. An example of using `TFcalcLC` function is given in Fig. 6.

$$H(s) = \frac{5}{36\,s^3 + 24\,s^2 + 21\,s + 5}$$
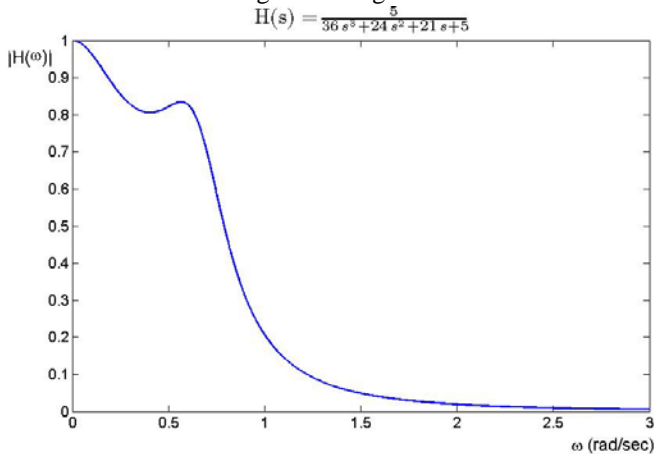
Fig. 6. Transfer function calculated and plotted for filter from Figs. 1 and 2

V. APPLICATIONS

This chapter brings four examples of using functions described in previous two chapters. They illustrate typical use of codes for automatic drawing of LC filters and calculation of transfer function and impact of parameter variation on transfer function.

A. Basics

This example illustrates straightforward application of `DrawLC` function for visualizing LC filters. Different order LC filters obtained using `DrawLC` function are shown in Figs. 7, 8 and 9.
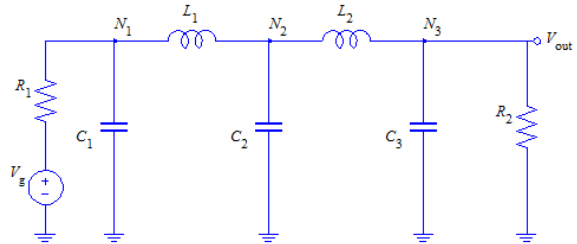
Fig. 7. Representation of LC filter with three nodes (capacitor/inductor representation)
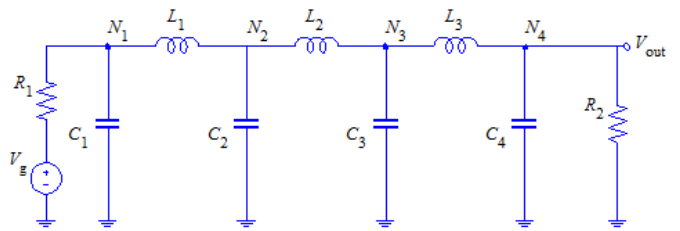
Fig. 8. Representation of LC filter with four nodes (capacitor/inductor representation)
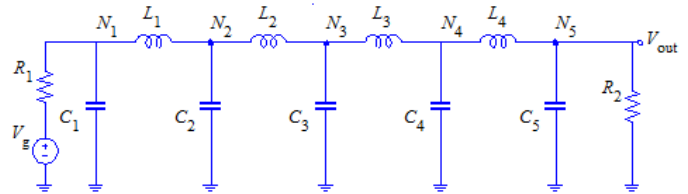
Fig. 9. Representation of LC filter with five nodes (capacitor/inductor representation)

B. Approximation Function

Approximation function [4], [5] is calculated using next lines:

```
M2 = 1/H1/subs(H1,'s',-s)-4;
syms w; M2 = subs(M2,'s',i*w);
pretty(factor(M))
```

As a result there is approximation function (Fig. 10).

$$w^2 (L_1 C_1^2 w^2 - 2 C_1 + L_1)^2$$

Fig. 10. Approximation function of symmetric LC filter with two nodes

Approximation function will be used in example D. Approximation functions for symmetric LC filters with different number of nodes are given in Figs. 11, 12 and 13 in factorized form.

$$w^2 \left(C_2 \, C_1{}^2 \, L_1{}^2 \, w^4 - 2 \, C_1{}^2 \, L_1 \, w^2 - 2 \, C_2 \, C_1 \, L_1 \, w^2 + 2 \, C_1 + C_2 \, L_1{}^2 \, w^2 - 2 \, L_1 + C_2 \right)^2$$

Fig. 11. Approximation function for symmetric LC filter with three nodes

$$w^2 \left(L_2 \, C_1{}^2 \, C_2{}^2 \, L_1{}^2 \, w^6 - 2 \, C_1{}^2 \, C_2 \, L_1{}^2 \, w^4 - 2 \, L_2 \, C_1{}^2 \, C_2 \, L_1 \, w^4 + 2 \, C_1{}^2 \, L_1 \, w^2 + \right.$$

$$L_2 \, C_1{}^2 \, w^2 - 2 \, L_2 \, C_1 \, C_2{}^2 \, L_1 \, w^4 + 4 \, C_1 \, C_2 \, L_1 \, w^2 + 2 \, L_2 \, C_1 \, C_2 \, w^2 - 2 \, C_1 +$$

$$\left. L_2 \, C_2{}^2 \, L_1{}^2 \, w^4 + L_2 \, C_2{}^2 \, w^2 - 2 \, C_2 \, L_1{}^2 \, w^2 - 2 \, L_2 \, C_2 \, L_1 \, w^2 - 2 \, C_2 + 2 \, L_1 + L_2 \right)^2$$

Fig. 12. Approximation function for symmetric LC filter with four nodes

$$w^2 \left(C_3 \, C_1{}^2 \, C_2{}^2 \, L_1{}^2 \, L_2{}^2 \, w^8 - 2 \, C_1{}^2 \, C_2{}^2 \, L_1{}^2 \, L_2 \, w^6 - 2 \, C_3 \, C_1{}^2 \, C_2 \, L_1{}^2 \, L_2 \, w^6 + 2 \, C_1{}^2 \, C_2 \, L_1{}^2 \, w^4 - \right.$$

$$2 \, C_3 \, C_1{}^2 \, C_2 \, L_1 \, L_2{}^2 \, w^6 + 4 \, C_1{}^2 \, C_2 \, L_1 \, L_2 \, w^4 + C_3 \, C_1{}^2 \, L_1{}^2 \, w^4 + 2 \, C_3 \, C_1{}^2 \, L_1 \, L_2 \, w^4 - 2 \, C_1{}^2 \, L_1 \, w^2 +$$

$$C_3 \, C_1{}^2 \, L_2{}^2 \, w^4 - 2 \, C_1{}^2 \, L_2 \, w^2 - 2 \, C_3 \, C_1 \, C_2{}^2 \, L_1 \, L_2{}^2 \, w^6 + 4 \, C_1 \, C_2{}^2 \, L_1 \, L_2 \, w^4 + 4 \, C_3 \, C_1 \, C_2 \, L_1 \, L_2 \, w^4 -$$

$$4 \, C_1 \, C_2 \, L_1 \, w^2 + 2 \, C_3 \, C_1 \, C_2 \, L_2{}^2 \, w^4 - 4 \, C_1 \, C_2 \, L_2 \, w^2 - 2 \, C_3 \, C_1 \, L_1 \, w^2 - 2 \, C_3 \, C_1 \, L_2 \, w^2 + 2 \, C_1 +$$

$$C_3 \, C_2{}^2 \, L_1{}^2 \, L_2{}^2 \, w^6 - 2 \, C_2{}^2 \, L_1{}^2 \, L_2 \, w^4 + C_3 \, C_2{}^2 \, L_2{}^2 \, w^4 - 2 \, C_2{}^2 \, L_2 \, w^2 - 2 \, C_3 \, C_2 \, L_1{}^2 \, L_2 \, w^4 +$$

$$2 \, C_2 \, L_1{}^2 \, w^2 - 2 \, C_3 \, C_2 \, L_1 \, L_2{}^2 \, w^4 + 4 \, C_2 \, L_1 \, L_2 \, w^2 - 2 \, C_3 \, C_2 \, L_2 \, w^2 + 2 \, C_2 + C_3 \, L_1{}^2 \, w^2 + 2 \, C_3 \, L_1 \, L_2 \, w^2 -$$

$$\left. 2 \, L_1 + C_3 \, L_2{}^2 \, w^2 - 2 \, L_2 + C_3 \right)^2$$

Fig. 13. Approximation function for symmetric LC filter with five nodes

C. Parameter Variation

This example shows the influence of one parameter in LC filter on transfer function. We assume symmetric LC filter with three nodes varying ratio of capacitors values in range 1-5 with step of 0.5. Transfer functions for different values of this parameter are shown in Fig. 14.

This procedure is used in analysis of pass-band width when it is possible to implement quartz crystals with different parameters. Iterative implementation of this analysis should provide crystal with optimal parameters according to amplification (or attenuation) in prescribed frequency range.

D. Inverse Problem

Contrary to previous example herein is illustrated problem of estimating one filter parameter with respect to given criterion. Thus Fig. 15 illustrates optimal inductance values considering amplitude characteristic being one for a range of frequencies around 2 rad/s. The inductance is calculated as the first derivative of approximation function in closed form:
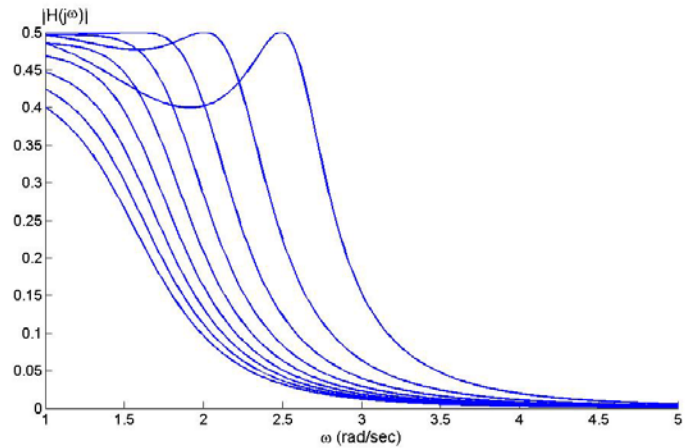
Fig. 14. Transfer functions for different ratios of capacitors values in the symmetric LC filter with three nodes

```
M=subs(H1,'s',1i*w)*subs(H1,'s',-
1i*w);
L = solve(simplify(diff(sqrt(1/M-
1))),'L0');
w = -2.7:0.05:2.7;
L = subs(L,'w',w);
L = min(L);
```
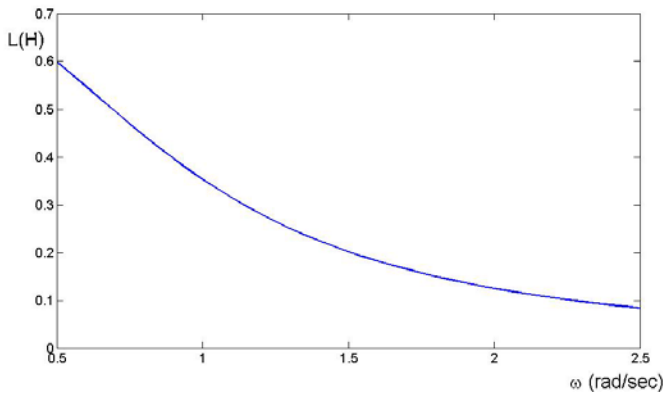
Fig. 15. Optimal inductance values w.r.t. amplitude characteristic being one for a range of frequencies around 2 rad/s

VI. CONCLUSION

In this paper we presented computer codes for drawing and analysing LC lossless filters primarily motivated by crystal filters design. We showed some useful applications of those codes that can be used instead of demanding manual calculations.

This is only the first step in our research towards crystal filter analysis in means of automatization. In our future work we shall consider computer analysis of LC filters with losses. As it was indicated in the paper non ideal model of crystal considers losses introduced by internal resistance of quartz crystal.

REFERENCES

[1] Dragi M. Dujković, Dubravka R. Jevtić, Snežana Dedić-Nešić, Lenkica Grubišić, Irini Reljin, Branimir Reljin, "*High-quality OCXO for Digital TV*" in Proc. *TELSIKS 2009*, vol. 1, pp. 281-284, Nis, Serbia, 2009.

[2] Dragi M. Dujković, Snežana Dedić-Nešić, Lenkica Grubišić, Branimir Reljin, Irini Reljin, "*Crystal Filter 50 MHz for Applications in Specific Environmental Conditions*" TELSIKS 2009, Vol. 1, pp. 253-256, Nis, Serbia, 2011.

[3] M. D. Lutovac, D. V. Tošić, *SchematicSolver Version2.2*, 2010:
http://books.google.com/books?id=9ue-uVG__JsC

[4] M. D. Lutovac, V. D. Pavlović,"*Symbolic Optimization of Symmetric Lossless Filters*", Infoteh 2011, Jahorina, March, 2011.

[5] V. D. Pavlović, M. D. Lutovac, "*Automatizovano projektovanje simetričnih LC lestvičastih filtara za zadati red filtra*", TELFOR 2010, Beograd, Nov. 23-25, 2010.

Network Simulator Tools and GPU Parallel Systems

Leonid Djinevski, Sonja Filiposka, and Dimitar Trajanov

Abstract – In this paper we discuss the possibilities for parallel implementations of network simulators. Specifically we investigate the options for porting parts of the simulator on GPU in order to utilize its resources and obtain faster simulations. We discuss few issues which are unsuitable for the GPU architecture, and we propose a possible work around for each of them. We introduce a design of parallel module that interconnects with a network simulator, while maintaining transparency in aspect of the simulation modeler.

Keywords – Network Simulator Tools, HPC, GPGPU, CUDA, OpenCL.

I. INTRODUCTION

Network simulators are tools used by researchers in order to test new scenarios and protocols in a controlled and reproducible environment, allowing the user to represent various topologies, simulate network traffic using different protocols, visualize the network and measure the performances. Although network simulators are very useful, most of the widely used network simulators do not scale [1]. Simulation of medium to large networks results in a long simulation time which is not practical for investigating protocols.

With the development of parallel systems, significant processing power is becoming available. The single instruction, multiple data (SIMD) models of parallel systems, more particular the Graphics Processing Units (GPUs) have provided a massive acceleration. Additionally, the low cost of these units have brought a huge performance in the insides of regular personal computers (PCs). The first attempts for utilizing the GPU hardware for general purpose computing proved to be a very complicated process [2]. However, with development of the Compute Unified Device Architecture (CUDA) programming model in 2007 [3], and also with the publishing of the standard Open Computing Language (OpenCL) late 2008 [4], general purpose computing on graphics hardware has significantly improved. Therefore, many general purpose applications have been ported for the GPU architecture.

Network simulators have traditionally been developed for execution on sequential computers. Developing a parallel implementation for a network simulator is not straight forward. There are many architectural issues that

Leonid Djinevski, Sonja Filiposka and Dimitar Trajanov are with the E-TNC Research Group, Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Rugjer Boshkovikj 16, 1000 Skopje, Macedonia, E-mail: {leonid.djinevski, sonja.filiposka, dimitar.trajanov}@finki.ukim.mk.

have to be taken in to account and they might prevent the complete utilizing of the GPU resources.

In this paper we review few of the most widely used network simulators. We also discuss the possibilities for parallel implementations of network simulators. Specifically we investigate the options for porting parts of the simulator on GPU in order to utilize its resources and obtain faster simulations. Additionally, we identify modules which carry the biggest workload as well as possible, issues that make the network simulators unsuitable for the GPU architecture, and we propose resolutions to work around these issues.

This rest of this paper is organized as follows: We review implementations of network simulator tools in Section 2, followed by a short overview of the GPU computing in Section 3. In Section 4 we identify which modules of the network simulator contain intensive workloads. Also in this Section we propose a framework which will utilize the GPU resources. In Section 5 we analyze performance, and we conclude and propose future work in Section 6.

II. RELETED WORK

There are two types of approaches for developing a parallel network simulator. One can create the parallel simulator from scratch, where all the simulation software is custom designed for a particular parallel simulation engine. For this approach a significant amount of time and effort are necessary to create a useable system. This is so, because new models must be developed, and therefore validated for accuracy.

An example of this approach is the Global Mobile Information System Simulator (GloMoSim), which is a scalable simulation library designed at UCLA Computing Laboratory to support studies of large-scale network models, using parallel and/or distributed execution on a diverse set of parallel computers [5]. GloMoSim beside sequential adopts parallel simulation model using libraries and layered API. The libraries are developed using PARSEC [6], which is a parallel C based programming language which uses message based approach.

Another example is the Scalable Simulation Framework (SSFNet) which claims that is a standard for parallel discrete event network simulation [6, 7]. SSFNET's commercial Java implementation is becoming popular in the research community, but SSFNet for C++ (DaSSF) does not seem to receive nearly as much attention, probably due to the lack of network protocol models. It is a high performance network simulator designed to transparently

utilize parallel processor resources, and therefore scales to a very large collection of simulated entities and problem sizes.

The second approach for developing parallel/distributed simulation involves interconnecting with existing simulators. These federated simulations may include multiple copies of the same simulator (modeling different portions of the network), or entirely different simulators. Few parallel implementations of this approach are presented in the following.

The NS-2 Simulator [8] is widely used in the networking research community and has found large acceptance as a tool to experiment new ideas, protocols and distributed algorithms. It is a discrete event driven sequential network simulator, developed at UC Berkeley by numbers of different researchers and institutions. NS-2 is suitable for simulating and analyzing either wired or wireless network sand is used mostly for small scale simulations. NS-2 is written in C++ and OTcl. The users define the network topology structure, the nodes, protocols and transmitting times in an OTcl script. The open source model of NS-2 encourages many researchers from institutions and universities to participate and contribute to improve and extend the project. NS-2 plays an important role especially in the research community of mobile ad hoc networks, being a sort of reference simulator [9]. Adding new network objects, protocols and agents requires creation of new classes in C++ and then linking them with the corresponding OTcl objects.

A parallel simulation extension for the traditionally widely used NS-2 simulator has been created at the Georgia Institute of Technology (PADS Research Group), but it is not in wide use. The Parallel/Distributed NS (*PDNS*) [10] was designed to solve the NS-2 problems with large scale networks by running the simulator on a network of workstations connected either via a Myrinet network, or a standard Ethernet network using the TCP/IP protocol stack. In that way the overall execution time of the simulation should be at least as fast at the original single workstation simulation, allowing simulating large scale networks.

Georgia Tech Network Simulator (GTNetS) is a network simulation environment which uses C++ as a programming language [11]. GTNetS is designed for studying the behavior of moderate to large scale networks. The simulation environment is structured as an actual network with distinct separation of protocol stack layers.

OMNeT++ is a network simulation library and framework, primary used for simulation of communication networks, but because of its flexible architecture can be used to simulate complex IT systems too. OMNeT++ offers an Eclipse based IDE and the programming language used is C++ [12, 13].

In this paper we introduce a different approach for parallelizing network simulators that is based on federation simulations. In order to fully utilize the available hardware we investigate the possibility to port the computing intensive network simulator modules to the GPU and thus obtain faster simulation time.

III. GPGPU, CUDA AND OPENCL

In this section we summarize some key fact of the GPU architecture so we can provide and discuss information about parallel module implementation of a network simulator. The origin of General-Purpose computing on Graphics Processing Units (GPGPU) comes from graphics applications, so in similar fashion, CUDA or OpenCL applications can be accelerated by data-parallel computation [14] of millions of threads. A thread in this context means an instance of a kernel, which is a program that is running on the GPU. This way, the GPU device can be visualized as a SIMD parallel machine. Therefore, understanding of the graphics pipeline to execute programs is not needed. In a nutshell, CUDA or OpenCL provide convenient memory hierarchy, allowing maximizing the performance, by optimizing the data access. The memory hierarchy of a GPU device is presented in Fig 1.
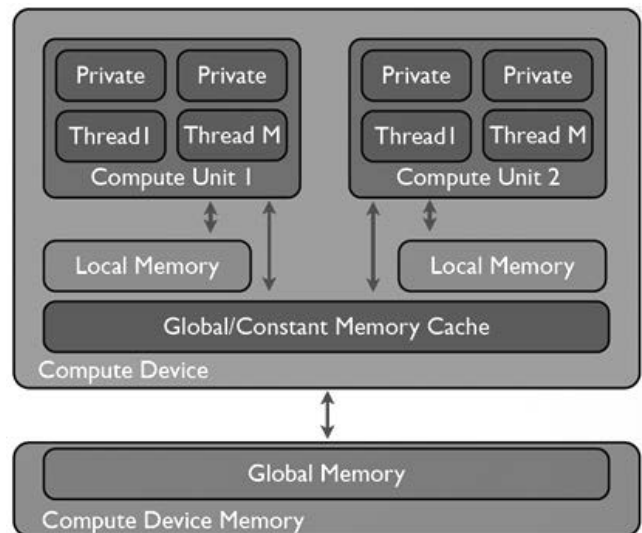
Fig. 1. GPU device memory hierarchy

The GPU device has off-chip memory, so called global memory. Since this memory is separated from the GPU, a single fetching of data takes at least 500 cycles. This is the slowest memory on the device, and therefore the most expensive performance wise.

The next level in the memory hierarchy is the local memory, which is shared by a number of threads organized in work groups. This memory is very small 16 – 48KB, and it can be accessed almost as fast as register memory denoted in Fig 1 as private memory which is exclusive to a single thread. Therefore, a program will compute correctly if there is no data dependence between threads in different work groups. Exception is that within the same work group thread can have dependence because they can exchange data using the local memory.

IV. NETWORK SIMULATOR MODULES

Network simulator algorithms are usually not so straight forward for mapping on the GPU, therefore we need to identify the workload of each module. The modules with the biggest workload are candidates for parallelization. Since, the GPU is a SIMD, in order to utilize the architecture, we look for segments of the algorithm code which are repeated regularly. Usually, these code segments are for loops or loops for which control flow can be predicted.

Once we identify which modules to parallelize, few issues have to be taken in to account. If the code segment works with small amount of data, the GPU device parallelism cannot be expressed. Another major issue is the control flow divergence. If the code segment contains much branching, the parallel code gets serialized, thus minimal or no performance increase is achieved. Nevertheless, in order to tweak the algorithm, few methods can be used to decrease the divergence. However, the worst divergence situation is presented in Listing 1.

```
LISTING 1. Unavoidable Divergence.
if (condition 1)
    do this block of operations
else if (condition 2)
    do that block of operations
else if (condition 3)
    do some block of operations
else
    do any block of operations
```

In this case the divergence can cause up to 75% efficiency reduction, because the block of operation requires hundreds of instructions, thus making the algorithm unsuitable for SIMD parallel execution.

A. Program transformations

In order to exploit more parallelism from the resources at hand, the program has to be transformed. The structure of the computations and their schedule need to be changes, so the program transformations will result with equivalent program which will have better performance.

Since data access is the most expensive part of the program execution, sometimes the program can be transformed so the data is not loaded from memory and calculated on the GPU device. In addition, another important factor is to have enough data to process in order to utilize the parallel resources. Therefore, it is prudent to introduce more calculation even if there are not needed at the moment, since in the following moments a requested calculation could already been obtained.

V. PERFORMANCE ANALYSIS

In order to obtain relevant results, we propose using a GPU device from the high-end segment. An example of a high-end GPU device is the Nvidia Tesla C2070 GPU, which is the flag holder device for Nvidia at the moment of writing this paper.

Regarding parallelism, the Amdahl Law is plotted in Fig. 1, where the x-axis is the number of processors p, and the y-axis is the achieved speedup.
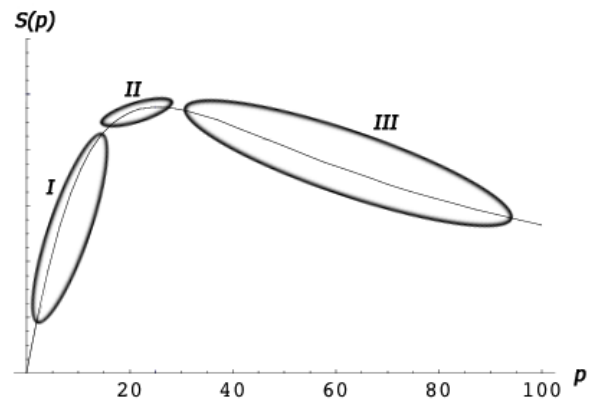
Fig. 2. Parallel speedup

There are three segments that can be noticed on the plot. The segment I represents a relation between the speedup and the number of processors, where by increasing the number of processors. In the second segment, a saturation is achieved, so the speedup stays constant with the increasing the number of processors. The segment III, indicates that increasing of the number of processors, can lead to decreasing of the speedup, which is a consequence of much more communication between the processors and much less computing achieved.

Since for a given GPU device, the number of cores is constant, the plotted curve will depend of the amount of data that is being computed as it is presented in Fig. 3.
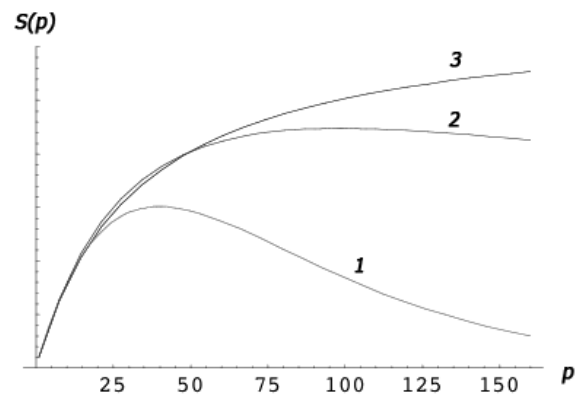
Fig. 3. Parallel speedups for different data amounts

The curve 1 is the same curve as plotted in Fig 2. Curves 2 and 3 present the speedup for larger data quantities. Hence, we can conclude that for larger data quantities, the curve achieves saturation much slower.

Therefore, the network simulator parallel module, should scale well over different sizes of networks, in such a way that the simulation scenarios of interest are in the linear segment I, and possibly, if unavoidable in the saturation segment II.

The parallel module should achieve maximal speedup of at least x25 on a high-end TESLA C2070 GPU for the overall execution of the network simulator. This is a reasonable performance increase that is consistent with many real-life applications ported to the GPU platform, thus providing another example of achieved acceleration by utilizing the computational power of modern programmable GPU devices.

VI. CONCLUSION

Specific modules of the network simulators demand high computational resources. Therefore, we propose a parallel module for the network simulator in order to utilize the computational performance of GPU devices. Usually the network simulator algorithms run in single precision, so the GPU devices are suitable, although the fact that the GPUs support double precision which is still significantly slower.

In our future work, we intend to develop an implementation of a parallel module for one of the few most widely used network simulators. Also, we would like to evaluate how the GPU implementation of the network simulator extension can perform in specific case network topologies. In addition, we would like to search for the best suitable data structures that can provide further optimization. Beside the stand alone machine setup, we would like to test our parallel module on a multi-GPU setup. Additionally we would like to combine MPI and OpenCL, in order to investigate how parallel module will perform on a cluster of computers, where each computer has a multi-GPU setup.

REFERENCES

[1] Weingartner, E., vom Lehn, H., Wehrle, K., "A Performance Comparison of Recent Network Simulators" in Conf. Rec. 2009. ICC '09. IEEE Int. Conf. Communications, pp. 1-5.

[2] Harris, M.J., "*General Purpose Computation on GPUs*", retrieved June 2011 from http://www.gpgpu.org/.

[3] NVIDIA CUDA, retrieved February 2010 from http://developer.nvidia.com/object/cuda.html/.

[4] The OpenCL Specification, Version 1.0, document Revision 43, 2009, retrieved February 2010 from http://www.khronos.org/opencl/.

[5] Zeng, X., Bagrodia, R., Gerla, M., "*GloMoSim: A Library forParallel Simulation of Large-Scale Wireless Networks*", in Proc.12th Workshop on Parallel and Distributed Simulation, Banff, Alta.Canada, 1998, p. 154-161.

[6] Parallel Simulation Environment for Complex Systems (PARSEC), retrieved June 2010 from http://pcl.cs.ucla.edu/projects/parsec/.

[7] Cowie, J.H., Nicol D.M., and Ogielski A.T., "Modeling the GlobalInternet", *Computing in Science and Engineering*, 1999.

[8] NS-2 Simulator, retrieved June 2010 from: http://nsnam.isi.edu/nsnam/index.php.

[9] Di Caro, G. A., "Analysis of simulation environments for mobile adhoc networks", Technical Report No. IDSIA-24-03, IDSIA / USISUPSI,BISON project, Switzerland, 2003.

[10] Riley, G., Fujimoto, R.M., Ammar, M., "A Generic Framework for Parallelization of Network Simulations", in Proc. 7th Int.Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 1999, p. 128-135.

[11] Riley, G.F., "The Georgia Tech Network Simulator", in Proc. of the Workshop on Models, Methods, and Tools for Reproducible Network Research (MoMe Tools), 2003.

[12] Varga, A., "The OMNeT++ discrete event simulation system", Proc. of the European Simulation Multiconference (ESM '2001), Prague, Czech Republic, 2001.

[13] Sekercioglu, Y. A., Varga, A., Egan, G. K., "Parallel Simulation Made Easy With Omnet++", in Proc. of the European Simulation Symposium (ESS2003), Oct. 2003, Delft, The Netherlands.

[14] Grama, A., Gupta, A., Karypis, G., Kumar, V., Introduction to Parallel Computing, 2nd Edition, Addison-Wesley, Reading, MA, 2003.

Modeling and Simulation of *L*-branch Selection Combining Diversity Receiver in Nakagami-*m* Environment using Matlab

Mihajlo Stefanović, Dragan Drača, Aleksandra Panajotović, and Nikola Sekulović

Abstract – In this paper, motivated by the fact that wireless channels are exposed to fading, *L*-branch selection combining (SC) diversity receiver operating over Nakagami-*m* fading environment in the presence of cochannel interference (CCI) is modeled and simulated using program package Matlab. Level crossing rate (LCR), as second order statistic, is chosen to indicate performance measure. Simulation results show great agreement with earlier published numerical results.

Keywords – Fading, Selection combining diversity, Level crossing rate, Sum-of-sinusoids-based simulator.

I. INTRODUCTION

In cellular mobile radio systems, the main causes for the performance degradation are fading due to multipath propagation and cochannel interference (CCI) due to frequency reuse [1]. In the open technical literature, several statistical models are used to describe fading in wireless environments. The most frequently used distributions are Rayleigh, Nakagami-*m*, Rician and Weibull.

Space diversity techniques, which combine input signals from multiple receive antennas, are the well known techniques that can be used to alleviate the effects of degradations [2]. Diversity improvement is achieved without increasing transmission power and bandwidth, but at the expense of increased system complexity and moderate increase in receive power consumption. The most popular ones are maximal-ratio combining (MRC), equal-gain combining (EGC), and selection combining (SC) [3]. MRC and EGC require all or some of the channel state information (fading amplitude, phase, and delay). In addition, a separate receiver chain is needed for each diversity branch which increases its complexity. In opposition to MRC and EGC, SC receiver is much simpler for practical realization because it processes only one of the diversity branches. In general, branch with the highest signal-to-noise ratio (SNR) (or equivalently, with the strongest signal assuming equal noise power among the antennas) is connected to the output. Efficient cellular

Mihajlo Stefanović, Dragan Drača and Aleksandra Panajotović are with the Department of Telecommunications, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {mihajlo.stefanovic, dragan.drača, aleksandra.panajotic}@elfak.ni.ac.rs.

Nikola Sekulović is with the Faculty of Information Technologies, Alfa University, Palmira Toljatija 3, 11000 Belgrad, Serbia, E-mail: sekulani@gmail.com

system designs are interference-limited [2], i.e. the level of the thermal noise is sufficiently low as compared to the level of CCI, so the thermal noise effect may be ignored. In that case, three different decision algorithms can be applied: the desired signal power algorithm, the total signal power algorithm and signal-to-interference ratio (SIR) algorithm [3]. We choose to investigate desired signal power algorithm for an interference-limited SC system since it has identical performance as the total signal power algorithm and it is easier to model [4]. Also, it was shown that SIR algorithm provides the best performance for interference-limited environment systems in the sense of outage probability and average fade duration (AFD), but it almost provides the worst performance for the average level crossing rate (LCR). In desired signal power algorithm, SC receiver selects the branch with the largest instantaneous desired signal power.

In the last years, there has been continuing interest in modeling various propagation channels with the Nakagami-*m* model, which describes multipath scattering with relatively large delay-time spreads, with different clusters of reflected waves [5]. It provides good fits to collected data in indoor and outdoor mobile-radio environments and is used in many wireless communications applications [6]-[9].

Motivated by the previous observations, in this paper, *L*-branch SC diversity receiver operating over Nakagami-*m* fading environment in the presence of CCI is modeled and simulated using program package Matlab. We use Nakagami-*m* fading simulator incorporating Pop's architecture with Zhang decomposition algorithm [10]. In other words, a random phase into low-frequency oscillators for gaining the wide-sense stationary property is inserted, while decomposing a real number of the fading figure, *m*, into two parts, an integer and a fraction to accomplish our design [11]. The average LCR of considered system is simulated to reflect the correlation properties of fading channels and provide a dynamic representation of the system outage performance. Furthermore, simulation results are compared with previously published numerical results in papers [4], [12].

II. NUMERICAL RESULTS

Regardless of the branch of science or engineering, theoreticians have always been enamored with the notation

of expressing their results in the form of closed-form expressions [3]. Therefore, in open technical literature, performance measures of wireless systems – outage probability, average bit error probability, channel capacity, amount of fading, AFD and average LCR - were obtained in closed-forms [.

The average LCR of the envelope ratio of desired signal signal and CCI, μ, at threshold μ_{th} is defined as the rate at which a fading process crosses level μ_{th} in a positive (or negative) going direction and is mathematically defined by the Rice's formula [17]

$$N_\mu(\mu_{th}) = \int_0^\infty \dot{\mu} p_{\mu\dot{\mu}}(\mu_{th}, \dot{\mu}) d\dot{\mu}, \qquad (1)$$

where $\dot{\mu}$ denotes the time derivative of μ and $p_{\mu\dot{\mu}}(\mu, \dot{\mu})$ is the joint PDF of random variables $\mu(t)$ and $\dot{\mu}(t)$ in an arbitrary moment t.

Expressions for the average LCR of dual and triple SC diversity system applying desired signal power decision algorithm over Nakagami-m fading channels in the presence of CCI are presented in [4], [12] as

$$N_\mu(\mu_{th}) = \frac{2\sqrt{2\pi} f_m m_I{}^{m_I-0.5} S^{m_I-0.5}}{\Gamma(m_I) m^{m_I} \mu_{th}{}^{2m_I} \Gamma(m)}$$

$$\times \sqrt{Sm_I + m\mu_{th}{}^2} \left\{ \frac{\Gamma(m+m_I-0.5)}{\left(1 + \frac{Sm_I}{m\mu_{th}{}^2}\right)^{m+m_I-0.5}} \right.$$

$$-2\left(\frac{m\mu_{th}{}^2}{\Omega}\right)^{m+m_I-0.5} \int_0^\infty y^{2m+2m_I-2} \exp\left(-\left(\frac{m\mu_{th}{}^2}{\Omega} + \frac{m_I}{\Omega_I}\right)\right)$$

$$\times Q_m\left(\sqrt{\frac{2m\rho\mu_{th}{}^2}{\Omega(1-\rho)}} y, \sqrt{\frac{2m\rho\mu_{th}{}^2}{\Omega(1-\rho)}} y\right) dy, \qquad (2)$$

and

$$N_\mu(\mu_{th}) = \frac{\sqrt{2\pi} f_m \mu_{th}{}^{2m-1} m_I{}^{m_I-0.5} m^{m-0.5} S^{m_I-0.5}}{\Gamma(m_I)}$$

$$\times \sqrt{Sm_I + m\mu_{th}{}^2} \sum_{i,j=0}^\infty \theta^j \alpha \left[\frac{2\Gamma(i+j+m)}{\Gamma(j+m)(1+\rho)^{i+j+m}} \right.$$

$$\times \left(\frac{\Gamma(j+m+m_I-0.5)}{\alpha_1{}^{j+m+m_I-0.5}} \right.$$

$$- \sum_{k=0}^{i+m-1} \frac{\Gamma(j+m+m_I+k-0.5)\theta^k}{k! \alpha_2{}^{j+m+m_I+k-0.5}}$$

$$- \sum_{l=0}^{j+i+m-1} \frac{\Gamma(j+m+m_I+l-0.5)\theta^l(1+\rho)^l}{l! \alpha_3{}^{j+m+m_I+l-0.5}}$$

$$+ \sum_{k=0}^{i+m-1} \sum_{l=0}^{j+i+m-1} \frac{\Gamma(j+m+m_I+k+l-0.5)\theta^{k+l}(1+\rho)^l}{k! l! \alpha_4{}^{j+m+m_I+k+l-0.5}} \right)$$

$$+ \theta^i \left(\frac{\Gamma(i+j+m+m_I-0.5)}{\alpha_5{}^{i+j+m+m_I-0.5}} \right.$$

$$- \sum_{l=0}^{j+m-1} \frac{\Gamma(i+j+m+m_I+l-0.5)\theta^l}{l! \alpha_3{}^{i+j+m+m_I+l-0.5}} \qquad (3)$$

$$- \sum_{k=0}^{i+m-1} \frac{\Gamma(i+j+m+m_I+k-0.5)\theta^k}{k! \alpha_3{}^{i+j+m+m_I+k-0.5}}$$

$$+ \sum_{k=0}^{i+m-1} \sum_{l=0}^{j+m-1} \frac{\Gamma(i+j+m+m_I+k+l-0.5)\theta^{k+l}}{k! l! \alpha_4{}^{i+j+m+m_I+k+l-0.5}} \right],$$

respectively, where f_m is Doppler shift frequency, ρ is the correlation coefficient, m and m_I are Nakagami parameters describing fading severity of desired signal and CCI, respectively, $Q_m(a,b)$ is the generalized Marcum Q-function, average SIR is $S = \Omega/\Omega_I$ and $\theta = m\mu_{th}{}^2/(1-\rho)$, $\chi = m_I S$, $\alpha = \rho^{i+j}/(i! j! \Gamma(m))$, $\alpha_1 = \chi + \theta$, $\alpha_2 = \chi + 2\theta$, $\alpha_3 = \chi + (2+\rho)\theta$, $\alpha_4 = \chi + (3+\rho)\theta$, $\alpha_5 = \chi + (1+\rho)\theta$.

III. SIMULATION RESULTS

The architecture of sum-of-sinusoids-based Nakagami-m simulator is depicted in Fig. 1 [11].
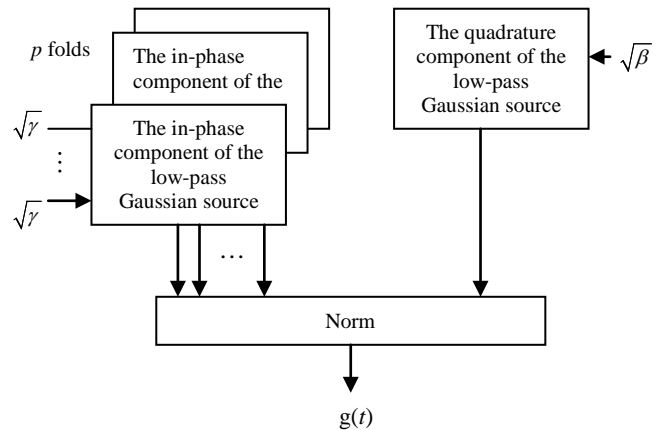
Fig. 1. The block diagram of sum-of-sinusoids-based Nakagami-m simulator

The corresponding composite signal is

$$g(t) = \sqrt{\gamma \sum_{k=1}^{p} g_{I,k}^2(t) + \beta g_Q^2(t)}, \quad (4)$$

where

$$g_I(t) = 2\sqrt{\frac{2}{N}}$$

$$\times \left[\sum_{n=1}^{M} \cos\Phi_n \cos(\omega_n t + \Psi_n) + \sqrt{2} \cos\Phi_N \cos(\omega_N t + \Psi_N) \right], \quad (5)$$

$$g_Q(t) = 2\sqrt{\frac{2}{N}}$$

$$\times \left[\sum_{n=1}^{M} \sin\Phi_n \cos(\omega_n t + \Psi_n) + \sqrt{2} \sin\Phi_N \cos(\omega_N t + \Psi_N) \right], \quad (6)$$

$$\gamma = \frac{2pm \pm \sqrt{2pm(1 + p - 2m)}}{p(1 + p)} \quad (7)$$

and

$$\beta = 2m - \gamma p \quad (8)$$

With $p = [2m]$, $N = 4M+2$, $\omega_n = 2\pi f_m \cos(2\pi n/N)$, $\Phi_n = n\pi/M$, $\Phi_N = 0$ and ψ_j is random phase uniformly distributed in the range $(-\pi, \pi]$.

Having in mind applied decision algorithm in *L*-branch SC receiver, the Fig. 2 describes average LCR simulation process of system operating in Nakagami-*m* environment in the presence of CCI.

Program package Matlab is used to model considered problem. Simulation and numerical results for uncorrelated ($\rho \to 0$) dual and triple SC diversity system in environments under different fading severity are presented in Figs. 3 and 4.

Fig. 3. Average LCR of dual SC diversity system
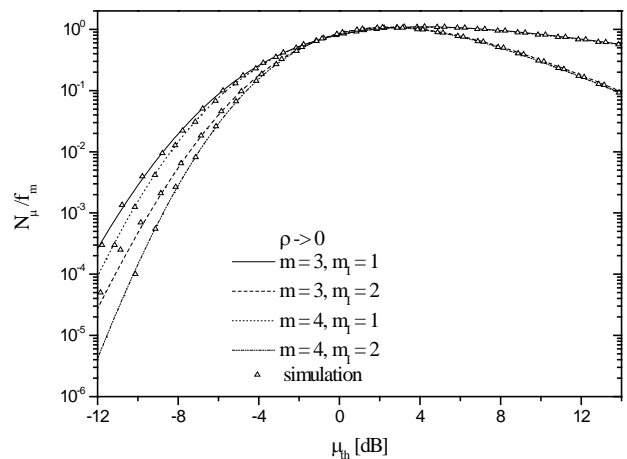
Fig. 4. Average LCR of triple SC diversity system

Fig. 2. The algorithm for simulation of average LCR of considered *L*-branch SC receiver

The great agreement between numerical and simulation results is evident regardless of number of diversity branches or fading severity.

For the reason of greater precision, number of choosen oscillators is $M = 500$. In all simulations maximum Doppler frequency is $f_m = 100$ Hz causing selected $\Delta t = 10 \ \mu s$.

IV. CONCLUSION

This work is result of intention to verify previously published theoretical results. Important and widely accepted performance indicator, LCR, is choosen to be simulated. SC diversity system with two and three uncorrelated branches in Nakagami-*m* fading environment in the presence of CCI is modeled. Simulation results obtained using program package Matlab show great agreement with earlier published numerical results calculated using program package Mathematica.

ACKNOWLEDGEMENT

REFERENCES

[1] Parsons, J. D., "*The Mobile Radio Propagation Channels*", 2nd ed. New York: Wiley, 2000.

[2] Goldsmith, A., "*Wireless communications*", Cambridge University Pres: New York, 2005.

[3] Simon, M. K., Alouini, M.-S., "*Digital Communication over Fading Channel*s", 1st ed. New York: Wiley, 2000.

[4] Yang, L., Alouini, M. -S., "*Wireless communications systems and networks, Average outage duration of wireless communication systems (ch. 8)*", US: Springer, 2004.

[5] Nakagami, M., "*Statistical methods in radio wave propagation. The m-distribution – A general formula if intensity distribution of rapid fading*", Oxford: Ed. Pergamon, 1960.

[6] Charash, U., "*Reception through Nakagami fading multipath channels with random delay*", IEEE Transactions on Communications, COM-27, 1979, pp. 657-670.

[7] Aulin, T., "*Characteristics of a digital mobile radio channel*", IEEE Transactions on Vehicular Technology, VT-30, 1981, pp. 45-53.

[8] Wu, K. T., Hou, J. H., "*Average error probability for quadriphase modulated DS-SSMA communications through Nakagami fading channels*", Telecommunication Systems, Vol. 2, No. 1, 1994, pp.144-158.

[9] Lee, Y. H., Chung, C. H., Cho, S. H., "*Performance analysis of a convolutional coded DS/CDMA system in Nakagami fading channels*", Telecommunication Systems, Vol. 14, No. 1-4, 2000, pp. 31-45.

[10] Zhang, Q. T., "*A decomposition technique for efficient generation of correlated Nakagami fading channels*", IEEE Journal of Selected Areas in Communications, Vol. 18, No. 11, 2000, pp. 2385-2392.

[11] Wu, T-M., Tzeng, S-Y., "*Sum-of-sinusoids-base Simulator for Nakagami-m fading channels*", 58th IEEE Vehicular Technology Conference, 2003. VTC 2003-Fall, Vol. 1, 2003, pp. 158-162.

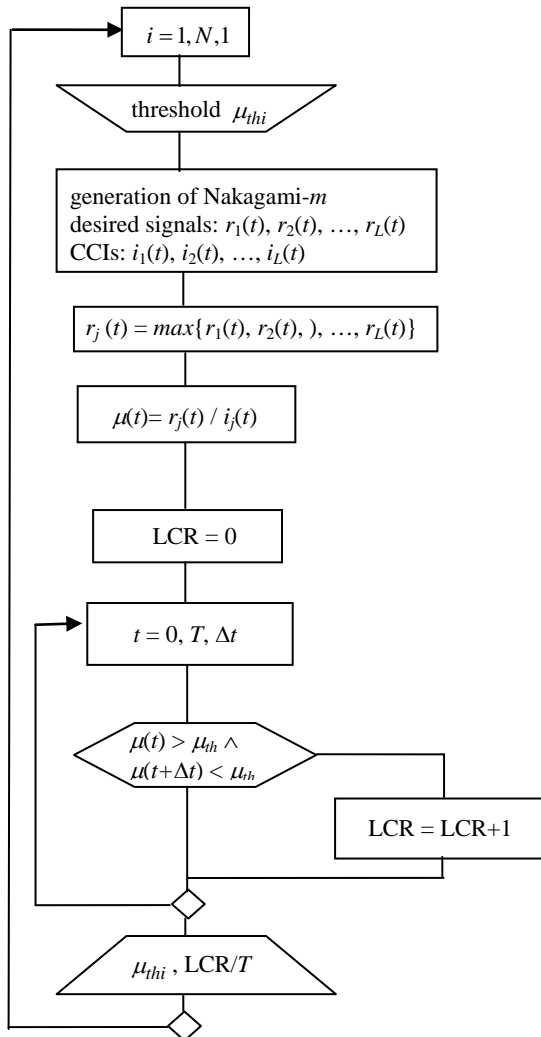[12] Panajotović, A., Stefanović M., Drača, D., Sekulović, N., Stefanović, D., "*Average level crossing rate and average fade duration of triple selection diversity over correlated Nakagami-m fading channels in the presence of cochannel interference*", Telecommunication Systems, under review.

[13] Panajotović, A., Stefanović, M., Drača, D., "*Performance Analysis of System with Selection Combining over Correlated Rician Fading Channels in the Presence of Cochannel Interference*", International Journal of Electronics and Communications-AEÜ, Vol. 63, No. 12, 2009, pp. 1061-1066.

[14] Stefanović, M., Drača, D., Panajotović, A., Sekulović, N., "*Performance Analysis of System with L-branch Selection Combining over Correlated Weibull Fading Channels in the Presence of Cochannel Interference*", International Journal of Communication Systems, Vol. 23, No. 2, 2010, pp. 139-150.

[15] Panajotović, A., Stefanović, M., Drača, D., Sekulović, N., "*Average Level Crossing Rate of Dual Selection Diversity in Correlated Rician Fading with Rayleigh Cochannel Interference*", IEEE Communication Letters, Vol. 14, No. 7, 2010, pp. 605-607.

[16] Panajotović, A., Sekulović, N., Stefanović, M., Drača, D., "*Average Level Crossing Rate of Microcellular Mobile Radio System with Selection Combining in the Presence of Arbitrary Number of Cochannel Interferences*", European Transactions on Telecommunications, accepted for publication.

[17] Dong, X., Beaulieu, N. C., "*Average level crossing rate and average fade duration of selection diversity*". IEEE Communications Letters, Vol. 5, No. 10, 2001, pp. 396-398.

The Decomposition of DSP's Control Logic Block

Borisav Jovanović, Milunka Damnjanović, Dejan Stevanović

Abstract – The paper considers the architecture and low power design aspects of the digital signal processing block embedded into a three-phase integrated power meter IC. Utilized power reduction techniques were focused on the optimization of control logic block. The operations that control unit performs are described together with power-optimization results.

Keywords - digital signal processing, power optimization.

I. INTRODUCTION

Nowadays, the most of circuits used for measurement of power line parameters embed digital signal processors (DSP). This paper proposes a DSP circuit which enables high performances at the level as those obtained with commercial DSP microprocessors, and at the same time, saves the occupied chip area and minimizes power consumption.

The proposed DSP circuit is incorporated into Integrated Power Meter (IPM) system-on-chip. DSP receives from AD converters [1] and digital filters [2] 16-bit digital samples of voltage, current and phase-shifted voltage signals at data-rate of 4096 samples per second, and calculates following power-line parameters:

- root mean square values for voltage and current,
- mean values for active power, reactive power, distortion and apparent power,
- active and reactive energy,
- power factor , and
- frequency.

The measurement range for current signal is from 10mA RMS to 100A RMS, and for voltage it is up to 300V RMS. The results are obtained for three power line phases.

The paper explains the operations performed by DSP, including the novel digital filtering methods, used for processing the instantaneous values of current- and voltage-sample signals. Besides, new circuit for distortion power measurement is presented. Since DSP's control unit is one of largest and most power consuming DSP's part, the paper presents the utilized techniques for power minimization, which are mainly focused on optimization of control logic block.

Borisav Jovanović and Milunka Damnjanović are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia,E-mail:
{borisav.jovanovic, milunka.damnjanovic} @elfak.ni.ac.rs.

Dejan Stevanović is with The Innovation Center, School of Electrical Engineering, University of Belgrade, d.o.o. (ICEF), Bul. Kralja Aleksandra 73, 11120 Belgrade, Serbia, E-mail:
dejan.stevanovic@venus.elfak.ni.ac.rs.

II. DSP'S OPERATION

A. Controller/datapath architecture

The architecture of DSP [3, 4] utilizes controller/ datapath architecture and consists of several blocks:
- Block 1 – the part which consists of arithmetical units used for I^2, V^2, P, Q accumulating and energy calculation
- Block 2 - including arithmetical operators used for calculation of current and voltage RMS, power factor, active, reactive, distortion and apparent power
- Block 3 - control unit that controls all other parts of DSP.
- Block 4 - frequency measurement circuit
- Block 5 - RAM memory block storing the measurement results

DSP's control unit (Block 3) is implemented as finite state machine (FSM). During DSP's measurement operation, the control unit periodically executes main state sequence that lasts 1024 clock periods [5], repeated 4096 times during the time interval of one second. The sequence is divided into four sub-sequences called R, S, T and E that lasts 256 clock periods each. The first three sub-sequences R, S and T control the calculations made for each phase of the three-phase energy system. The fourth sub-sequence, denoted E, manages the calculations that are periodically repeated every second [5].

The control unit is composed of four smaller finite state machines: named F0, F1, F2 and F3. The reason for dividing the control unit is significant power consumption reduction which will be examined in following sections. Two sub-FSMs, F1 and F2, perform arithmetical operations within the Block1 during the phases R, S and T, while sub-FSM F3 - performs operations within Block 2 during E period. The F0 is intended for RAM memory initialization and F0 is active only at the beginning of chip operation, after the main reset state. The operations that F1 and F2 perform will be described in detail.

B. The operation of F1

The FSM F1 executes the state sequence during the phases R, S and T and consists of one hundred and two states.

At the beginning of the F1 operation sequence, the AC part of instantaneous samples of current m_I_{ac} (stored in RAM block) is squared in the multiplication unit within the

Block1. The squared value $m\_I_{ac}^2$ is then passed through the digital Low Pass Filter (LPF), and after, it is accumulated into the accumulation register $m\_AccI_{ac}^2$.

The LPF is implemented as Infinite Impulse Response (IIR) digital filter and helps in reducing the I_{RMS} calculation error. The error could exist because the time interval of one second (that is, accumulating time of the value $m\_I_{ac}^2$) is not always equal to the integer number of power-line-signal half-periods. LPF has cut-off frequency 10Hz and its transfer function is given by Eq.1.

$$H_{LPF}(z) = \frac{2^{-6}}{1 - z^{-1}(1 - 2^{-6})} \qquad (1)$$

The filter transfer function can be transformed into the following equations performed by DSP:

$$m\_FI_{ac}^2 x64_{NEW} = m\_FI_{ac}^2 x64\left(1 - \frac{1}{2^6}\right) + m\_I_{ac}^2 \qquad (2)$$

$$(m\_I_{ac}^2)_{DC} = (m\_FI_{ac}^2 x64)/64 \qquad (3)$$

All these operations are done by arithmetical circuits within the Block 1. The structure of Block 1 is given in Fig.1 and includes one multiplication unit and one circuit for addition and subtraction. Only the inputs (the AC part of current signal m_I_{ac}, values of LPF register $m\_FI_{ac}^2 x64$ and accumulation register $m\_AccI_{ac}^2$) are stored in the RAM memory block. The transfer of data between the Block 5 (RAM memory) and Block 1 is achieved through 24-bit data bus. The intermediate results of operations are temporarily stored in the registers RegA and RegB of Block1 (Fig.1).
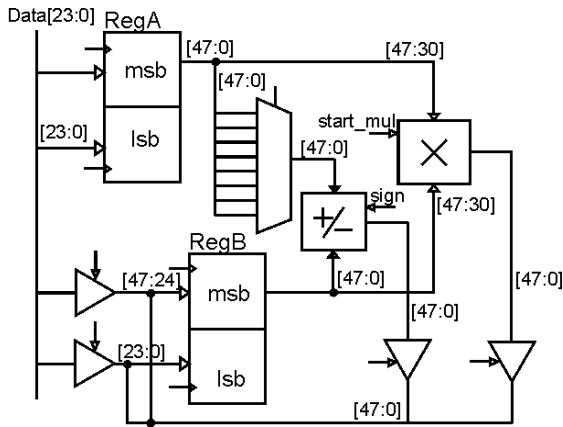
Fig.1 The structure of Block 1

The sequence of operations for the accumulation of squared current values is given by the Fig. 2. The sequence consists of simple data transfer, shifting, multiplication and addition operations which are performed at registers RegA and RegB.

The operations utilize contents of RAM memory registers:

- m_I_{ac} – which contains the AC part of instantaneous current sample
- $m\_FI_{ac}^2 x64\_h$ and $m\_FI_{ac}^2 x64\_l$ are the 24-bit MSB and 24-bit LSB parts of 48-bit LPF register $m\_FI_{ac}^2 x64$, which contains the DC value of I_{ac}^2, multiplied by constant value equal to 64.
- $m\_AccI_{ac}^2$ is 48-bit register for the accumulation of squared current samples.

$$m\_I_{ac} \rightarrow RegA\_h, RegB\_h$$
$$RegA\_h \times RegB\_h \rightarrow RegA$$
$$m\_FI_{ac}^2 x64\_h \rightarrow RegB\_h$$
$$m\_FI_{ac}^2 x64\_l \rightarrow RegB\_l$$
$$RegA - (RegB >> 6) \rightarrow RegA$$
$$RegA + RegB \rightarrow RegA$$
$$RegA\_h \rightarrow m\_FI_{ac}^2 x64\_h, RegB\_h$$
$$RegA\_l \rightarrow m\_FI_{ac}^2 x64\_l, RegB\_l$$
$$m\_AccI_{ac}^2\_h \rightarrow RegA\_h$$
$$m\_AccI_{ac}^2\_l \rightarrow RegA\_l$$
$$RegA + (RegB >> 6) \rightarrow RegA$$
$$RegA\_h \rightarrow m\_AccI_{ac}^2\_h$$
$$RegA\_l \rightarrow m\_AccI_{ac}^2\_l$$

Fig.2 The sequence of accumulation of squared current values controlled by F1

The similar procedure is performed by Block 1 for processing the V_{ac}^2 (necessary for obtaining V_{RMS}) and instantaneous values of active and reactive power. The results are stored in the RAM registers: $m\_AccV_{ac}^2$, m_AccP and m_AccQ. The difference is in the multiplication operands: voltage samples are multiplied to obtain V_{RMS}; voltage and current sample values for active power, and current-sample value is multiplied with phase-shifted voltage-sample for reactive power processing.

C. The operation of F2

The F2 is active during the phases R, S and T. It controls the energy pulses generation for measured active and reactive energy. It consists of one hundred and ninety three states. A pulse is generated when measured energy exceeds some predetermined energy level. The default energy level is one Whr (Watt-hour) for active and VAR (Volt-Ampere reactive) for reactive energy.

The DSP has four outputs producing the narrow pulses: Ea_pos – for consumed active, Ea_neg – generated active, Eq_pos –inductive reactive, and Eq_neg –capacitive reactive energy. The energy level is stored in m_Whr register, the part of RAM memory block, and can be modified. The operations are carried out by Block 1 using

the adder/subtractor and registers RegA and RegB.

The sequence of operations is given in Fig.3. At the beginning of each sequence, performed exactly 4096 times during the time interval of one second, the active power value m_P, is added to the value of 48-bit register m_AccEa. The m_AccEa consists of two parts: the MSB part - m_AccEa_h and the LSB part - m_AccEa_l, both stored in RAM. After addition operation is done, the value of m_P and new value of m_AccEa are compared with zero. If value of m_P is positive and if new value of m_AccEa is greater than the energy level equivalent (given by m_Whr), a pulse on Ea_pos is generated and m_AccEa is subtracted by the m_Whr value. Else, if both m_P and m_AccEa are negative, a pulse on Ea_neg is generated, and value of m_Whr is added to m_AccEa.

The similar procedure stands for the reactive energy processing. Accompanied registers are m_AccEq_h and m_AccEq_h.

$$m\_P \rightarrow RegB\_l$$
$$m\_AccEa\_h \rightarrow RegA\_h$$
$$m\_AccEa\_l \rightarrow RegA\_l$$
$$RegA + RegB \rightarrow RegA$$
$$if\ (RegB > 0)\ \{$$
$$\quad m\_Whr \rightarrow RegB\_l$$
$$\quad if\ ((RegA - (RegB << 12)) > 0)\ \{$$
$$\quad\quad RegA - (RegB << 12) \rightarrow RegA$$
$$\quad\quad genarate\ pulse\ for\ positive\ Ea;$$
$$\quad \}$$
$$\}\ else\ \{$$
$$\quad if\ (RegA < 0)\ \{$$
$$\quad\quad m\_Whr \rightarrow RegB\_l$$
$$\quad\quad RegA + (RegB << 12) \rightarrow RegA$$
$$\quad\quad genarate\ pulse\ for\ negative\ Ea;$$
$$\quad \}$$
$$\}$$
$$RegA\_h \rightarrow m\_AccEa\_h$$
$$RegA\_l \rightarrow m\_AccEa\_l$$

Fig.3 The sequence of operations producing the energy pulses on Ea_neg and Ea_pos pins

Besides dealing with energy pulses, the F2 eliminates DC offsets from instantiations current and voltage signals that are derived from digital filters. This is necessary for the calculation of current and voltage RMS value. The DC offset will give a DC component after squaring operation. Since this DC component is extracted by LPF, this offsets can induce the error to RMS values. This problem is avoided by introducing the HPF in voltage and current signal processing chains. The HPF, applied to instantaneous current and voltage signals, is implemented as Infinite Impulse Response (IIR) digital filter with cut-off

frequency 5Hz and transfer function as given by Eq.4:

$$H_{HPF}(z) = (1 - 2^{-10}) \frac{(1 - z^{-1})}{1 - z^{-1}(1 - 2^{-9})} \quad (4)$$

The HPF transfer function can be transformed into the equations (5) and (6) performed by DSP.

$$m\_FIx1024_{NEW} =$$
$$m\_FIx1024(1 - \frac{1}{2^9}) + (2^{10} - 1)(m\_I - m\_I\_p) \quad (5)$$

$$m\_I_{ac} = m\_FIx1024 / 1024 \quad (6)$$

The following registers values are used in the equations (5) and (6):

- m_I and m_I_p - two consecutive current samples
- m_FIx1024 is 48-bit HPF register, which contains the AC value of I, multiplied by constant value 1024. The register consists of two parts: the MSB part - m_FIx1024_h and LSB part - m_Fix1024_l.
- m_I$_{ac}$ is AC part of instantaneous sample of current signal. It represents the result of filtering operation and it is further used by FSM F1.

$$m\_I\_p \rightarrow RegA\_l$$
$$m\_I \rightarrow RegB\_l$$
$$RegA\_l - RegB\_l \rightarrow RegA\_l$$
$$RegA \rightarrow RegB$$
$$RegA - (RegB << 10) \rightarrow RegA$$
$$//RegA = ((1024 - 1)(m\_I - m\_I\_p))$$
$$m\_FIx1024\_h \rightarrow RegB\_h$$
$$m\_FIx1024\_l \rightarrow RegB\_l$$
$$RegA + RegB \rightarrow RegA$$
$$RegA - (RegB >> 9) \rightarrow RegA$$
$$//RegA = (1024 - 1)(m\_I - m\_I\_p) +$$
$$//\qquad m\_FIx1024\ (1 - 2^{-9})$$
$$RegA\_h \rightarrow m\_FIx1024\_h$$
$$RegA\_l \rightarrow m\_FIx1024\_l$$
$$RegA \rightarrow RegB$$
$$0 \rightarrow RegA\_l$$
$$RegA + (RegB >> 10) \rightarrow RegA$$
$$RegA\_l \rightarrow m\_I_{ac}$$

Fig.4 The sequence for high pass filtering of instantiations current sample signals, done by F2

The operation sequence for the offset elimination,

performed by F2, is given in the Fig.4. The operations are carried out by Block 1.

The similar procedure is made for processing of m_Vac (necessary for obtaining V_{RMS}). The intermediate results are stored in 24-bit RAM registers: m_FVx1024_h and m_FVx1024_l.

D. The operation of F3 FSM

The fourth sub-sequence of the control unit, manages the calculations that are periodically repeated every second and consists of one three hundred and four states.

Based on accumulating sums m_AccI$_{ac}^2$, m_AccV$_{ac}^2$, m_AccP and, m_AccQ, arithmetical operations are performed by Block 2 to generate voltage and current root mean square values m_I$_{RMS}$ and m_V$_{RMS}$ and mean active and reactive power values m_P and m_Q. The sequence of operations is performed by FSM F3.

The interior structure of Block 2 is given in Fig.5. It consists of two registers named RegC and RegD and arithmetical units that implement square rooting, subtraction, multiplication and division.
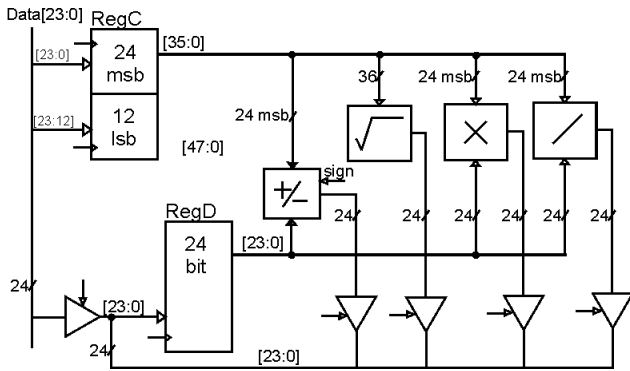
Fig.5 The structure of Block 2

The sequence, controlled by F3 that generates current root mean square m_I$_{RMS}$, is given in Fig.6.

To generate m_I$_{RMS}$, accumulated sum m_AccI$_{ac}^2$ is stored into RegC and then, it is divided by 4096. Next, square rooting operation is performed over the average value of voltage square. Then, current offset m_I$_{ACoff}$ is subtracted, multiplied with gain correction m_Igain and root mean square of current is obtained (Fig.6).

The similar processing steps are conducted for m_V$_{RMS}$. For mean active and reactive power calculation the square root calculation is avoided. Apparent power m_S is obtained by multiplying m_I$_{RMS}$ and m_V$_{RMS}$, and power factor m_CosF – by dividing active m_P and apparent power m_S.

$$m\_AccI^2\_h \to RegC\_h$$
$$m\_AccI^2\_l \to RegC\_l$$
$$\sqrt{RegC} \to RegD$$
$$0 \to m\_AccI^2$$
$$m\_I_{ACoff} \to RegC\_h$$
$$RegC\_h - RegD \to RegD$$
$$m\_Igain \to RegC\_h$$
$$RegC\_h \times RegD \to RegD$$
$$RegD \to m\_I_{RMS}$$

Fig.6 The sequence that generates current root mean square m_IRMS

In addition to finding mean active (m_P), reactive (m_Q) and apparent power (m_S), the distortion power [6] (stored in the register m_D) calculation is provided.. F3 controls the operations producing the m_D. Arithmetical operators used to calculate the value of m_D, belong to blocks 1 and 2. The structure of Block 1 had to be slightly modified. The new input is introduced to RegB which makes the connection from the multiplication unit from Block2. The result of multiplication operation, done by arithmetical operator within Block2, has to be transferred to the RegB in Block1. The sequence is given in Fig.7.

At the beginning, the register RegA is reset to zero, and the content of register m_S is copied to both of the registers RegC and RegD. The squaring operation is performed and the result is moved to the RegA. Then, the active power m_P is moved to RegC and RegD, and the multiplication is performed. The result is subtracted from register RegA. The same operations are done with the value m_Q. After, the content of RegA is moved to the RegC, and square root operation is performed. Finally, the result is moved from RegD into the m_D, which is stored in the RAM memory.

$$0 \to RegA\_h, RegA\_l$$
$$m\_S \to RegC\_h, RegD$$
$$RegC\_h \times RegD \to RegB$$
$$RegA + RegB \to RegA$$
$$m\_P \to RegC\_h, RegD$$
$$RegC\_h \times RegD \to RegB$$
$$RegA - RegB \to RegA$$
$$m\_Q \to RegC\_h, RegD$$
$$RegC\_h \times RegD \to RegB$$
$$RegA - RegB \to RegA$$
$$RegA \to RegC$$
$$\sqrt{RegC} \to RegD$$
$$RegD \to m\_D$$

Fig.7 The sequence that generates distortion power m_D

III. THE IMPLEMENTATION RESULTS

The most of optimization process considered the DSP's control unit. The control unit incorporates over six hundred states and this large number of states required huge combinational logic of synthesized FSM. The implementation occupies large portion of DSP's area. Also, it represents one of the largest power consumers among DSP's blocks.

The following power minimization techniques were used: FSM decomposition [7, 8], clock gating and Grey code encoding [9]. The first technique divides large control unit into several smaller state machines, simplifying their combinatorial logic blocks. The division of control unit into smaller state machines has positive effect on power dissipation. Furthermore, the clock gating disables inactive parts of FSM by stopping its clock signal, and, reduces the switching activity within the combinatorial logic blocks. Besides, Gray binary encodings are assigned to the FSM's states.

The transition graph of original FSM was considered first, and after, divided into four sub-graphs (F0, F1, F2 and F3) that jointly produce the equivalent behaviour as the original FSM. The decomposition is performed by considering the datapath architecture. The states within one subset control the arithmetical operations performed by same part of DSP. As stated earlier, F1 and F2 perform the operations within the Block1, while F3 – mainly within the Block 2.

After the FSM decomposition is done, the clock gating is introduced in the FSM's implementation. New circuit is added into control logic block which identifies currently active sub FSM. The circuit also provides clock input signals to sub FSMs. The clock signal is present only at the input of active sub FSM, and the other three sub FSMs are blocked.
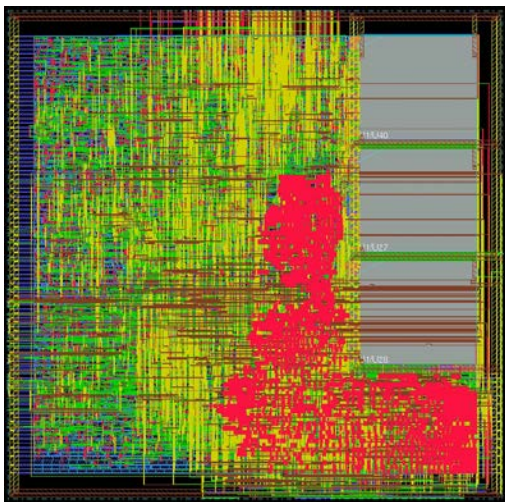
Fig.8 The layout of DSP

The DSP was implemented in technology AMI CMOS 350nm with power supply voltage of 3.3V.

When the design was verified by RTL simulation, the RTL descriptions were loaded into program for logical synthesis, Cadence's RTL Compiler that generated the netlist of digital library cells. The extracted netlist was loaded back to Verilog simulator and the simulation was performed using Cadence' NCsim tool.

SoC Encounter has performed floorplanning, placement and routing, as well as clock and reset trees generation for complete circuit (Fig.8). At the end of logical verification process, Verilog file was extracted from layout and brought back to NCsim simulator where final check of the total digital part of the IC was performed.

During the post-layout simulation, switching activity file was obtained and the power consumption results are obtained by the SoC Encounter taking account the parasitic capacitances from layout and switching activity file.

The estimation of DSP's power consumption gave the valuable information about the energy budget and identified all power hungry components. Three power analyses were performed: for the: (a) original design (before the power minimization techniques where applied), (b) DSP design which is optimized by gating and FSM decomposition, and finally, (c) design where all proposed techniques were applied: FSM decomposition, clock gating and Gray state encoding. The Table 1 gives the simulated power consumption values of different DSP cores, derived after layout generation. The power consumption of not-optimized design was 1.82mW .When all these techniques were applied, the total power became only 1.043mW resulting in the 42% switching power reduction, comparing to the non-optimized implementation.

TABLE I

THE RESULTS OF POWER OPTIMIZATION

	Not optimized	Decomposition & clock gating	Decomp., clock gating & Grey encoding
Area	$1.84mm^2$	$1.831mm^2$	$1.823mm^2$
Clock tree power	0.732mW	0.263mW	0.227mW
Control unit power	0.407mW	0.172mW	0.172mW
DSP's power	1.82mW	1.117mW	1.043mW

IV. CONCLUSION

The architecture and the low power design aspects of the digital signal processing block embedded into a three-phase integrated power meter IC, are considered. The operations that control unit performs are described together with power-optimization results.

The power reduction techniques were successfully implemented on the optimization of the control logic block.

The control unit of DSP block, implemented as finite state machine, was decomposed into four smaller state machines, clock gating was completely introduced and Gray finite state machine encoding used. The resulting effect was the significant reduction of the power consumption.

REFERENCES

[1] Mirković, D., Petković, P. *"Multi channel Sigma-Delta A/D converter for integrated power meter"*, Proceedings of the Small Systems Simulation Symposium 2010, Niš, ISBN 987-86-6125-006-4, Feb., 2010, pp. 90-93.

[2] Marinković, M., Andjelković, B., Petković, P. "*Compact Architecture of Digital Decimation Filters in Solid-State Energy Meter*", Electronics, Vol. 10, No. 2, University of Banja Luka, ISSN 1450, December 2006, pp. 28-32.

[3] Jovanović, B., Damnjanović, M., Petković, P. *" Digital Signal Processing for an Integrated Power Meter"*, Proceedings of 49. Internationales Wissenschaftliches Kolloquium, Technische Universirtat Ilmenau, Ilmenau, ISBN 3-8322-2824-1, Vol.2, September, 2004, pp. 190-195.

[4] Jovanović, B., Zwolinski, M., Damnjanović, M. *"Low power digital design in Integrated Power Meter IC"*, Proceedings of the Small Systems Simulation Symposium 2010, Niš, ISBN 987-86-6125-006-4, Feb., 2010, pp. 49-55

[5] Jevtić, M., Jovanović, B., Brankov, S.*" Upravljačka jedinica sistema na čipu za registrovanje potrošnje električne energije"*, Zbornik radova XLVIII konferencije Etran 2004, Čačak, ISBN 86-80509-49-3, Vol.1, June, 2004, pp. 75-78

[6] Stevanović, D., Jovanović, B., Petković, P., Litovski, V. *"Korišćenje snage distorzije za identifikaciju izvora harmonijskog zagađenja na mreži"*, Tehnika - Elektrotehnika, 6/2011, Savez inženjera i tehničara Srbije, ISSN 0040-2176, 2011, accepted for publication.

[7] Chow, S.H., Yi-Cheng Ho, Y.C., Hwang, T. "*Low power realization of finite state machines - a decomposition approach*", ACM Transactions on Design Automation of Electronic Systems (TODAES) Volume 1, Issue 3 (July 1996) pp.: 315 – 340, ISSN:1084-4309

[8] Lee, W.K., Chi-Ying Tsui, C.Y. "*Finite state machine partitioning for low power*", Circuits and Systems, 1999. ISCAS'99, Proceedings of the 1999 IEEE International Symposium, Volume 1, June 1999, pp. 306 – 309

[9] Benini, L., De Micheli, G. "*State assignment for low power dissipation*", Solid-State Circuits, IEEE Journal of Volume 30, Issue 3, Mar 1995 pp.:258 – 268

Efficient Fault Effect Extraction for an Integrated Power Meter's ΣΔ ADC

Dejan Mirković, Dejan Stevanović and Vančo Litovski

Abstract - Analog to digital converter (ADC) is the vital part of many mixed-signal ICs because it interfaces analog signals from real world with digital logic on a chip. Errors made during conversion are hard to eliminate in the digital part that follows. Therefore, functional testing of ADC is a very important especially during developing of prototypes. Testing techniques for the ADC implemented in an integrated power-meter are considered first. A testing procedure based on efficient fault effect extraction will be proposed and implemented to a ΣΔ ADC. Encouraging results were observed.

Keywords – IMPEG, ΣΔ AD Converter**, DFT, SC integrator

I. INTRODUCTION

Modern power meters are usually realized as solid state integrated SoC circuits. The functionality of this kind of circuits is based on acquiring instantaneous values of voltage and current. These values are obtained through some kind of analog-to-digital data conversion and further processed in digital domain. One such solution is realized as integrated circuit in LEDA laboratory and named IMPEG [1]. Basic block diagram of IMPEG chip is shown on Fig. 1.
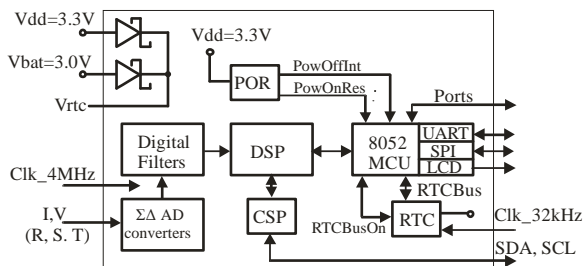
Fig. 1: IMPEG block diagram

From Fig. 1 one can notice the main functional blocks of IMPEG. One, in particular, block which is of prime concern of this paper, is ADC block. In the IMPEG solution ADC is realized as ΣΔ type of converter with second order noise shaping loop [2]. Since the accuracy of the ADC determines the quality of the power measurement, a proper function of this block is very important. So the ADC has to be fully tested.

Dejan Mirković and Vančo Litovski are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {dejan, vanco}@elfak.ni.ac.rs.

Dejan Stevanović is with The Innovation Center, School of Electrical Engineering, University of Belgrade, d.o.o. (ICEF), Bul. Kralja Aleksandra 73, 11120 Belgrade, Serbia, E-mail: dejan.stevanovic@elfak.ni.ac.rs.

It is well known that time consumed for testing is the key parameter to be minimized from the manufacturing point of view. This fact holds primarily because the testing process increases time to market cost. In widely accepted automatic test equipment (ATE) testing of ADCs is based on the verification of a subset of converter performance parameters that include offset, gain, signal-to-noise ratio (*SNR*), total harmonic distortion (*THD*), integral nonlinearity (*INL*), differential nonlinearity (*DNL*) and power consumption. Dynamic performance parameters are estimated from FFT analysis performed over output bit stream. Static performance parameters can be determined from measured code transition edges (feedback loop test) or from the number of code occurrences in response to a periodic signal (histogram testing). The accuracy of such parameter measurements requires a large number of samples (approx. 16536 for 90 dB THD and *SNR* measurements) and up to 10 s of test time for a single channel 16-bit audio ADC.

There is a number of techniques and methods (e.g. built-in self-test (BIST)) for analog and mixed-signal circuitry in high resolution ADCs. All those techniques address the measurement of just one or of only a limited number of specification parameters using conventional techniques. One of such techniques that focus on minimization of test time and data required for FFT analysis is published in [3]. Often, most of these methods tends to require significant die area and dedicated additional test patterns for verification. In this paper a recent solution of design for testability (DFT) for ΣΔ ADC implemented in an integrated power meter is discussed [4].

This paper is organized as follows. The next section deals with different methods for ADC testing. After that we describe a practical way for testing the analog part of ΣΔ AD Converter. The paper concludes with important results obtained after testing the particular ΣΔ ADC.

II. ADC TESTING

Typical ADC tests are: ADC code edge measurement, DC tests, transfer curve tests, and dynamic ADC tests [5]. Each one of them will be briefly explained.

The aim of ADC code edge measurement is to find the input voltage threshold between two successive ADC codes that causes an output code to change. To measure the ADC linearity one needs to derive transfer curve of an ADC. Two well-known methods for transfer curve derivation are *center code testing* and *edge code testing* [4].

The center code testing gives artificially low DNL

value, and because of that this technique should be avoided. There are several different ways to search for the code edges. One of the most common techniques is the histogram method.

The simplest way to perform a histogram test is to apply a rising or falling linear ramp to the input of the ADC and collect samples from the ADC at constant sampling rate. The ADC samples are captured while the input ramp slowly moves from one end of the ADC conversion range to the other. The number of occurrences of each code is plotted as a histogram. It shows which codes are hit more often, indicating that they are wider codes. After obtaining the histogram, a code edge transfer curve must be derived using a mathematical equation that sums the code widths.

To compensate for the poor linearity of the ramp generators, the alternative, sinusoidal histogram method can be used. It is easier to produce a pure sinusoidal waveform than to produce a perfectly linear ramp. This method also allows testing in more dynamic, real-world situation, since ramps are varying very slowly. By using a sinusoidal signal instead of a ramp, one would expect to get more code hits at the upper and lower codes than at the center of the ADC transfer curve, even when testing a perfect ADC. The effects of the non-uniform voltage distribution can be removed after normalization.

DC Tests and Transfer Curve Tests comprise: DC Gain, DC offset, *INL*, *DNL*, monotonicity and missing codes tests. Once the ideal transfer curve has been established, DC gain and offset can be measured. The gain and offset are measured by calculating the slope and offset of the best-fit line.

Dynamic ADC parameters are: maximum sampling frequency, maximum conversion time, and minimum recovery time. More information about these parameters can be found in [5].

Considering all tests listed and an existing ADC architecture, it is very important to determine the significance and the feasibility of tests to be performed. Tests such as *INL* and *DNL* are not well suited for ΣΔ converters. Instead, channel tests like gain, offset, *SNR*, idle channel noise, etc., are commonly specified. When the resolution exceeds 12 or 13 bits, it becomes very expensive to perform transfer curve test such as *INL* and *DNL* because of the large number of code edges that must be measured. Fortunately, transmission parameters such as frequency response signal to distortion ratio (*SNDR*) and idle channel tests are much less time-consuming to measure.

III. ARCHITECTURE OF DFT FOR ΣΔ ADC

Simplified diagram of implemented DFT architecture for ΣΔ ADC is illustrated on Fig. 2. According to this diagram there are two test points available at the chip pins named *analog_out* and *digital_inout*. The pin *analog_out* provides access to two functionally important nodes of ADC's analog part. Practically it buffers multiplexed outputs of the first and the second integrator of ΣΔ

modulator, respectively. The pin *digital_inout* represents bidirectional digital port which has dual role. Firstly, it provides access to the bit stream which represents coarsely quantized, oversampled value of analog input signal. Secondly through this pin the digital filters, that follow the ΣΔ modulator, can be fed with some externally generated bit sequence. So this pin enables implementation of testing algorithms weather for the whole ADC or just for the digital filters.

It is worth to mention that the converted value of analog input signal is stored in a 24 bit wide register which is refreshed with a rate of 4096 kHz. This register can be accessed using three wire serial communication port (SCP) at rate of 400 kHz. More about architecture illustrated in Fig. 2 can be found in [4] and [6].
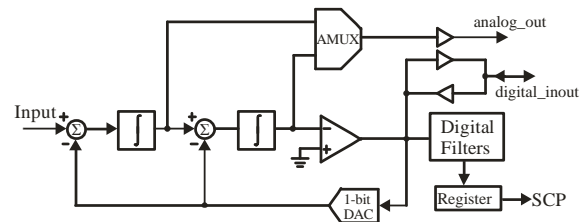
Fig. 2: DFT for ADC

In this section we dealt with the structure and operation of the implemented DFT architecture for ΣΔ ADC. This will help as to explain and establish a testing procedure for the analog part of ADC.

IV. TESTING PROCEDURE FOR ΣΔ ADC

From earlier discussion in section one; one can conclude that ΣΔ ADC class of converters is most often characterized by dynamic parameters such as *SNR* or *SNDR*. Considering this fact the testing procedure presented in this paper adopts *SNR* and *THD* parameters as qualitative measure of ADC functionality. These parameters are extracted from FFT analysis of appropriate signal in one of previously discussed test points.

A. Test procedure

It is important to clarify that the proposed procedure is developed for the analog part of the ADC, namely ΣΔ modulator, and it is confirmed by means of simulation. Proposed testing procedure can be divided in several steps.

First, a test signal needs to be adopted. Because of the coarse quantizing, one bit, it is suitable to scan the output of the quantizer (*digital_inout* in Fig. 2) since there are only two values (logic zero or one). So this signal is measured (observed, simulated) for appropriate amount of time to get enough data for performing FFT. FFT analysis is preformed over collected set of data, and *SNR* and *THD* are estimated.

In order to create the fault dictionary, at this point, the previous steps should be repeated but with defects inserted
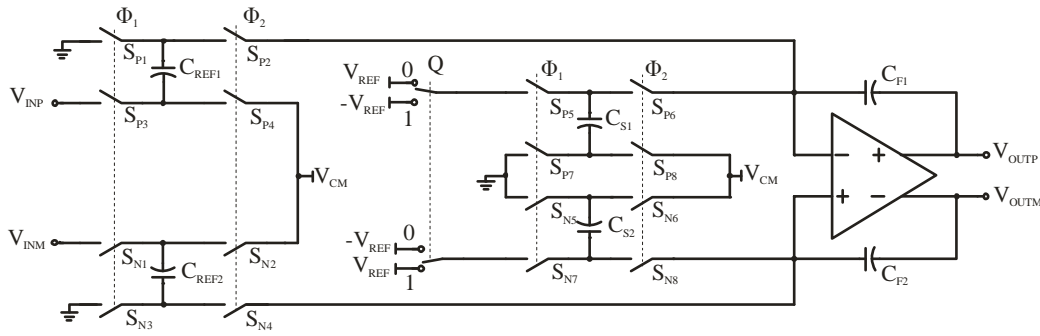
Fig. 3. Structure of the SC integrator in the sigma-delta modulator

in the analog part of the $\Sigma\Delta$ modulator (quantizer, SC integrator etc.). Single fault analysis is conceived. Finally, the responses of the fault-free and the faulty circuits are compared in order to establish testability for every inserted fault.

If there is deviation in *SNR* and/or *THD* values between the responses of the fault-free and the faulty circuit, the fault is covered with the test signal adopted in first step. Otherwise, the fault is masked so different test signal should be adopted, and the procedure repeated. Practically, the procedure is iteratively repeated until the given defect is covered i.e. appropriate test signal discovered. Besides *SNR* and *THD* parameters, visual inspection of the signal's spectrum can be helpful to detect presence of defect in the circuit.

Fortunately, because of the specificity of the circuit, every defect that forces any of the integrators, and the further quantizer, to saturate, will lead to malfunction. So sometimes is just enough to observe the waveform of the quantizer output. But when the effect of defect is not visible from output signal waveform then proposed testing procedure should be used.

B. Functionality of SC integrator

As it was mentioned in previous sections noise shaping is obtained with second order loop filter i.e. $\Sigma\Delta$ modulator. Each of two modulator stages is realized as SC, delaying, parasitic insensitive, integrator. The simplified, fully differential, realization of SC integrator is depicted in Fig. 3. Since the circuit is symmetrical to the horizontal axis for sake of simplicity one can concentrate on the top half of the circuit.

Every analog switch, S_{Ni}/S_{Pi}, where $i= 1, ..,8$, is realized as an NMOS transistor. The voltages V_{INP}, V_{OUTP} and V_{INM}, V_{OUTM} represent the input and output for a positive and negative analog signal, respectively. The port Q is the output of the quantizer while Φ_1 and Φ_2 are clock signal ports. The common mode voltage is denoted by V_{CM}, and the reference voltage with V_{REF}.

The operation of the SC integrator takes place in two non overlapping clock phases provided by Φ_1 and Φ_2. In the first phase, when the clock signal Φ_1 is active, the sampling capacitor (C_{S1}) is charged with charge that

corresponds to instantaneous value of analog input signal, V_{INP}. In the second phase (Φ_2 active) the charge maintained on C_{S1} is transferred to the feedback capacitor C_{F1} and further to the output of the opamp, V_{OUTP}. Similarly, the charge proportional to $k \cdot V_{REF}$, where k is the ratio C_{REF1}/C_{F1}, is added to the feedback capacitor with plus or minus sign depending on the state of Q.

Knowing this, one can notice that the analog switch S_{P2}/S_{P4} (equivalently S_{N2}/S_{N4}) is vital for proper functioning of the SC integrator. During its activation charge is transferred to the output of the integrator. That was why malfunction of S_{P2} is picked to be tested with the proposed procedure. The testing procedure is applied to detect shorts and opens between all four transistor terminals (drain, gate, source and bulk) of the switch.

V. RESULTS

The proposed test procedure is applied on SPICE macro model of $\Sigma\Delta$ modulator which includes adequate MOS transistor model for 0.35 μm CMOS technology process. After simulation, FFT analysis, *SNR* and *THD* estimation are performed using appropriate MATLAB® script. All results are summarized in Table I.

The values for *THD* and *SNR* are given in homonymous columns. Column *detected* provides information weather defect is detected or not with the adopted excitation signal. The second row refers to the fault-free circuit, while the others represent faulty ones. The letters D, G, S, and B correspond to drain, gate, source, and bulk transistor terminals, respectively. Sine wave with 200mVpp magnitude and 50Hz frequency is adopted for excitation. After applying the test procedure to the circuit, this kind of excitation signal is shown to be adequate test signal for all defects in S_{P2}.

As can be seen from Table I all shorts between analog switch terminals, except gate-balk combination, result with malfunction of the circuit. So these defects could be detected only by observation of the waveform at the quantizer output. However, for gate-bulk short and drain, source and gate opens, the functionality of the circuit apparently stays unchanged. Because of that, presence of these defects cannot be detected by purely observing waveform of output signal.

TABLE I
SUMMARY OF RESULTS

			THD [%]	SNR [dB]	detected
Fault Free Circuit			0.067	78.53	-
Circuit with faults in S_{P2}	shorts	DS	98.742	-2.81	Yes
		GD	98.742	-2.81	Yes
		GS	98.741	-2.82	Yes
		DB	98.742	-2.81	Yes
		GB	5.11	57.9	Yes
		SB	98.742	-2.81	Yes
	opens	D	0.0204	86.13	Yes
		S	0.016	86.91	Yes
		G	0.028	85.79	Yes

So here is where proposed procedure pays its worth. For gate-bulk short there is a decrease of about 11dB in *SNR* and approximately 76 times increase in *THD* percentage.

For opens the effect is opposite as there are increases in *SNR* of about 8dB and approximately 3 times decreases in *THD* percentage.
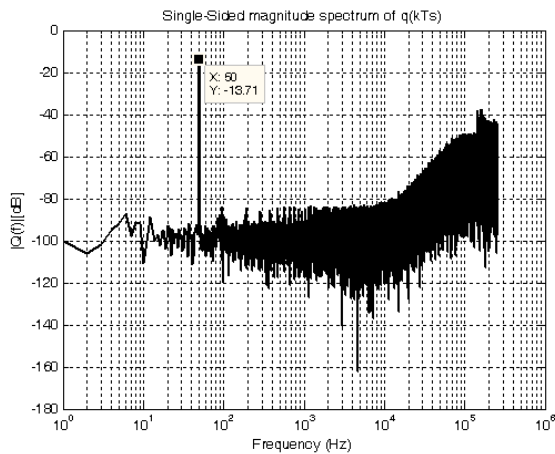
Fig. 4. Single-sided magnitude spectrum of fully functional circuit
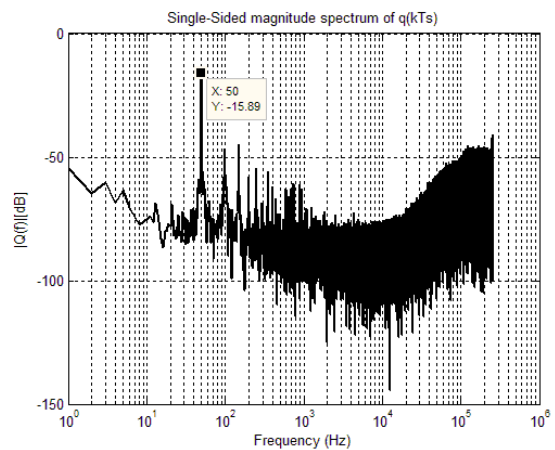
Fig. 5. Single-sided magnitude spectrum with shorted gate and bulk terminals of S_{P2} analog switch

As mentioned earlier, graphical representation of observed signal spectrum can be useful for defects detection. Fig. 4 illustrates single-sided magnitude spectrum of a quantizer output signal when there is no defect in circuit. From Fig. 4, single tone at 50Hz with -13.71dB magnitude can be noted. In band (below 2 kHz) noise floor is at about -100dB while out of band noise (over 2 kHz) is shaped with approximately 40dB/dec slope which corresponds to a second order modulator. For the sake of illustration on Fig. 5 single-sided spectrum of quantizer output whit presence of the S_{P2} gate-bulk (GB) defect in the circuit is shown. As one can notice from Fig. 5 the magnitude spectrum, along with *SNR* and *THD* parameters (see Table I), deviates from the spectrum of the fault-free circuit shown on Fig. 4.

VI. CONCLUSION

Review of standard test methods and principles which apply to mixed-signal data converter circuits is given. Some of the most popular ADC testing techniques are covered. Suitable parameters for functional verification of $\Sigma\Delta$ type of data converters are explained and adopted.

The architecture and functionality of the implemented DFT structure for integrated power meter $\Sigma\Delta$ ADC is commented. Appropriate test procedure for testing the analog part of $\Sigma\Delta$ ADC is developed and described along with the basic operation of the tested circuit. The test procedure is confirmed by means of SPICE simulation. Application of the proposed test procedure is presented through an example. Finally, the obtained simulation results are presented and commented.

ACKNOWLEDGEMENT

REFERENCES

[1] http://leda.elfak.ni.ac.rs/projects/IMPEG/impeg.htm
[2] R. Schreier, and G. C. Temes., "*Understanding Delta-Sigma Data Converters*", John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.
[3] De Venuto D., "*Testing high resolution SD ADC's by using the quantizer input as test access*", Microelectronics Journal, Vol. 36, No. 9, 2005, pp. 810–819.
[4] Nikolić, M., Sokolović, P., and Petković, P., "*Laboratory ADC Tester Based on NI-6251 Acquisition Card*", Proc. of 25th MIEL, Belgrade, 14-17 May, 2006
[5] M. Burns, and G. W. Roberts, "*An Introduction to Mixed-Signal IC Test and Measurement*", Oxford University Press, New York, 2001.
[6] Sokolović, M., Savić, M., Nikolić, M., Litovski, V., Jevtić, M., and Petković, P., "*Testing and Diagnostic of ADC and Integrated Powermeter*", Electronics, Vol. 9, No. 1, 2005, ISSN 1450-5843.

High level simulation of multiplexed incremental ADC for Integrated Power Meter

Dejan Mirković and Predrag Petković

Abstract – This paper presents an architectural solution for multiplexed ADC designed for a new generation of integrated power meter, IMPEG3. Basic problems related to the use of classic ΣΔ ADCs will be discussed. Proposed solution along with appropriate procedure for determining sampling frequency will be explained. High level simulation confirmed the developed behavioral model of the multiplexed ADC.

Keywords – Multiplexed ΣΔ ADC, behavioral modeling, integrated power meter.

I. INTRODUCTION

Contemporary power meters are based on integrated circuits dedicated to calculate energy using samples of current and voltage. Accuracy of these calculations is mainly determined by the quality and resolution of its analog frontend, namely analog-to-digital converter (ADC). The design team of LEDA laboratory at Faculty of Electronic Engineering, University of Niš continuously develops a series of own ASIC for electric power measuring applications. Fig. 1 shows the block diagram of the first version of a single-phase power meters (named IMPEG1) that has been prototyped in 2005. The analog frontend has been design as classic ΣΔ ADC.
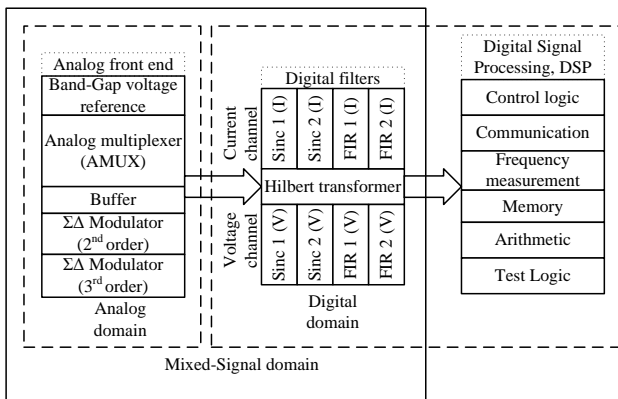
Fig. 1: Block diagram of IMPEG1 circuit

The prime purpose of ADC is to convert samples of current and voltage to digital domain. Conversion takes place in two separate channels for current and for voltage. The analog part of ADC consists of ΣΔ modulator [1].

Dejan Mirković and Predrag Petković are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {dejan, predrag}@elfak.ni.ac.rs.

Modulator serves to *shape* the noise generated during analog-to-digital signal discreatization. The nature of ΣΔ modulator is to operate with oversampled high frequency but with few number of output bits. IMPEG1 operates as single bit at output channel. Practically, the carrier of 50Hz is oversampled with 524288 Hz. The modulator suppresses HF noise out of the signal base band that has been limited to 2kHz. This implies noise rejection to high frequencies while the signal is processed with a low pass frequency (LP) filter. Integrators are used for LP filtering. The order of LP filters depends on needed ADC parameters. Signal to noise Ratio (SNR) and SFDR Spurious Free Dynamic Range (SFDR) are the most important input parameters in ADC design. They characterize the required resolution in terms of the number of output bits [1]. Initially the third order filter has been used for current and second order filter for voltage channel. However measurements on the prototyped samples have shown that the second order filtering is sufficient for both channels. In order to obtain multi-bit output one needs decimation digital filter. It accumulates the output signal with the oversampling frequency and provides wider digital word at the output at lower frequency. This word is a digital representation of instantaneous values of voltage and current, respectively.

The digital data is further fed into the Digital Signal Processing (DSP) block. DSP calculates RMS values of current and voltage, active power, reactive power, apparent power, corresponding energies, power factor (displacement factor) and frequency. This single-phase version of IMPEG chip is prototyped in CMOS 0.35 μm AMI Semiconductor technology.

The next generation of IMPEG solid-state power meters, named IMPEG2, was enhanced to fit for three-phase systems. Analog part of ADC was simply tripled for all three phases per channel. Therefore IMPEG2 was designed to have six ΣΔ modulators (three for currents and three for voltages of every phase). The digital part of the ADC has been redesign in a quite different manner comparing to IMPEG1. An original solution published in [2] describes the unique digital filter for processing of all three phases. Besides, IMPEG2 has been enhanced by other features mainly related to the digital part. The most important are embedded MCU8052, drivers for display and UART ports. More about these previous versions can be found at [3].

The subsequent generation of IMPEG required improvements in analog part.

The idea was to make compact ADC without lack of functionality. Consequently, an implicit solution was to

utilize multiplexed input that uses the same hardware. Some possible results for IMPEG3 were published in [4]. All of them considered separated current and voltage channel. This approach was motivated by different dynamic ranges and noise sensitivity. Fig. 2 depicts block diagram of the proposed ADC architecture.
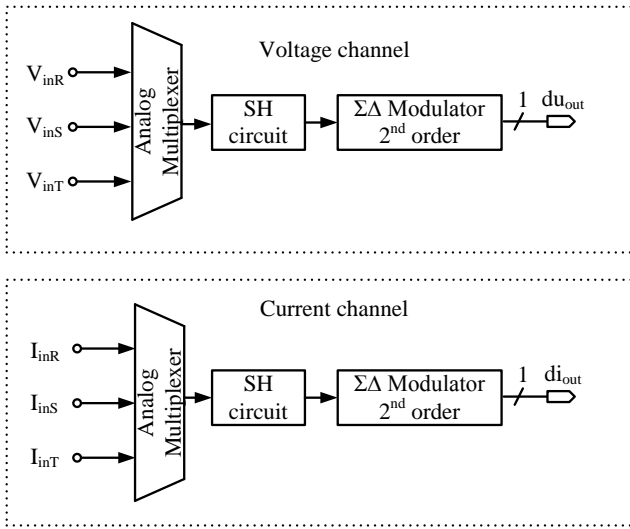

Fig. 2: Block diagram ADC in IMPEG3

In Fig. 2, V_{inR}, V_{inS} and V_{inT} represent analog line voltages while I_{inR}, I_{inS} and I_{inT} represent voltage equivalents of analog line currents (obtained with e.g. shunt resistors). Signals di_{out} and du_{out} denote one-bit ΣΔ modulator's outputs for voltage and current channels, respectively. Concept illustrated in Fig. 2 can be expanded to four inputs where the fourth port may be used for zero line current or for an external temperature sensor. The new version of IMPEG that is the subject of this paper will be designed to accept eight analog signals. This paper considers a new concept for multiplexed analog inputs. It relays on experiences obtained with the previous versions. Therefore, they will be shortly explained before we suggest a new architecture.

The next section describes different ADC architectures starting from ΣΔ modulator that has been used in previous IMPEG version. Thereafter the new architecture will be proposed together with procedure for determining fundamental operational parameter, namely sampling frequency, will be presented. It concludes with the subsection that elaborate stability of the proposed ΣΔ modulator architecture. The third and forth section describe the behavioral model and the corresponding results of simulation for the new ADC, respectively. The final section reviews conclusion remarks.

II. Architectures of ΣΔ modulator

A. Previous architecture of ΣΔ modulator

Architecture of modulators built-in IMPEG1-3 ADC is

CIFB (Cascade of Integrators Feed-Back) [1]. Fig. 3 shows block diagram of the second order ΣΔ modulator with CIFB architecture.
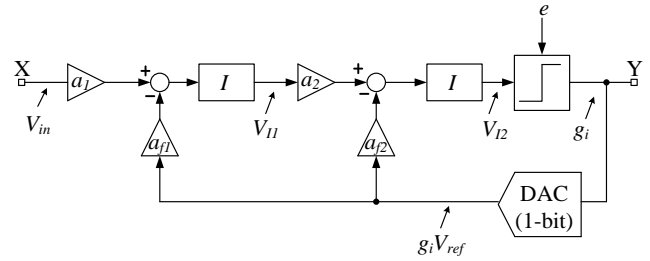
Fig. 3: Block diagram of 2nd order ΣΔ modulator with CIFB architecture

Blocks denoted with I represent DAI (Discrete Analog Integrators) with delay whose Z domain transfer function is given as:

$$I = \frac{z^{-1}}{1 - z^{-1}},$$ (1)

where z is a complex variable. Input voltage is denoted with V_{in}, while V_{I1} and V_{I2}, represent the output of the first and the second integrator, respectively. Output of single-bit quantizer after i^{th} step is denoted with g_i. It can take values from set $\{1, -1\}$. Constants a_1, a_2, a_{f1} and a_{f2} are gains in direct path and feedback loops. The modulator is realized as Switched Capacitor (SC) circuit. More about architecture and realization of modulator shown on Fig. 3 can be found in [5].

Multiplexed implementation implies constant input signal during conversion of each channel. Therefore, the classic ΣΔ modulator architecture is not suitable for multiplexed application. Consequently, it is necessary to introduce a new architectural solution for ΣΔ modulator. In other words, classic ΣΔ type of ADC architecture is not favorable in combination with multiplexer on its input. The main problem with classic ΣΔ converter is too long settling time of digital part, i.e. digital filters that cause unacceptable delay between multiplexed channels. Therefore the architecture published in [6] is more suitable for multiplexing. This architecture is known as CIFF (Cascade of Integrators Feed-Forward) [7]. In order to be ready for accepting data from the new channel, ADC has to be reset after every conversion cycle. This implies that reset mast be introduced in ADC system. This kind of oversampling ADCs which contains resettable ΣΔ modulators are usually named *charge transfer*, *single shot* or *incremental* ADC [7].

B. Proposed architecture of ΣΔ modulator

Fig. 4 represents CIFF architecture of a second order ΣΔ modulator. Obviously, this architecture has only one feedback path from output (Y) to input (X), in contrast to CIFB architecture (Fig. 3) that has two. According to Fig. 3, blocks denoted with I represent DAI with Z domain transfer function given in (1).
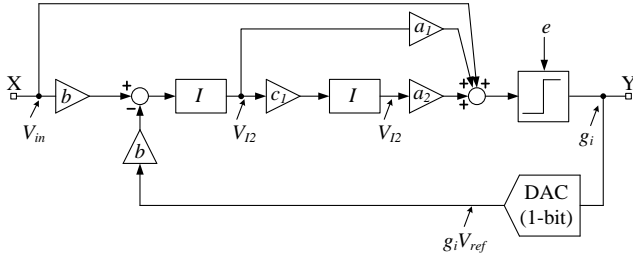
Fig. 4: Block diagram of 2nd order ΣΔ modulator with CIFF architecture

In Fig. 4 V_{in}, V_{I1} and V_{I2}, denotes input voltage and output voltage of the first and the second integrator, respectively. The output voltage of the single-bit DAC, denoted by term $g_i V_{ref}$, takes values $\pm V_{ref}$. Constants a_1, a_2, c_1 and b are modulator coefficients.

The modulator operates as follows.

Let suppose that the conversion of one input value requires n clock cycles. Each conversion starts from the initial state with $V_{I1}[0] = V_{I2}[0] = 0V$. The subsequent n outputs at the first integrator will take the following values:

$$V_{I1}[0] = 0V$$
$$V_{I1}[1] = b(V_{in}[0] - g_0 V_{ref}) + V_{I1}[0]$$
$$V_{I1}[2] = b(V_{in}[1] - g_1 V_{ref}) + V_{I1}[1]$$
$$V_{I1}[2] = b(V_{in}[0] + V[1] - (g_{10} + g_1)V_{ref}) , \quad (2)$$
$$\vdots$$
$$V_{I1}[n] = b\sum_{i=0}^{n-1}(V_{in}[i] - g_i V_{ref})$$

where $V_{in}[i]$ and $V_{I1}[i]$ stand for input analog voltage and the output voltage of the first integrator after i clock cycles, respectively while g_i denotes state of the quantizer after i^{th} cycle. Similarly, the output of the second integrator will be:

$$V_{I2}[0] = 0V$$
$$V_{I2}[1] = c_1 V_{I1}[0] + V_{I1}[0]$$
$$V_{I2}[2] = c_1 V_{I1}[1] + V_{I2}[1] = c_1(V_{I1}[0] + V_{I2}[1])$$
$$V_{I2}[3] = c_1 V_{I1}[2] + V_{I2}[2] = \quad ,(3)$$
$$c_1(V_{I1}[0] + V_{I1}[1] + V_{I2}[2])$$
$$\vdots$$
$$V_{I2}[n] = c_1 \sum_{j=0}^{n-1} V_{I1}[j] = c_1 b \sum_{j=0}^{n-1}\sum_{i=0}^{j-1}(V_{in}[i] - g_i V_{ref})$$

where $V_{I2}[j]$ is output voltage of the second integrator after j^{th} clock cycle. This voltage takes value in range $\pm V_{ref}$.

$$-V_{ref} < V_{I2}[j] < +V_{ref} . \quad (4)$$

During the conversion, SH circuit provides constant analog input signal $V_{in}[j] = V_{in0}$, for $j=1,...,n$. Therefore its contribution to the total sum in (4) is:

$$\sum_{j=0}^{n-1}\sum_{i=0}^{j-1} V_{in}[i] = \sum_{j=0}^{n-1}\sum_{i=0}^{j-1} V_{in0} = \binom{n}{2} V_{in0} = \frac{n(n-1)}{2!} V_{in0} . \quad (5)$$

Substitution of (3) and (5) into (4) provides important result given in (6).

$$-\frac{V_{ref}}{c_1 b}\frac{2!}{n(n-1)} < V_{in0} - \sum_{j=0}^{n-1}\sum_{i=0}^{j-1} g_i V_{ref} < +\frac{V_{ref}}{c_1 b}\frac{2!}{n(n-1)} . \quad (6)$$

The middle term of (6) represents the difference between analog input and the converted digital signal. Practically, it stands for conversion error that is bounded with $\pm (1/2) \cdot V_{LSB}$, where V_{LSB} is the analog voltage equivalent of the least significant bit. Therefore, it is defined as:

$$V_{LSB} = \frac{V_{ref}}{c_1 b}\frac{2 \cdot 2!}{n(n-1)} . \quad (7)$$

Resolution of the converter can be expressed as:

$$n_{bit} = \log_2\left(\frac{V_{in\,max}}{V_{LSB}}\right) =$$
$$\log_2(n(n-1)) + \log_2(c_1 b) + \quad , \quad (8)$$
$$\log_2\left(\frac{V_{in\,max}}{V_{ref}}\right) - 2$$

where, n_{bit} is a number of bits. Assuming $n \gg 1$ one gets:

$$n_{bit} \approx 2 \cdot \log_2 n + \log_2(c_1 b) + \log_2\left(\frac{V_{in\,max}}{V_{ref}}\right) - 2 . \quad (9)$$

For the full dynamic range of the input signal of $V_{inmax} = +V_{ref} - (-V_{ref}) = 2 \cdot V_{ref}$, and for normalized architecture (all modulator's coefficients equal to one) the resolution is reduced to:

$$n_{bit} \approx 2 \cdot \log_2 n - 1 . \quad (10)$$

Expressions given in (9) or (10) allow calculation of the minimum number of clock cycles, n, needed to obtain n_{bit} resolution. Supposing that the required dynamic range is 96bB (knowing that each bit gives 6dB) it is easy to find that $n_{bit} = 96/6 = 16$ will met the request. According to (10) one easily calculates the minimum number of clock cycles to be $n_{min} = 362$. This means that the conversion of one channel lasts for 362 clock intervals. According to [7], in order to prevent saturation of integrators it is good practice to adopt $V_{inmax} \sim 0.67 V_{ref}$ for second order modulator. Consequently, this increases the minimum number of clock cycles to $n_{min} = 512$. More information about evaluating n_{min} for higher order modulators can be found in [7].

So far the architecture and operation condition of the modulator has been determined. The next step is to define decimation filter that should average n_{min} bit stream to provide n_{bit} digital output. It is good practice to choose

filter to be at least one order higher than the modulator. Therefore this paper will consider a third order *Sinc* filter.

In general, transfer function of the Lth order *Sinc* filter expressed in z-domain is:

$$H(z) = \frac{1}{M^L} \left(\frac{1 + z^{-M}}{1 - z^{-1}} \right)^L, \qquad (11)$$

where $M = f_s/f_n$ is the oversampling ratio, f_n is Nyquist frequency defined as $f_n = 2 \cdot BW$ and BW represents signal base band-width. In the particular case of IMPEG $BW = 2048$Hz. Realization of the *Sinc* filter is simplified if $M = 2^K$, where K is an integer.

According to [7] procedure for determining sampling frequency should be:

1. For given n_{min} ($n_{min} = 512$) and adopted order of the modulator *La* (*La* = 2) determine the order of digital *Sinc* filter $L = La + 1$ ($L = 3$).
2. Increase n_{min} until obtain $M = n_{min}/L$ to be the first larger 2^K number. (M= $n_{min}/L = 512/3 = 170.667$; the first larger 2^K is 256; choose $M = 256$ and recalculate $n_{min} = L \cdot M = 768$).
3. Determine sampling frequency as $f_s = M \cdot f_n = M \cdot (2 \cdot BW)$ ($f_s = 2^{20}$ Hz ≈ 2 MHz).

After n_{min} clock cycles filter will provide 16-bit digital word at the output. Therefore, one conversion cycle requires $n_{min}/f_s \approx 732.42$μs. The existing $\Sigma\Delta$ ADC in previous IMPEG versions provides digital data three times faster with 4096Hz rate, i.e. at every ~ 214.14μs. In order to maintain the same data rate towards digital part of the chip, a new three times higher sampling frequency is adopted ($f_{snew} = 3 \cdot 2^{20}$ Hz ≈ 3.14 MHz). Eventually, the output of the digital filter provides 16-bit wide word after $n_{min}/f_{snew} = 1/4096$s ~ 214.14μs.

Let consider a multiplexed ADC with *N* input signals (channels) based on architecture from Fig. 4. According to the fact that n_{min} depends on the required resolution (which was not the case with classic $\Sigma\Delta$ ADC), the multiplexing with the specified precision is feasible if $N \cdot n_{min}$ clock periods which fit within the available time window. In a case of IMPEG3, goal is to obtain $N = 4$ conversions within the same time window of 214.14μs. This means that the 16-bit output word appears for all four channels with the rate of 4096Hz. Therefore the conversion time should be four times shorter for each channel and consequently, the new sampling frequency has to be $N = 4$ times higher ($f_{sN} = N \cdot f_s$). After n_{min} clock periods, the obtained output of one channel is buffered, the whole converter (modulator and digital filter) is reset and another channel is fed to the input of ADC. This cycle repeats for all of *N* input channels. When N^{th} channel is converted, the buffered digital words are fed to DSP unit in parallel. The algorithm of conversion for *N*-channel multiplexed, incremental ADC contains the following steps:

1. First channel is selected with analog multiplexer.
2. With SH circuit selected signal is sampled and held for whole conversion cycle at the ADC's input.
3. After n_{min}/f_{sN} (1/2^{14} s ≈ 61μs) digital output word is obtained and buffered for the selected channel; ADC is reset; and the subsequent new channel is selected from analog multiplexer to the ADC input.
4. Step 3 is repeated *N* times for all input channels.
5. After conversion of Nth channel; N n_{bit}-bit words are red with rate $f_{sN}/(n_{min} \cdot N)$ ($f_{sN}/(n_{min} \cdot N) = 4096$ Hz); the procedure repeats from step 1.

One could conclude that due to the multiplexing a delay of $n_{min}/f_{sN} = 61$μs appears between each channel that could reflect on signal phase. Because the signal frequency is 50Hz, i.e. period 20ms, the delay is only about 0.3% of the signal's period. Therefore, the phase error is negligible. Nevertheless it is necessary to examine its influence on overall power and energy calculation. Obviously it is a systematic error that can be eliminated later in the digital part of the chip by the built-in programmable compensation.

C. Stability of proposed $\Sigma\Delta$ modulator architecture

So far the stability has not been considered. However it is an important feature of modulators. According to the common practice the stability will be appraised by analyzing the modulator behavior in *z*-domain. In $\Sigma\Delta$ convertors noise signal is fed back to the input and consequently it is responsible for stability. Therefore it is of important interest to analyze distributions of poles in Noise Transfer Function (*NTF*). *NTF* is defined as *Y/e* ratio and according to Fig. 4 the *NTF* can be written as:

$$NTF = \frac{z^2 - 2z + 1}{(z - p_1)(z - p_2)}, \qquad (12)$$

where poles of the transfer function, p_1 and p_2, are related with modulator's coefficients as:

$$\begin{aligned} p_1 + p_2 &= 2 - a_1 b \\ p_1 p_2 &= 1 + a_2 c_1 b - c_1 b \end{aligned}. \qquad (13)$$

Obviously, the second order filter gives second order *NTF*. Modulator will be stabile if pair of *NTF* poles, p_1 and p_2, are placed inside the unit circle in z-plane. As mentioned earlier, for a second order modulator, it is good practice to limit integrator output voltages on $0.67V_{ref}$. The design task is to determine coefficient values that provide stabile and reliable operation. For adopted $M = 256$, $V_{Imax} = 0.67V_{ref}$ and maximum *NTF* out of band gain of 2, MATLAB® *Delta-Sigma Toolbox* [8] a value of $p_{1/2} = 0.3819 \pm j \cdot 0.3004$. Mapping these poles on CIFF architecture resulted in following coefficients: $b = 0.475$, $c_1 = 0.598$, $a_1 = 2.59$, $a_2 = 2.755$.
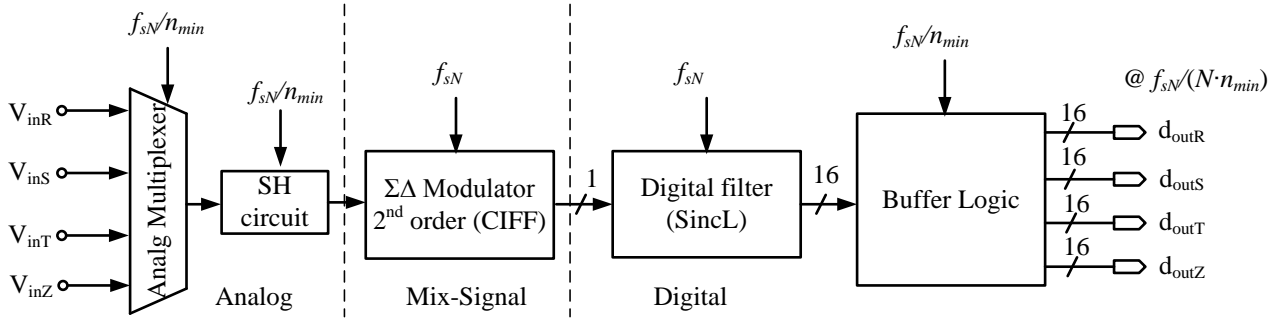
Fig. 5: Data flow block diagram of voltage channel ADC

Now, when all relevant design parameters of ΣΔ modulator and digital filter (f_s, M, L, La, NTF ...) are explained, behavioral model of ADC can be considered.

III. MODELING OF ΣΔ MODULATOR

As Fig. 5 indicates, ADC is a mix-signal circuit. It can be roughly divided in three sub sections namely analog, mix-signal and digital. Analog block contain analog multiplexer and SH circuit. In this block signals are conditioned in analog domain i.e. can take any real value. Through the loop filters (integrators) signal is conditioned in the same manner i.e. in analog domain, while amount of the feedback is controlled digitally by quantizer and one-bit DAC outputs. Quantizer output is the point of connection between mix-signal and purely digital world. From this point signal is digitally processed and its value is constrained with two logic levels ("0" and "1").

Taking the previous in to account, behavioral model of proposed ADC architecture is developed. Model is built combining MATLAB®'s *Simulink* environment and appropriate scripts. The reset signal requires the use of time enabled and triggered *Simulink* blocks. These blocks provide a solid behavioral description of clock and reset dependent discrete system components. Therefore, DAI and digital filter/buffer logic registers are described in this manner.

All previously discussed characteristics of the proposed ADC architecture are built-in behavioral model. Obtained simulation results will be presented and commented in the subsequent sections.

IV. SIMULATION RESULTS

The architecture of four inputs multiplexed, incremental ADC was verified on developed behavioral model. Sampling frequency is f_{sN} =12MHz (12582912Hz), while analog multiplexer, SH circuit and reset are clocked with f_{sN}/n_{min} =16384Hz (see Fig. 5). ADC is verified in presence of four sinusoidal signals with the same frequency of 50Hz and with amplitudes of 200mV, 100mV, 50mV and 25mV at inputs V_{inR}, V_{inS}, V_{inT} and V_{inZ}, respectively. Phases of V_{inR}, V_{inS}, and V_{inT} are shifted for for 120° while V_{inZ} have the same phase as V_{inR}. Fig. 6 illustrates waveforms of

input and the output signals.
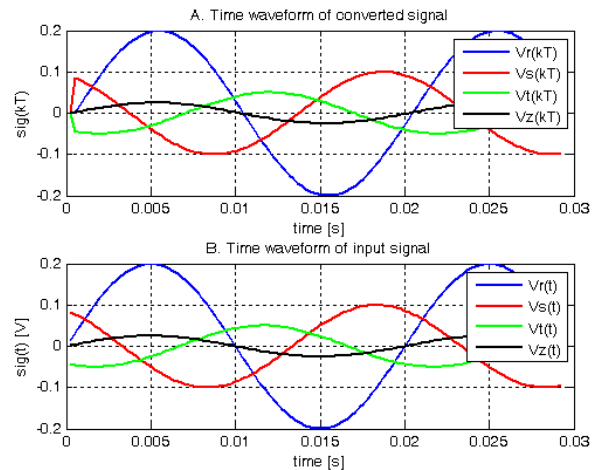
Fig. 6: Waveforms of input (A) and converted (B) signals of incremental ADC in voltage channel

The best insight into circuit behavior gives signal spectrum obtained by FFT analyzes over ADC's outputs. Fig. 7 illustrates the obtained results. Figure 7.a presents the complete, single-sided, signal spectrum up to the base band of 2kHz. Figure 7.b depicts magnified part of the spectra around 50Hz in order to clearly verify presence of all four signals. Obviously the obtained result confirms functionality of the proposed multiplexed, incremental, ADC architecture.

It should be mentioned that the current channel has identical architecture. Experience with previously prototyped IMPEG version suggests that the obtained SFDR better than 130dB at behavioral level have good chances to satisfy requested dynamic range of 80dB in current channel after fabrication.

V. CONCLUSION

One architectural solution for ADC that is going to be built-in integrated power meter, IMPEG3, suitable for multiplexed applications was presented in this paper. Review of the previous versions of IMPEG was given as well. The paper discussed some basic drawbacks of classic ΣΔ ADCs in multiplexing conditions. The central part of
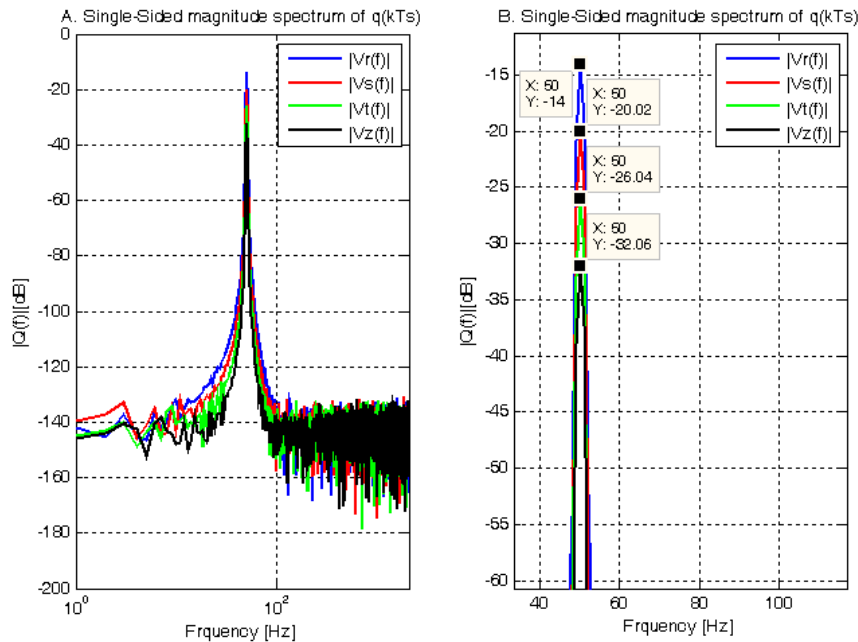
Fig. 7: One side magnitude spectrum of four channel ADC output (A) Full spectrum, logarithmic frequency scale;
(B) Part of spectrum around 50Hz, linear frequency scale

this work proposed the new architecture with detailed derivation of mathematic expressions needed for behavioral modeling. In addition, the attention was paid to the appropriate algorithm for determining sampling frequency and to the stability criterion.

Basic modeling approaches are given. The developed behavioral model of the proposed architecture was confirmed by simulation using *Simulink*. The obtained simulation results are presented and commented.

Further research should consider other $\Sigma\Delta$ modulator and/or digital filter architectures. It has been shown in [7] that good results can be obtained with simpler digital filter architectures (e.g. cascade of two digital integrators), as well. Therefore, the expectation of better matched ADCs and reduced area comes through. Every, new, architecture should be well examined for design requirements in terms of dynamic range, resolution and noise sensitivity.

ACKNOWLEDGEMENT

REFERENCES

[1] R. Schreier, and G. C. Temes., ”*Understanding Delta-Sigma Data Converters*”, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.

[2] Marinković, M., „*Decimacioni filtri u trofaznom integrisanom meraču potrošnje električne energije*“, Magistarska teza, Univerzitet u Nišu, Elektronksi fakultet Niš, 2008. Ovde bolje da stavimo nešto na engleskom.ili

[2] Marinković, M., „*Decimation filters in tree-phase integrated power meters*”, Master Thesis, University of Niš, Faculty of Electronic Engineering, Nič, 2008. (In Serbian)

[3] http://leda.elfak.ni.ac.rs/projects/IMPEG/impeg.htm

[4] Mirković, D., Petković, P.: „*Multi channel Sigma-Delta A/D converter for integrated power meter*“, Proceedings of the Small Systems Simulation Symposium 2010, Niš, 12-14 February, 2010, pp. 90-93, ISBN 987-86-6125-006-4

[5] Nikolić, M., „Layout design of Mixed Signal CMOS Integrated Circuits“, Master Thesis, University of Niš, Faculty of Electronic Engineering, Nič, 2006. (In Serbian)

[6] C. Lyden, C. A. Ugarte, J. Kornblum, and F. M. Yung, “*Single shot sigma-delta analog-to-digital converter for multiplexed applications*” in *Proc. IEEE Custom Integrated Circuit Conf.*, Santa Clara, CA, May 1-4, 1995, pp. 203-206.

[7] J. Márkus, J. Silva, and G. C. Temes, “*Theory and applications of incremental delta-sigma converters*”, IEEE Trans. Circuit Syst. I, Reg. Papers, vol. 51, no. 4, pp 678-690, Apr. 2004.

[8] R. Schreier. (2003) *The Delta–Sigma toolbox v6.0 (Delsig)*. Mathworks, Natick, MA. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange

Privacy Issues in Smart Grids

Slobodan Bojanić, Srdan Đorđević and Octavio Nieto-Taladriz

Abstract - The smart grid brings an entirely new and complex model of inter-relationships which poses challenges for data privacy. It is an emerging area where new data privacy problems evolve as more smart meters are installed. Since mass rollout of smart meters is already happening, there is urgency how to process personal data and to treat some issues of general concern which warrant serious privacy consideration.

Keywords – Smart Grid, Privacy, Smart Meter, Privacy by Design.

I. INTRODUCTION

The benefits of smart energy use include opportunities for consumers to cut their bills by changing their habits, perhaps using energy at different times to take advantage of lower tariffs, as well as opportunities for industry to more accurately forecast demand, reducing expensive electricity storage costs. The realisation of climate change targets relies to some extent on consumers releasing personal data, but this needs to be achieved in such a way that all parties involved in programmes to introduce smart meters and the development of the smart grid ensure that the fundamental rights of individuals are protected and respected [1]-[5].

Without such protection there is a risk not only that processing of personal data will be in breach of national laws but also that consumers will reject these programmes on the basis that the collection of personal data is unacceptable to them. Such rejection may arise even if there is no breach of the law. While the potential benefits of these programmes are far-reaching and significant, they also have the potential to process increasing amounts of personal data, unprecedented in this industry, and to make that personal data more readily available to a wider circle of recipients than at present.

II. BACKGROUND

A variety of definitions of smart grid have been available. It can be assumed that the Smart grid is an intelligent electricity network that combines information from users of that grid in order to plan the supply of electricity more effectively and economically that was possible in the pre-smart environment. It can cost

Slobodan Bojanić and Octavio Nieto-Taladriz are with Universidad Politécnica de Madrid, ETSIT Avenida Complutense nº 30, 28040 Madrid, Spain, e-mail: {slobodan, octavio}@die.upm.es.

Srdan Đorđević is with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: srdjan.djordjevic@elfak.ni.ac.rs.

efficiently integrate the behaviour and actions of all users connected to it – generators, consumers and those that do both – in order to ensure economically efficient, sustainable power system with low losses and high levels of quality and security of supply and safety. In a further step, energy optimization crossing the domains of electricity, gas and heat will be a further challenge.

The electricity networks have provided the vital links between electricity producers and consumers with great success for many decades. The fundamental architecture of these networks has been developed in most countries to meet the needs of large, predominantly carbon-based generation technologies. Europe is committed to the 20-20-20 targets to reduce carbon emissions and to secure energy supply. Energy efficiency and renewable energy are seen as key to reach this goal. Both measures call for changes in the energy supply system leading to smart grids as key enablers for the required innovation.

Smart meters allow for the generation, transmission and analysis of data relating to consumers, much more than is possible with a 'traditional' or 'dumb' meter. Consequently, they also allow the network operator (also known as Distribution Service Operator or DSO), energy suppliers and other parties to compile detailed information about energy consumption and patterns of use as well as make decisions about individual consumers based on usage profiles. Whilst it is acknowledged that such decisions can often be to the benefit of consumers in terms of energy savings, it is also emerging that that there is potential for intrusion into the private lives of citizens through the use of devices which are installed in homes. It also marks a shift in our fundamental relationship with energy suppliers in that consumers have traditionally simply paid suppliers for the electricity and gas that has been supplied.
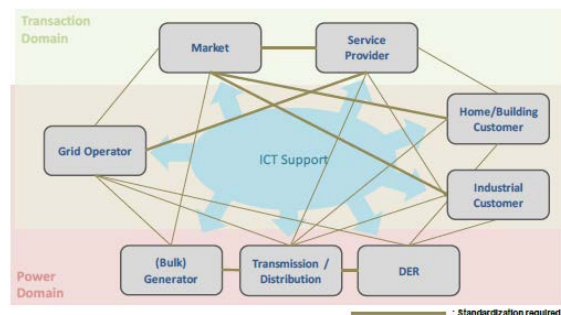

Fig. 1 A conceptual model of the Smart Grid

With the advent of smart meters, the process is more complex in that the data subject will provide suppliers with insights into personal routines. There is a huge variation in

circumstances between countries, ranging from those where rollout is largely complete following government mandate to those where no meters have been installed.

II. STAKEHOLDERS

As there is no one universal, internationally accepted definition of "privacy," it can mean many things to different people. At its most basic, privacy can be seen as the right to be left alone. Privacy is not a plainly delineated concept and is not simply the specifications provided within laws and regulations. Furthermore, privacy should not be confused, as it often is, with being the same as confidentiality; and personal information is not the same as confidential information. Confidential information is information for which access should be limited to only those with a business need to know and that could result in compromise to a system, data, application, or other business function if inappropriately shared.

The smart metering brings with it the potential for numerous novel ways for processing data and delivering services to consumers. Whatever the processing, whether it is similar to that which existed in the pre-smart environment, or unprecedented, the data controller must be clearly identified, and be clear about obligations arising from data protection legislation including Privacy by Design, security and the rights of the data subject. Data subjects must be properly informed about how their data is being processed, and be aware of the fundamental differences in the way that their data is being processed so that when they give their consent it is valid.

The following Smart Grid stakeholders
• Grid users including/composed of grid operators, grid customers and meter operators
  • End customer (domestic or commercial)
  • Municipalities including energy retailers
  • Politics
  • Industries
  • Consumer organizations
  • Politics/society
can also be viewed through various domains interconnected
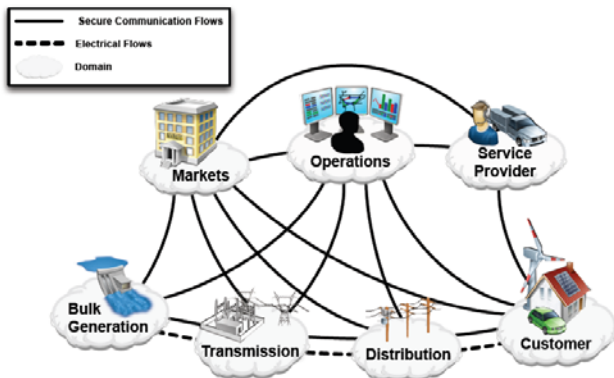
Fig. 2 Interaction among actors in Smart Grid

by secure communication flows and flows of electricity as presented in Fig. 2.

Basicaly the smart meter takes a reading which reflects the energy usage at the property. At some point that reading, along with other information, can be transmitted outside the property. In some models it will be sent directly to a central communications hub where the smart meter data are managed. Once there, it can be accessed by DSOs, suppliers and ESCOs. It appears that the DSOs will have to face the greatest changes to make smart grids a reality. The reasons for that are the growing distributed character (resulting in growing bidirectional power flow at all voltage levels) and variability of generation, customer privacy issues, system security, data and information processing for new applications and concepts such as Virtual Power Plants, etc.

There are also multiple and complex methods of communication, with additional entry points and data paths creating complicated security challenges requiring solutions that encompass them all. Given the complex and disparate landscape, the task of producing privacy solutions is quite challenging, and at this stage it seems that they can only be general, rather than specific.

The disparity of the current position does not allow presenting a comprehensive view on all specific aspects of smart metering programmes across member states. There is a huge variation in circumstances between countries, ranging from those where rollout is largely complete following government mandate to those where no meters have been installed. There is also much variation in the level of involvement from DPAs and in the nature of the market across member states, and where responsibility lies with installation of meters. In some countries, publicly owned utility companies are responsible. Elsewhere, there is a competitive supplier market. Distribution system operators have a more prominent role in some countries.

As the smart grid brings with it an entirely new and complex model of inter-relationships that poses challenges for the application of data protection. This is an emerging area of work where it is fully expected that new data protection problems and solutions will evolve as more smart meters and smart grid components are installed. What is inarguable is that mass rollout of smart meters is already happening, so there is urgency to collectively understand the way that smart meters process personal data, and the issues that this raises. There are some issues of general concern which warrant serious consideration by all those involved in this area.

Given that that data in Smart Grids might contain privacy sensitive information it is advised that principles such as privacy by design and Default should be encouraged. The personal data is being processed by the meters, so data protection laws apply. The smart metering brings with it the potential for numerous novel ways for processing data and delivering services to consumers. Whatever the processing, whether it is similar to that which existed in the pre-smart environment, or unprecedented, the
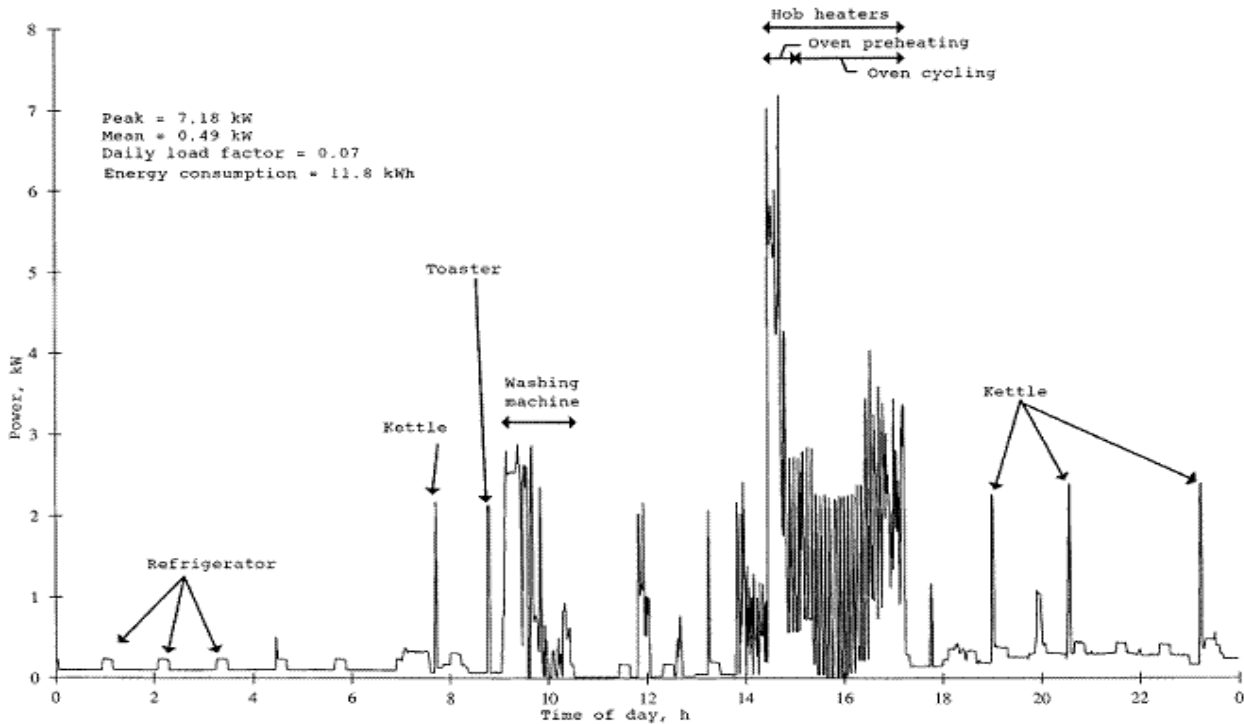
Fig. 3 Consumer profiling by energy

data controller must be clearly identified, and be clear about obligations arising from data protection legislation, security and the rights of the data subject. Data subjects must be properly informed about how their data is being processed, and be aware of the fundamental differences in the way that their data is being processed so that when they give their consent it is valid.

III. PRIVACY THREATS

There are numerous privacy implications identified for smart grid technology deployment centers on the collection, retention, sharing, or reuse of electricity consumption information on individuals, homes, or offices. Fundamentally, smart grid systems will be multi-directional communications and energy transfer networks that enable electricity service providers, consumers, or third party energy management assistance programs to access consumption data. Further, if plans for national or transnational electric utility smart grid systems proceed as currently proposed these far reaching networks will enable data collection and sharing across platforms and great distances [7]-[11].

Consumer privacy is a key aspect in the change towards smart energy systems thus data access and ownership and the permission to gather data need to be very carefully considered. At the same time, consumers should be well-informed about who deals with their data. It has to be remembered that it is the consumer who owns his data, no-one else, and therefore he is entitled to appropriate rights and protections.

A list of potential privacy consequences of Smart Grid systems include:
- Identity Theft
- Determine Personal Behavior Patterns
- Determine Specific Appliances Used
- Perform Real-Time Surveillance
- Reveal Activities Through Residual Data
- Targeted Home Invasions (latch key children, elderly, etc.)
- Provide Accidental Invasions
- Activity Censorship
- Decisions and Actions Based Upon Inaccurate Data
- Profiling
- Unwanted Publicity and Embarrassment
- Tracking Behavior of Renters/Leasers
- Behavior Tracking (possible combination with Personal Behavior Patterns)
- Public Aggregated Searches Revealing Individual Behavior.

Plans are underway to support smart grid system applications that will monitor any device transmitting a signal, which may include non-energy-consuming end use items that are only fitted with small radio frequency identification devices (RFID) tags may be possible.

Whereas, in Europe energy theft and privacy are the most important concerns related to Smart Grid implementation, in other parts of the world (e.g. in the US) it is energy theft and malevolent attacks that are the main concerns.

IV. PRIVACY PRINCIPLES

The increased amount of personal data being processed, the possibility of remote management of connection and the likelihood of energy profiling based on the detailed meter readings make it imperative that proper consideration is given to individuals' fundamental rights to privacy.

Privacy by Design (PbD) is a concept to address the ever-growing and systemic effects of Information and Communication Technologies, and of large-scale networked data systems [12]. The objectives of Privacy by Design — ensuring privacy and gaining personal control over one's information and, for organizations, gaining a sustainable competitive advantage — may be accomplished by practicing the following seven Foundational Principles:

1. Proactive not Reactive; Preventative not Remedial measures by anticipating and preventing privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred — it aims to prevent them from occurring. In short, Privacy by Design comes before-the-fact, not after.

2. Privacy as the Default Setting i.e. ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy — it is built into the system, by default.

3. Privacy Embedded into design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality.

4. Full Functionality — Positive-Sum, not Zero-Sum, *by* accommodating all legitimate interests and objectives in a positive-sum "win-win" manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made such as privacy vs. security, demonstrating that it is possible to have both.

5. End-to-End Security — Full Lifecycle Protection extending securely throughout the entire lifecycle of the data involved — strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion.

6. Visibility and Transparency — thus its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify.

7. Respect for User Privacy — Keeping it User-Centric appropriate notice, and empowering user-friendly options.

Yet, privacy concerns still need to be transposed into specific, precise and non-ambiguous technical requirements if they are to allow the security industry to competitively design and develop privacy-compliant solutions and services. The Privacy by Design concept should, at its turn, be better detailed in order to allow for its practical implementation in concrete cases.

There are also OECD Privacy Guidelines:

1. Collection Limitation Principle: There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

2. Data Quality Principle: Personal data should be relevant to the purposes for which they are to be used and, to the extent necessary for those purposes, should be accurate, compete and kept up-to-date.

3. Purpose Specification Principle: The purposes for which personal data are collected should be specified not later than at the time of collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

4. Use Limitation Principle: Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Principle 3 except— with the consent of the data subject; or by the authority of law.

5. Security Safeguards Principle: Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.

6. Openness Principle: There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

7. Individual Participation Principle: An individual should have the right: a. To obtain from the data controller, or otherwise, confirmation of whether or not the data controller has data relating to him; b. To have communicated to him, data relating to him.

8. Accountability Principle: A data controller should be accountable for complying with measures that give effect to the principles stated above. Data can be sent to the controller in real-time or be stored in the smart meter. In both cases however, under the Data Protection Directive, it is considered that the data have been collected by the controller.
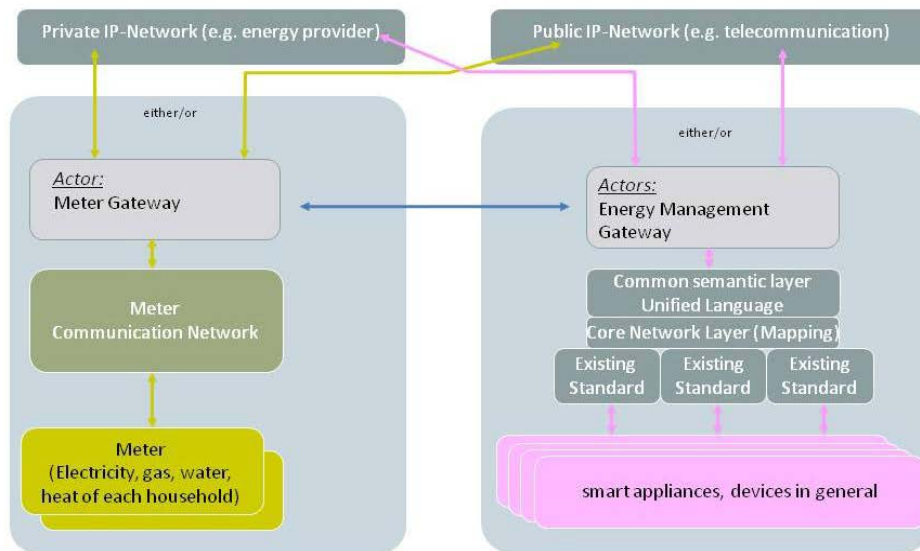
Fig. 4 – Logical separation of metering and energy management

As part of the Privacy by Design process, security and privacy risk assessments will identify the potential risks to data security. Given the novel and vast prospect that is in store with the smart grid and its associated technologies, the task of anticipating security requirements is a challenging one. In order to mitigate risk, the approach should be end-to-end, incorporating all parties and drawing on a broad range of expertise. Security should also be designed in at the early stage as part of the architecture of the network rather than added on later. Appropriately robust security safeguards must be in place that should apply to the whole process including the in-home elements of the network, the transmission of personal data across the network and the storage and processing of personal data by suppliers, networks and other data controllers. Security is a path, not a destination. Security is about risk management and implementing effective counter measures.

The technical and organisational safeguards should cover at least the following areas:

- The prevention of unauthorised disclosures of personal data;
- The maintenance of data integrity to ensure against unauthorised modification;
- The effective authentication of the identity of any recipient of personal data;
- The avoidance of important services being disrupted due to attacks on the security of personal data;
- The facility to conduct proper audits of personal data stored on or transmitted from a meter;
- Appropriate access controls and retention periods;

- The aggregation of data whenever individual level data is not required.

IV. TECHNICAL SOLUTIONS

The reference architecture for the home/building, pointing out the different logical blocks, and can be easily integrated in the whole system architecture is shown in Fig. 4. It is not related to a specific hardware design, but merely shows a logical separation of functions without predefining where and how those functions are implemented.

Final report of the CEN/CENELEC/ETSI Joint Working Group on Standards for Smart Grids [5] presents WAN interface to AMI subsystem & Head-End is used to connect the meter, a Local Network Access Point, or a Neighbourhood Network Access Point to a Central Data Collection system. Typical interface platforms for these interfaces are PSTN networks, public G2 (GPRS) and G3 (UMTS) networks, DSL or broadband TV communication lines, power line communications (PLC), either in narrowband or broadband.

The Head-End systems are the central Data Collection Systems for the Advanced Metering Subsystem. Head-end systems are typically part of an AMR (automatic meter reading) or AMM (automatic meter management) solution. The interface towards the gateways and data concentrators (Network Access Points) is being standardized with Mandate M/441 whilst the interface from head-end systems towards central ERP and meter data management systems is covered by other IEC TCs, e.g. IEC TC 57 (61968-9).

Little work exists on the design of technical solutions to protect privacy in the smart grid [13]. Wagner et al.

propose a privacy-aware framework for the smart grid based on semantic web technologies. Garcia and Jacobs design a multiparty computation to compute the sum of their consumption privately. The NIST privacy subgroup suggests anonymizing traces of readings, as proposed by Efthymiou et al., but also warns of the ease of reidentication. Molina et al. highlight the private information that current meters leak, and sketch a protocol that uses zero-knowledge proofs to achieve privacy in metering. Kumari et al. propose usage control mechanisms for data shared by smart meters connected to web based social networks.

It is equally important to make the principle of privacy-by-design mandatory, including principles of data minimization and data deletion when using privacy-enhancing technologies. As it is currently almost impossible to ensure the full anonymisation of personal data and it is often possible to 'reidentify' or 'de-anonymise' individuals hidden in anonymised data with astonishing ease, only aggregated data should be used to the maximum possible extend. Considering significant privacy threats, we ask for privacy impact assessment to be conducted prior to the smart meter roll out.

Moreover, technical standards and systems should be developed with a focus on upgradeability to safeguard end-to-end security ensuring the overall intelligent metering system is future-proof and ready to cope with future challenges.

Standardization of smart grids is not a business as usual due to the huge number of stakeholders, the necessary speed, the many international activities and the still changing solutions make it a difficult task.

Specific for the data privacy aspects, the European consumer groups are asking for clear regulation around frequency of meter reading and usage of data. It is stressed that only data necessary to perform Smart Grid tasks should be collected and utilised. At the same time, whilst acknowledging benefits, Smart Grid/Meters should be designed for privacy and security.

IV. CONCLUSION

The smart grid brings with it an entirely new and complex model of inter-relationships which poses challenges for the application of data protection. Because of the wide ranging nature of the issues presented by smart metering, it is not possible to provide an exhaustive list of privacy and security points. It is an emerging area of work where it is fully expected that new data protection

problems and solutions will evolve as more smart meters are installed. What is inarguable is that mass rollout of smart meters is already happening, so there is urgency to understand the way that smart meters process personal data, and the issues that this raises. There are some issues of general concern which warrant serious consideration by all those involved in this area.

ACKNOWLEDGEMENT

REFERENCES

[1] ANEC/BEUC POSITION ON ENERGY EFFICIENCY, Joint ANEC/BEUC position paper on the Commission's Communication "Energy Efficiency Plan 2011"

[2] Article 29 Data Protection Working Party, Opinion 12/2011 on smart metering, WP 183, 4.4.2011

[3] BEUC Response To CEER Public Consultation On Demand Response Programmes

[4] Cavoukian A., Privacy By Design …Take The Challenge, Book.

[5] CEN/CENELEC/ETSI Joint Working Group, Standards for Smart Grids, Final report, 4 May 2011.

[6] Elster's White Paper, Privacy Enhancing Technologies for the Smart Grid, 4.10.2011

[7] European Commission, COM(2010) 609, A comprehensive approach on personal data protection in the European Union Brussels, 4.11.2010

[8] European Commission, COM(2011) 202, Smart Grids: from innovation to deployment, Brussels, 12.4.2011

[9] European Technology Platform SmartGrids, Strategic Deployment Document for Europe's Electricity Networks of the Future

[10] Kursawe, K., Danezis, G. and Kohlweiss, M., Privacy-friendly Aggregation for the Smart-grid.

[11] NISTIR 7628, Guidelines for Smart Grid Cyber Security, September 2010

[12] PbD, SmartPrivacy for the Smart Grid, November 2009.

[13] Rial, R. and Danezis G., Privacy-Preserving Smart Metering, WPES11.

Resistance of XOR/XNOR NSDDL cell to Side Channel Attack

Milena Stanojlović and Predrag Petković

Abstract - Complex cryptographic algorithms guard the content of data from unauthorized persons. The security level depends directly on the coding complexity. The complicated algorithm prevents, or impedes the searching for possible combinations that breaks the code in real time. However the attackers use additional information about the behavior of an electronic crypto-system to reduce the number of combinations needed to explore the key. Collecting such information is referred to as Side Channel Attack - SCA. This paper describes simulation results that illustrate resistance of XOR/NXOR logic cell designed by NSDDL method to SCA. The cells are designed in CMOS TSMC035 technology using Mentor Graphics design tools.

Keywords - crypto-system, SCA.

I. INTRODUCTION

The importance of data being transferred through open or semi-closed communication networks provokes unauthorized users to discover their contents. Any unauthorized attempt to access to encrypted content is treated as an attack to the cryptographic system. A common way to prevent unauthorized attack is to increase number of combinations needed to detect the cryptographic key. However it is proven that additional information about cryptosystem behavior reduces required number of combinations [1]. Any attempt for illegal data collection about system behavior that does not rely on direct data reading is known as Side Channel Attack (SCA). The most popular methods for SCA relay on monitoring of dynamics consumption at electronic crypto-system. The most effective are SPA (Simple Power Analysis), DPA (Differential Power Analysis) and EMA (Electromagnetic Analysis) [2, 3].

The waveform of the supply current (IDD) hides very useful additional information about the behavior of cryptographic systems. An abrupt change of current IDD in a CMOS digital circuit occurs only during transition of the logical state. When changing from 0 to 1, output capacitances are charged to the VDD through PMOS network. As the state changes from 1 to 0, capacitances are discharge to ground. In addition within the transition some short-circuit current flows when PMOS and NMOS transistors are on simultaneously. Attackers are able to provide stimulus data, but cannot access the points in

Milena Stanojlović and Predrag Petković are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: milena@venus.elfak.ni.ac.rs,
    predrag.petkovic@elfak.ni.ac.rs.

which they could register the response. The only source of information about the behavior of a circuit is activity expressed through the change of the supply current.

During this research, authors were gained significant experience at a physical level implementation of data protection from SCA into LEDA Laboratory of Electronics, University of Nis. The research team is developing library of CMOS cells that are resistant to DPA attacks. Resistance is measured by the degree of masking and it is larger if the correlation between IDD and circuit behavior is diminished. The focus of our interest is NSDDL (No Short-circuit current Dynamic Differential Logic) method [5, 6].

We have already developed a restricted set of simple cells resistive to SCA [7, 8]. The aim of this paper is to present SCA resistivity of a more complex cell that is composed of already developed. To illustrate types of cells required for implementation of the RSA public key cryptosystem we refer to [4]. The squaring with serial squarer block requires NOT, AND, delay element and full adder. As we already have designed NOT, AND and D-flip-flop, the next step is to design XOR that would be a building block for the full adder. This paper reports the resistance to SCA of the new cell.

Simulation results were obtained using *ELDO* simulator of *Mentor Graphics Design Architect* environment. To draw the layout was used *IC studio Mentor Graphics tools*, while the DRC (*Design Rule Check*), LVS (*Layout Versus Schematics*) and PEX (*Parasitic Extraction*) perform by using *Calibre*. The technology chosen for the design is TSMC035.

The subsequent section reviews the core of NSDDL method. The third section explores design methodology and SCA resistivity of AND/NAND/OR/NOR NSDDL cell. This cell is a building block for XOR/XNOR NSDDL cell that is described in the forth section.

II. NSDDL METHOD

Cells resistant to SCA are based on the idea that each combination of input signals results in the same power consumption. This is possible when every logic cell has the counterpart that will react complementary. Therefore every cell has two outputs denoted as *true* and *false*. The hardware is doubled, but the effect of masking the true function of the cell is gained.

NSDDL method is based on the three phase clocking. The first phase named *pre-charge* is aimed to drive all outputs (true and false) of all logic cells go to high logic

level. In the second phase, known as *evaluation* phase true outputs takes desired value and false output takes the complementary value. The third phase is named *discharged* because all outputs go to the low logic level.

The advantage of this method compared to other popular solutions, like WDDL [3] is its immunity to imbalance loads at true and false output. This is achieved by using a dynamic NOR circuit (DNOR) which minimizes the impact of short-circuit currents in the CMOS circuit. It is the integral part of the control logic and NSDDL cells. Figure 1 illustrates circuitry of DNOR cell.
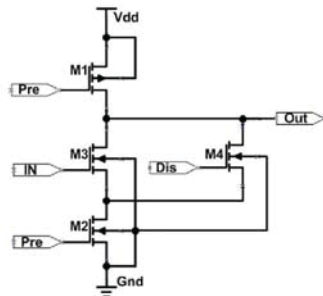

Fig. 1. DNOR circuit

Figure 2 illustrates waveforms of control signals. During the pre-charge phase signals PRE=0 and DIS=0, transistor M1 is *on*, while the other transistors are *off*. The output goes to logic 1, regardless of the input signal IN. The *evaluation* phase begins when signal PRE=1 And DIS=0. Then M1 and M4 turns off, M2 is *on*, and the input signal IN controls the state of the transistor M3. If the signal IN=0, M3 is *off*, so that the output remains at logical 1. If IN=1, M3 and M2 are *on* and output switches to 0. It is obvious that the output achieves inverting function of the input signal. Discharging phase occurs when PRE=1 and DIS=1. Therefore M3 is *off* and M4 is *on* and output goes to low logic level regardless to input signal.
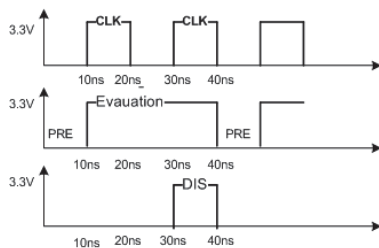
Fig. 2. Time waveforms of control signals for DNOR cell

III. RESISTANCE TO SIDE CHANNEL ATTACKS OF AND/NAND/OR/NOR NSDDL CELL

This section recalls to the results obtained for AND/NAND and OR/NOR NSDDL cells [8] that will be used in the design of XOR/NXOR NSDDL cell.

Block diagram of NAND/AND and NOR/OR NSDDL, SCA resistant cells are presented in Figures 3 and 4, respectively. According to the fact that these cells explore mutually complementary function, it is obvious that they can be realized using the same hardware. The only difference makes the meaning of the true and the false output.
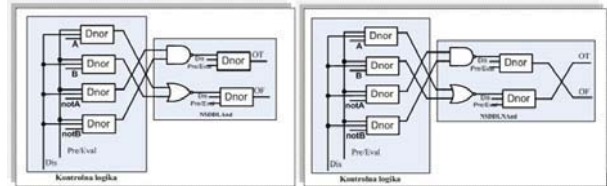
Fig. 3. Block scheme of NSDDL AND and NAND SCA resistance cell
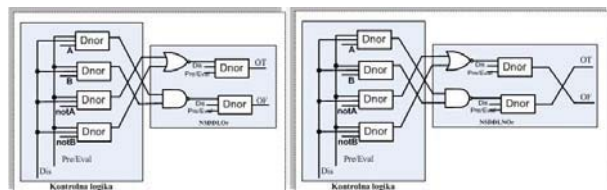
Fig. 4. Block scheme of NSDDL OR and NOR SCA resistance cell

Figures 3 and 4 illustrate that both cells need mutually complement input signals A/notA and B/notB. Using de Morgan rules it is easy to see that simple permutation of input signals (A, notA, B, notB) provides four different logic functions with the same hardware. Therefore this structure is named AND/NAND/OR/NOR SCA resistant cell.

It is important to note that all functions are implemented using native logic circuits with negative logic (NAND i NOR) which can be easily implemented in CMOS technology.

DNOR circuit represents basic element for all SCA resistant cells in NSDDL technique. Prime role of this circuit is to decrease short-circuit current in CMOS circuit Moreover, it provides inverting function when transforming from standard to NSDDL logic.

In order to estimate SCA resistance we consider the energies needed for output state transition during different combinations of input signals. As reference we use standard AND, NAND, OR and NOR cells and compare behavior of standard and NSDDL cell. For standard cells one can expect strong correlation between energy required for particular transition and combination of input signals. In particular any neutral event requires minimal energy while rise transition at the output needs more current to charge the output capacitance. NSDDL cells are designed with intention to mask cell operation regarding I_{DD}. Therefore they should provide minimal correlation between stimulus signals and I_{DD}.

Table I systematizes results of comparison.

Columns 1 and 2 indicate input combinations. Symbols "↑" and "↓" denote the rise and fall transition, respectively. Columns 3, 4, 5 and 6 present results obtained for standard AND, NAND, OR and NOR cells, respectively, while column 7 refers to NSDDL cell.

TABLE I
CHARACTERISTICS COMPARISON OF CLASSIC AND NSDDL CELLS

1	2	3	4	5	6	7
A	B	E_{ANDc} [pJ]	E_{NANDc} [pJ]	E_{ORc} [pJ]	E_{NORc} [pJ]	E_{NSDDL} [pJ]
0	↑	0.05	0.05	-0.49	-0.46	-2.80
0	↓	-0.05	-0.05	-0.674	-0.47	-2.77
↑	0	0.05	0.05	-0.50	-0.50	-2.77
↓	0	-0.05	-0.05	-0.76	-0.55	-2.74
↑	↑	-0.72	-0.69	-0.44	-0.43	-2.75
↓	1	-0.86	-0.65	-0.05	-0.05	-2.82
↑	1	-0.65	-0.62	0.05	0.05	-2.77
1	↓	-0.93	-0.73	-0.007	-0.007	-2.79
1	↑	-0.69	-0.66	0.007	0.007	-2.74
↓	↓	-0.97	-0.76	-0.71	-0.52	-2.76
E_{max} [pJ]		0.05	0.05	0.05	0.05	-2.74
E_{min} [pJ]		-0.97	-0.76	-0.76	-0.55	-2.82
E_{av} [J]		-0.48	-0.41	-0.36	-0.30	-2.77
δE [%]		210.2	196.98	222.05	202.67	2.81
σ [fJ]		405.4	337.7	310.3	243.1	24.31
NSD[%]		83.91	82.23	85.64	82.59	0.87

Energy consumption is expressed as integral in time of power ($I_{DD} \cdot V_{DD}$) during one cycle of input signal change. For AND, NAND, OR and NOR this cycle lasts as all three operational phases needed for NSDDL cell. In order to get better insight into behavior of every cell we derived from the simulation results the following parameters:

- maximum energy (E_{max}),
- minimum energy (E_{min})
- average energy (E_{av})
- relative difference in respect to E_{av} (δ)
- standard deviation (σ)
- normalized standard deviation in respect to E_{av} (NSD).

As a measure of SCA resistance we consider normalized standard deviation.

For standard logic cells this parameter reaches 85%. Obviously this indicates strong correlation between energy (practically the current, because V_{DD}=const) and input signal transition. However, NSDDL cell has NSD <1%. This is sufficient to conclude that AND/NAND/OR/NOR NSDDL cell is immune to SCA using DPA.

Figure 5 illustrates layout of SCA resistant AND/NAND/OR/NOR2 cell. Layout of NSDDL cells that perform particular logic function AND2, NAND2, OR2 and NOR2 cells differs only regarding the order of input and output ports which form desired functions. By rule of symmetry, true and false parts of the circuit are mirrored.
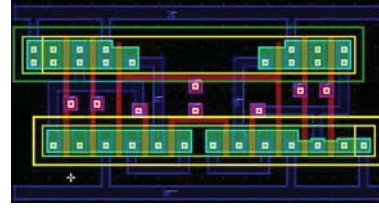
Fig. 5. Layout of SCA resistant AND/NAND/OR/NOR2 cell

IV. RESISTANCE TO SIDE CHANNEL ATTACKS OF XOR/XNOR NSDDL CELL

Figure 6 illustrates block diagram of XOR/XNOR, SCA resistant cell. As all other NSDDL cells it has true and false inputs and output. It is clear that the same structure provides the XOR function at the true output (OT) and XNOR function at the false output (OF). Therefore it is referred to as XOR/XNOR NSDDL cell.

Comparing Figures 3 and 4 with Fig. 6 one easily concludes that it is composed of three AND/NAND/OR/NOR2 cells described above.

Therefore it is interesting to track how the property defined as resistance to SCA is being transferred from lower hierarchical design level to the higher. Aiming that goal we performed similar set of simulation as for AND/NAND/OR/NOR2 NSDDL cell.

Referent cells were standard XOR and standard XNOR cell. They are compared for energy consumption with XOR/XNOR NSDDL cell.
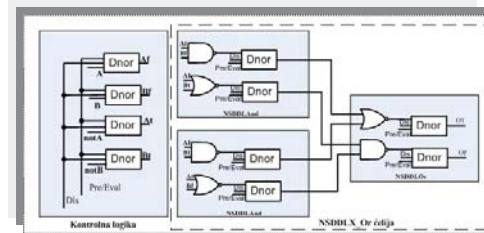
Fig. 6. Block diagram of NSDDL XOR SCA resistance cell

Table II summarizes results of the comparison. NSD parameter that was less than 1% (0.87%) for AND/NAND/OR/NOR2 NSDDL cell remained almost the same. Although slightly increased to 0.91%, it is still less than 1% that qualifies this cell as resistant to SCA. Actually NSD has increased for 4.6% in respect to AND/NAND/OR/NOR2 NSDDL cell. The total improvement of the resistivity to SCA in comparison with standard cells overcomes 2500% for XOR and 5000% for XNOR cell.

Figure 7 shows layout of SCA resistant XOR/XNOR NSDDL cell. Layout of XOR and XNOR cells differs only in respect to the order of input and output ports which form desired functions.

The layout complies with all rules for symmetry of true

TABLE II
CHARACTERISTICS COMPARISON OF CLASSIC AND NSDDL CELLS

1	2	3	4	5
A	B	E_{XORc}[pJ]	E_{XNORc}[pJ]	E_{NSDDL}[pJ]
0	↑	-0.35	-0.48	-6.38
0	↓	-0.51	-0.30	-6.21
↑	0	-0.34	-0.47	-6.22
↓	0	-0.48	-0.33	-6.22
↑	↑	-0.28	-0.05	-6.27
↓	1	-0.35	-0.47	-6.16
↑	1	-0.48	-0.31	-6.27
1	↓	-0.34	-0.47	-6.19
1	↑	-0.52	-0.32	-6.21
↓	↓	-0.27	-0.05	-6.23
E_{max} [pJ]		-0.27	-0.05	-6.16
E_{min} [pJ]		-0.52	-0.48	-6.38
E_{av} [pJ]		-0.39	-0.33	-6.24
δE [%]		63.64	131.76	3.53
σ [fJ]		91.77	154.18	56.58
NSD[%]		23.51	47.43	0.91
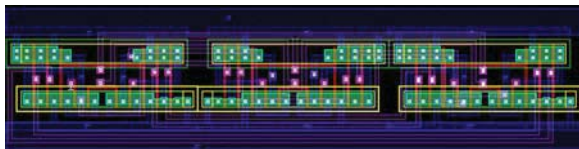
Fig. 7. Layout of SCA resistant XOR/XNOR cell

and false parts in order to suppress unequal consumption in complementary parts of the cell.

V. CONCLUSION

This paper presents simulation results that prove resistance of XOR/XNOR cell designed by NSDDL method to side channel attack. NSDDL method characterizes the implementation of duplicated hardware that provides true and false output. The false output has the same function as inverted true output. The basic idea is to mask the correlation between the supply current and the activity of the cell. This is possible to obtain if input signals are doubled Three-phase clock signal guarantee that all outputs will start from the high level during the pre-charging and that will take low level during the third phase. The cell operates the desired logic function in the middle phase. Then the true output takes the desired output state and, simultaneously the false output has opposite transition. Due to duplicated hardware the same cell is able to generate both XOR and XNOR functions and consequently named XOR/XNOR NSDDL cell. This cell is composed of three simple NSDDL cells that perform AND/NAND /OR/NOR function. The resistance to SCA was monitored through energies required for output transition under

different combination of input signal. The cell is resistive if all changes require the same energy.

Therefore as a measure for a cell resistance to SCA we considered standard deviation normalized to the average energy (NSD). The resistance of AND/NAND/OR/NOR NSDDL cell is less than 1% (0.87%). When implemented into XOR/XNOR NSDDL cell the resistivity decreased for less than 5% relatively to AND/NAND/OR/NOR NSDDL but still remained less than 1% (0.91%). This proves that NSDDL cells transfer their resistivity to SCA into the complex circuit where they are build-in.

REFERENCES

[1] Koc, Cetin Kaya (Ed.) *Cryptographic Engineering*, Springer, 2009.

[2] P. M. Petković, M. Stanojlović, V. B. Litovski "*Design of side-channel-attack resistive criptographic ASICS*",Forum BISEC 2010, Zbornik radova druge konferencija o bezbednosti informacionih sistema, Beograd, Srbija, Maj 2010, pp 22-27.

[3] M. Stanojlović, P. Petković, *"Hardware based strategies against side-channel-attack implemented in WDDL"*, Electronics, Vol. 14, No. 1, Banja Luka, June, 2010, pp. 117-122

[4] Danger, J.-L. Guilley, S. Bhasin, S. Nassar, M., *Overview of Dual Rail with Precharge Logic Styles to Thwart Implementation-Level Attacks on Hardware Cryptoprocessors,* Proc. of International Conference on Signals, Circuits and Systems SCS'2009, Djerba, Tunisia, November 5-8 2009, pp. 1-8

[5] M. Bucci, L. Giancane, R. Luzzi, A. Trifiletti: *"Three-Phase Dual-Rail Pre-Charge Logic"*. In: Goubin, L., Matsui, M. (eds.) CHES 2006. LNCS, vol. 4249, pp. 232–241. Springer, Heidelberg (2006).

[6] J. Quan and G. Bai, "*A new method to reduce the side-channel leakage caused by unbalanced capacitances of differential interconnections in dualrail logic styles*", 2009 Sixth International Conference on Information Technology: New Generations, DOI 10.1109/ITNG. 2009.185, pp. 58-63.

[7] Stanojlović, M., Petković, P.: *Otpornost na bočne napade ASIC kripto sistema zasnovanog na standardnim ćelijama*, VIII Simposium on Industrial Electronics INDEL 2010, Banja Luka, Bosnia and Herzegovina, 4-6 November, 2010, pp. 110-114, ISBN 978-99955-46-03-8

[8] Petković, P., Stanojlović, M.: *Hardverska zaštita od napada na kripto-sistem zasnovana na primeni ćelija koje maskiraju informaciju o potrošnji*, Zbornik LV konferencije ETRAN, Banja Vrućica, Bosna i Hercegovina, ISBN 978-86-80509-66-2.

Simulation of defects in sequential NSDDL Master/Slave D flip flop circuit

Milena Stanojlović and Vančo Litovski

Abstract - Testing of the NSDDL Master/Slave D flip flop (MSDFF) that represents a sequential cell, being part of NSDDL (No Short-circuit current Dynamic Differential Logic) side-channel-attack-resistant library, will be presented in this paper. Fault dictionary will be created based on repetitive simulation preformed on the circuit level description of the flip-flop with faults inserted one by one. Only open-circuit and short-circuit will be considered.

Keywords – Testing, sequential logic, encryption, open circuits, short circuits.

I. INTRODUCTION

The misuse of data is increasingly common. It became necessary to develop new methods, both in software and in hardware, in order to protect data. The domain of this paper is the use of cryptographic methods in ASIC hardware, based on applying standard cell design. The cryptographic algorithm in hardware protects the information leaks out of the device trough so called "side channel". Attacks are based on analyses of the leaked data are known as side channel attacks (SCA) [1]. Important information, such as secret keys, can be obtained by observing the power consumption, the electromagnetic radiation, the timing information etc.

After a long study of different cryptographic methods in hardware, for data protection, we chose one that meets the set criteria. This is the so-called NSDDL logic [2] (*No Short-circuit current Dynamic Differential Logic*). The method is based on a modification TDPL (Three-Phase Dual-Rail Pre-Charge Logic) approach [3] which introduces a third phase of work, during which all the capacitors in the circuit are empty. An important novelty in NSDDL method is immunity on unbalanced load true and false output. In addition, the method requires only one a new cell that is combined with standard logic cells.

Further in this paper, special attention will be devoted to testing NSDDL Master/Slave D flip flop circuit. For intentional introduction of defects, shorts and opens, in fault free circuit, output signal and supply current for each defect for certain combinations of input signals will be monitored. A number of simulations will depend on num-

Milena Stanojlović is with The Innovation Center, School of Electrical Engineering, University of Belgrade, Bul. Kralja Aleksandra 73, 11120 Belgrade, Serbia, E-mail: milena@venus.elfak.ni.ac.rs

Vančo Litovski is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia., E-mail: vanco@elfak.ni.ac.rs.

ber of defects which are tested. The authors decided for this way of testing because of establishing the test sequence. Therefore with given sequence success of the test is determined. As Coverage of defects with given sequence is better, testing is more successful. With this, it can be shown that one test cover more defects which significantly speeds up process of testing. Besides examining logic function of the circuit, it is also very important to compare supply currents of faulty and fault free circuits. When defect is present in the circuit, it is very possible that it will be mapped in to change of mentioned supply current [4].

II. CELL TESTING

A. NSDDL Master/Slave D flip flop circuit

Block schemes of NSDDL Master/Slave D flip flop (MS DFF) cell is presented on figures 1. This structure is composed of two identical standard MS DFFs, invertors and Dnor circuits. Each of MS DFFs inputs are connected to appropriate output of Dnor circuit in crisscross manner. Outputs of MS DFFs are connected to the Dnor circuit as well, but this time over inverting logic gate.
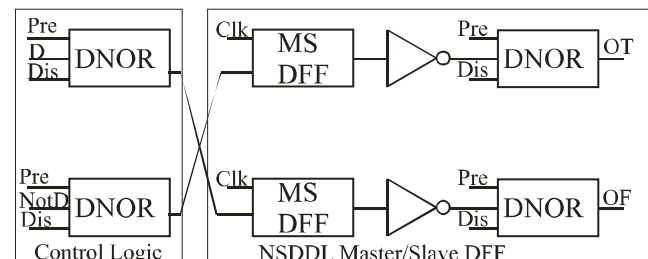
Fig. 1. Block scheme of SCA resistant NSDDL MS DFF cell

B. Testing of Master/Slave D flip flop circuit

Since a number of transistors in MSDFF is big, so marking defects for each transistor on MSDFF schematic is irrational. In order to perform simulations a number of defects which are to be simulated have to be determined. After that defect is inserted in the circuit and appropriate observing point is adopted. This point should provide visibility of the defect's effect [5, 6]. Since circuit contains eighty-eight transistors, five hundred and eight defects of mentioned type can occur. As can be seen from Figure 1

symmetric circuit structure in respect for true and false output is considered. This enables to half the total number of defects. Taking the previous in to account there are still forty four transistors to examine. Therefore the simulation of defects for each transistor for its self is a very tedious but unavoidable work. For all allowable combinations of input signals two hundred and sixty four simulations for faulty and one for fault free circuit are performed. For each transistor six defects are examined where each defect is introduced one after another. Transistors are denoted with *Pi_KSxy/Prex*, or *Nj_KSxy/Prex*, where P and N represent type of the transistor. Counters marked as *i*=0,1...,20, and

j=0,1...,22 represents index of pMOS and nMOS transistor, respectively. With *KSxy* short circuit is denoted while xy determines between transistor connections these shorts occurs. Therefore xy can take values from set {GD, GS, DS} where GD stands for gate-drain, GS for gate-source and DS drain-source. Similar is valid for Prex as well. In this case Prex represent open circuit of connection denoted with x. Here x is from set {G, D, S} where G, D and S represents gate, drain and source transistor terminals, respectively.

The goal is to perform exhaustive test regardless this kind of test is very demanding and tedious.
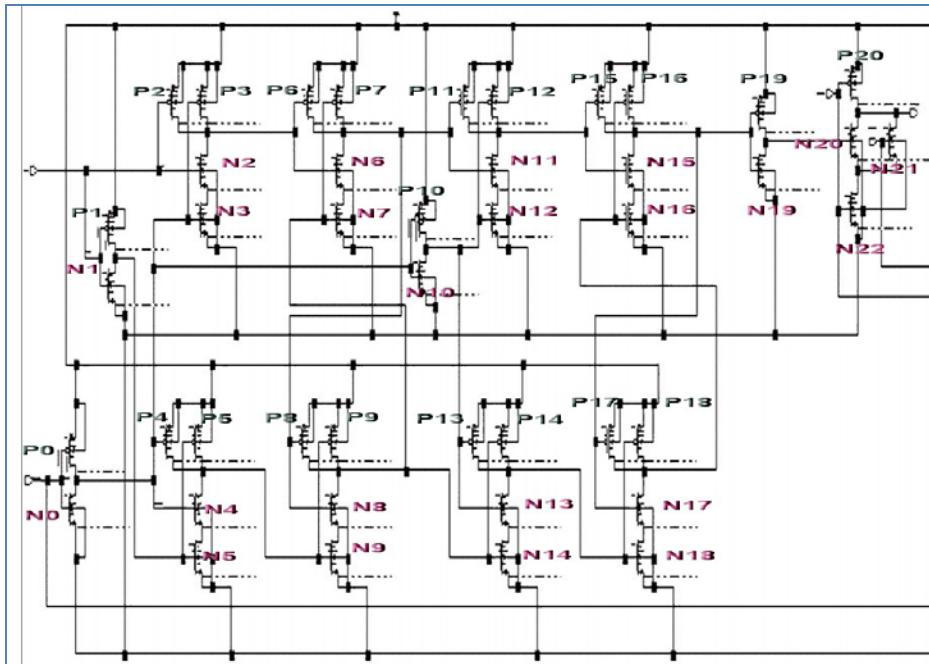
Fig. 2 NSDDL MSDFF, half circuit schematic with denoted transistors
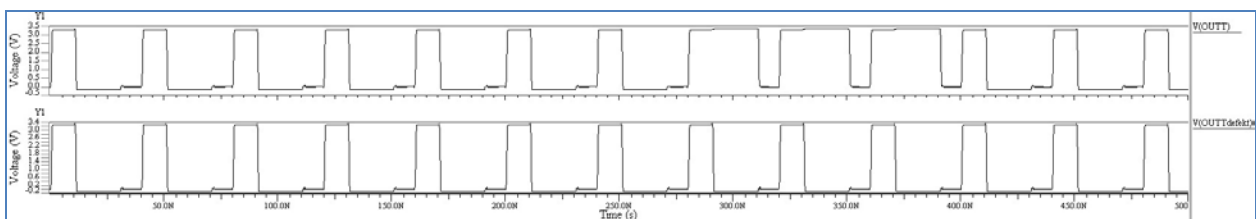
Fig. 3 Time waveforms of output voltage fault free and faulty circuit with defect P4_PreS

This is primarily reflected on time needed for simulations, processing and systematization of obtained results which makes this kind of testing very time consuming.

On Figure 1 block scheme of NSDDL MSDFF is shown which consist of is eighty eight transistors. Respecting symmetry, only half of the circuit is observed so figure 2 illustrates half regarding true output.

Erect of every defect is firstly observed with a respect to a logic function of the circuit. When logic function is violated in can be considered that defect is detected. An important number of defects in the circuit were detected in this way. From two hundred and sixty four defects, two hundred and thirty two defects were detected by only observing output signal. Figure 3 illustrates one such case for inserted defect of open circuit at source of pMOS trans-
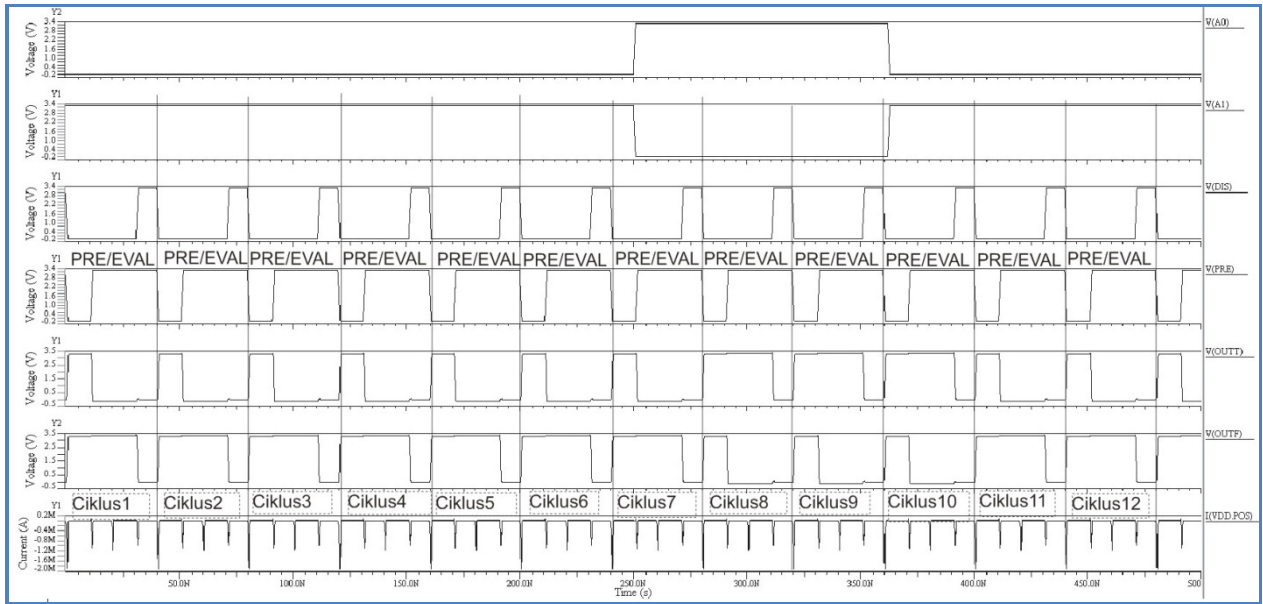
Fig. 4. Time waveforms of inputs, outputs and *idd* of fault free NSDDL MSDFF circuit

istor with index four (*P4_PreS*).

First waveform represents response of fault free circuit, while second represents response for faulty one. It can be clearly seen that these two responses are different which automatically implies defect detectability. Response of the circuit can be different depending on the type of a defect that is inserted in to it. Hence, at the output of the circuit distorted or fixed value (logic zero or one) signal can occur which is enough for detecting the presence of the defect since logic function is violated.

It can be noted that with this kind of testing good results are archived because large number of defects are detected in a quite easy way. For defects that do not violate logic function, additional analysis of *idd* is required. Namely, autocorrelation function of *idd* for fault free and correlation function between *idd*s for fault free and faulty circuit are compared.

Autocorrelation function of *idd* for fault free circuit is defined with (1) while correlation function between *idd*s for fault free and faulty circuits is defined with (2).

$$R_{idd idd}(\tau) = \frac{1}{T}\int_0^T i_{dd}(t) \cdot i_{dd}(\tau + t) \cdot dt \qquad (1)$$

$$R_{idd i^L_{dd}}(\tau) = \frac{1}{T}\int_0^T i_{dd}(t) \cdot i^L_{dd}(\tau + t) \cdot dt \qquad (2)$$

Practically, root mean square (RMS) values of these functions are compared in order to detect defect. Table I gives results for thirty two defects which were not detectable with logic simulations.

For this reason it was necessary to introduce a new method for defect detection. According to results given in the third column of Table I, where relation between R^L_{idd} $_{idd}{}^L$ and $R_{idd\ idd}$ is expressed in percentages, influence of the defects on *idd* can be seen. It can also be concluded that this approach provide detection of nine of thirty two undetected defects (colored rows in Table I). Remaining twenty there defects stay unrevealed.

Observing results given in Table I one can see that deviation of RMS value of correlation function from RMS value of autocorrelation function is mostly very small (few percent). Therefore it is not safe to adopt vary low threshold for defect detection. Since first significant deviation occurred for *N_11KS_DS* defect (≈ 22%), it was meaningful to adopt 20% deviation for threshold of defect detection in this case.

Besides previously discussed method for defect detection, time integral of the *idd* can be used in this purpose as well.

Since operation of the circuit is very specific, time integral of *idd* is calculated during PRE and EVALUATION phases separately for all combinations of input signals. Therefore, time integral of *idd* for fault free and faulty circuits are compared under same input conditions. Time interval occupied with PRE and EVALUATION phases represents one cycle. Practically, no this interval time integration of *idd* is performed. On figure 4 these intervals are marked as cycles.

Every deviation in value of the integral for each cycle is expressed in percentage and given in Table II. Value of this integral for faulty circuit is compared with fault free one in every cycle. With this method eleven of remaining twenty three defects are detected.

TABLE I
DETECTION OF DEFECTS BASED ON CORRELATIONS OF CURRENTS FAULT FREE AND FAULTY CIRCUITS

Type of defect on the transistor	RMS_R$iddidd$ [A^2]	(RMS$_{ispravno}$- RMS$_{Loše}$)/RMS$_{ispravno}$*100
Ispravno kolo	8.45E-6	
P_0KS_GS	8.06E-2	953586.10%
P_2PrekD	8.59E-6	1.59%
P_2PrekS	8.62E-6	2.03%
P_3prekD	8.27E-6	-2.13%
P_3prekS	8.26E-6	-2.28%
P_5PrekD	8.87E-6	4.91%
P_5PrekS	8.96E-6	6.05%
P_7PrekD	8.51E-6	0.72%
P_7PrekS	8.55E-6	1.10%
P_8PrekD	8.44E-6	-0.16%
P_8PrekS	8.47E-6	0.28%
P_11PrekD	8.63E-6	2.12%
P_11PrekS	8.51E-6	0.68%
P_11PrekG	1.19E-5	41.11%
P_12PrekD	8.23E-6	-2.65%
P_12PrekS	8.16E-6	-3.47%
P_13PrekD	8.29E-6	-1.86%
P_13PrekS	8.33E-6	-1.42%
P_14PrekD	8.41E-6	-0.52%
P_14PrekS	8.51E-6	0.66%
P_14PrekG	1.05E-5	24.36%
P_16PrekD	8.44E-6	-0.10%
P_16PrekS	8.50E-6	0.58%
P_17PrekD	8.40E-6	-0.60%
P_17PrekS	8.46E-6	0.06%
P_20KS_GS	4.32E-2	510526.44%
N_2KS_DS	1.19E-5	40.27%
N_3KS_GD	1.48E-5	75.17%
N_3KS_DS	8.67E-6	2.56%
N_11KS_DS	1.03E-5	21.97%
N_13KS_DS	1.09E-5	29.23%
N_14KS_DS	1.26E-5	48.91%

Therefore, number of undetected defects is reduced to only twelve. Comparing this number with total number of defects (two hundred and sixty four) one can conclude that defect coverage is quite good using these test methods.

Remaining twelve defects do not significant influence on *idd* so they can hardly be detected this way. These defects are: *P7_PrekD, P7_PrekS, P8_PrekD, P8_PrekS, P13_PrekS, P14_PrekS, P14_PrekD, P16_PrekS, P16_PrekD, P17_PrekS and P17_PrekD*.

It can be concluded that combination of three test methods, i.e. logic function violation, comparison of autocorrelation and correlation functions and comparison of time integral of *idd* for fault free and faulty circuits gives solid defect coverage.

This means that for safe testing a different methods and techniques should be combined.

From two hundred and sixty four defects two hundred and fifty two were detected. Since this is result for only a half of circuit, total defect coverage is five hundred and four from five hundred and twenty eight which is nearly 96%. It can be said that the testing was successful.

TABLE II
COVERED DEFECTS

	N2_ KSDS	P10_ PrekD	P11_ PrekS	P11_ PrekD	P12_ PrekS	P1_ PrekS	P1_ PrekD	P2_ PrekS	P2_ PrekD	P4_ PrekS	P4_ PrekD
1	6.94%	18.38%	-8.40%	5.71%	-67.49%	-5.81%	30.16%	6.30%	-0.10%	84.54%	81.70%
2	0.77%	18.35%	0.17%	-0.22%	31.21%	0.83%	24.14%	0.18%	-0.07%	81.20%	64.43%
3	0.87%	18.34%	0.16%	-0.45%	130.6%	0.82%	24.09%	0.13%	-0.09%	81.27%	64.40%
4	0.88%	18.34%	0.16%	-0.15%	230.1%	0.82%	24.10%	0.24%	-0.08%	81.73%	64.41%
5	0.88%	18.34%	0.16%	-0.37%	329.5%	0.82%	24.10%	0.09%	-0.08%	81.87%	64.41%
6	0.88%	18.34%	0.16%	-0.37%	429.0%	0.82%	24.10%	0.17%	-0.08%	80.77%	64.41%
7	239.60%	0.13%	-3.51%	-3.07%	528.0%	6.49%	-1.80%	-4.86%	-4.91%	91.98%	76.74%
8	240.06%	-0.15%	-5.45%	-4.55%	629.1%	6.12%	-1.48%	-4.87%	-4.82%	11.82%	0.65%
9	239.70%	-0.22%	-5.52%	-4.65%	728.2%	6.23%	-1.59%	-4.98%	-4.89%	11.83%	0.38%
10	1.27%	18.51%	-1.75%	-1.83%	829.1%	0.90%	24.45%	0.13%	0.11%	11.11%	-0.46%
11	0.79%	18.35%	0.17%	-0.10%	928.0%	0.83%	23.97%	0.13%	-0.07%	80.79%	64.55%
12	0.87%	18.33%	0.15%	-0.22%	1027.%	0.82%	24.13%	0.17%	-0.09%	81.71%	64.40%

III. CONCLUSION

This paper presents some of the techniques for testing applied on encrypted NSDDL MSDFF cell. First basic operation of unit under test was explained. Two proper methods for testing this sequential logic are adopted, namely logic function violation and testing based on power supply current. From last method two techniques are chosen to be applied on the circuit, i.e. comparison of autocorrelation and correlation functions and comparison of time integral of *idd* for fault free and faulty circuits. These techniques were briefly commented and explained. A number of simulations were performed in order to make appropriate fault dictionary for defects of short/open circuit type. Obtained results are presented and commented as well.

ACKNOWLEDGEMENT

REFERENCES

[1] Petković, P. M., Stanojlović, M., and Litovski, V. B. "Design of side-channel-attack resistive criptographic ASICS", Forum BISEC 2010, Zbornik Radova Druge Konferencija o Bezbednosti Informacionih Sistema, Beograd, Srbija, Maj 2010, pp 22-27.

[2] Quan J., and Bai, G., "A new method to reduce the sidechannel leakage caused by unbalanced capacitances of differential interconnections in dual-rail logic styles", 2009 Sixth International Conference on Information Technology: New Generations, DOI 10.1109/ITNG.2009.185, pp. 58-63

[3] Bucci, M., Giancane, L., Luzzi, R., and Trifiletti, A., "Three-Phase Dual-Rail Pre-Charge Logic". In: Goubin, L., Matsui, M. (eds.) CHES 2006. LNCS, vol. 4249, pp. 232–241. Springer, Heidelberg (2006)

[4] Litovski, V., "Projektovanje elektronskih kola", ISBN 86-7369-015-3, DGIP Nova Jugoslavija, Vranje, 2000.

[5] Litovski, V., Osnovi testiranja elektronskih kola, ISBN 978-86-85195- 71-6, Elektronski fakultet, Niš, 2009.

[6] Milovanović, D., and Litovski, V., "Fault Models of CMOS Circuits", Microelectronics and Reliability, 1994, Vol.34, No. 5, pp. 883-896.

Quantitative Analysis of Reactive Power Calculations for Small Non-linear Loads

Marko Dimitrijević and Vančo Litovski

Abstract — In this paper we will present quantitative analysis of reactive power calculated by various definitions. The analysis will be performed on small non-linear loads, such as CFL and LED lamps. All measurements and calculations are realized using virtual instrument for three-phase power factor and distortion analysis.

Keywords — reactive power, virtual instrument

I. INTRODUCTION

In linear circuits, with sinusoidal voltages and currents, active, reactive and apparent power are correlated with well-known quadratic formula: $S^2 = P^2 + Q^2$. When nonlinear loads are present one should introduce new quantities in the calculations emanated by the harmonics and related power components [1]. Now, the apparent power includes harmonic components. This is of importance in characterization and design of practical power systems which contain non-linear loads such as switched-mode power supplies [2].

Electronic loads are strongly related to the power quality thanks to the implementation of switched-mode power supplies that in general draw current from the grid in bursts. In that way, while keeping the voltage waveform almost unattached, they impregnate pulses into the current so chopping it into seemingly arbitrary waveform and, consequently, producing harmonic distortions [3]. The current-voltage relationship of these loads, looking from the grid side, is nonlinear, hence nonlinear loads. The existence of harmonics gives rise to interference with other devices being powered from the same source and, having in mind the enormous rise of the number of such loads, the problem becomes serious with serious, sometimes damaging, consequences and has to be dealt with properly.

There are a number of power definitions for non-sinusoidal conditions in order to characterize nonlinear loads and measure the degree of loads' non-linearity. As more general term, non-active power was introduced. All definitions have some advantages over others. Although tend to be general, there is no generally accepted

Marko Dimitrijević and Vančo Litovski are with the Department of Electronics, Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia,
E-mail: (marko,vanco)@elfak.ni.ac.rs.

definition.

The idea of quantitative analysis of reactive power and power decompositions is presented in the literature [4]. In study [4] widely recognized power decompositions proposed by Budeanu, Fryze, Kimbark, Shepherd and Zakikhani, Sharon, Depenbrock, Kusters and Moore, and Czarnecki are analysed quantitatively. This analysis is performed on simple test circuit.

In this study, we will perform analysis of various reactive power definitions by measuring characteristics of small non-linear loads, such as CFL and LED lamps. First, we will introduce definitions of reactive power proposed by Budeanu, IEEE standard, Kimbark, Sharon, Fryze, Kusters and Moore. Then, the virtual instrument for non-linear loads will be described. Finally, we will present and discuss measured and calculated values.

II. REACTIVE POWER DEFINITIONS

A sinusoidal voltage source

$$v(t) = \sqrt{2} V_{\text{RMS}} \sin(\omega_0 t) \qquad (1)$$

supplying a linear load, will produce a sinusoidal current

$$i(t) = \sqrt{2} I_{\text{RMS}} \sin(\omega_0 t - \varphi) \qquad (2)$$

where V_{RMS} is the RMS value of the voltage, I_{RMS} is the RMS value of the current, ω is the angular frequency, φ is the phase angle and t is the time. The instantaneous power is

$$p(t) = v(t) \cdot i(t) \qquad (3)$$

and can be represented as

$$p(t) = 2V_{\text{RMS}} I_{\text{RMS}} \sin \omega t \cdot \sin(\omega_0 t - \varphi) = p_p + p_q. \quad (4)$$

Using appropriate transformations we can write:

$$\begin{aligned} p_p &= V_{\text{RMS}} I_{\text{RMS}} \cos \varphi \cdot (1 - \cos(2\omega_0 t)) = \\ &= P \cdot (1 - \cos(2\omega_0 t)) \end{aligned} \qquad (5)$$

and

$$p_q = -V_{\text{RMS}} I_{\text{RMS}} \sin \varphi \cdot \sin(2\omega_0 t) = -Q \sin(2\omega_0 t) \quad (6)$$

where

$$P = V_{\text{RMS}} I_{\text{RMS}} \cos \varphi, \ Q = V_{\text{RMS}} I_{\text{RMS}} \sin \varphi \qquad (7)$$

represent real (P) and reactive (Q) power.

It can be easily shown that the real power presents the average of the instantaneous power over a cycle:

$$P = \frac{1}{T} \int_{t_0}^{t_0+T} v(t) \cdot i(t) \cdot dt \qquad (8)$$

where t_0 is arbitrary time (constant) after equilibrium, and T is the period (20 ms in European and 1/60 s in American system, respectively).

The reactive power Q is the amplitude of the oscillating instantaneous power p_q. The apparent power is the product of the root mean square value of current times the root mean square value of voltage:

$$S = V_{RMS} I_{RMS} \qquad (9)$$

or:

$$S = \sqrt{P^2 + Q^2}. \qquad (10)$$

In the presence of nonlinear loads the system no longer operates in sinusoidal condition and use of fundamental frequency analysis does not apply any more. The non-sinusoidal voltage and current are expressed by Fourier series:

$$v(t) = V_0 + \sum_{k=1}^{+\infty} \sqrt{2} V_{k,RMS} \cos\left(k\omega_0 t + \theta_k\right)$$
$$i(t) = I_0 + \sum_{k=1}^{+\infty} \sqrt{2} I_{k,RMS} \cos\left(k\omega_0 t + \psi_k\right). \qquad (11)$$

where $V_{k,RMS}$ and $I_{k,RMS}$ represent RMS values, and θ_k and ψ_k phases for k-th harmonic of voltage and current, respectively. V_0 and I_0 represent DC values.

The instantaneous power $p(t)$ calculated by equation (3) can be represented as Fourier series:

$$p(t) = P + \sum_{k=1}^{+\infty} P_k \cos\left(k\omega_0 t + \zeta_k\right) \qquad (12)$$

However, expressing components of instantaneous power (P_k, ζ_k) as function of voltage and current spectral components (V_0 , I_0, $V_{k,RMS}$, $I_{k,RMS}$, θ_k and ψ_k) in non-sinusoidal conditions is not an easy procedure. The first addend in sum (12), the real power P, determined as constant energy flow and calculated using equations (8) and (11) is

$$P = V_0 I_0 + \sum_{k=1}^{+\infty} I_{k,RMS} \cdot V_{k,RMS} \cdot \cos\left(\theta_k - \psi_k\right)$$
$$P = P_0 + P_1 + P_H \qquad (13)$$

where P_0, P_1 and P_H stands for DC power, active power of fundamental harmonic and harmonic active power, respectively.

There are a number of reactive power definitions and proposed relations with active and apparent power.

A. Budeanu's definition

The most common definition of reactive power is Budeanu's definition, given by following expression for single phase circuit:

$$Q_b = \sum_{k=1}^{+\infty} I_{k,RMS} \cdot V_{k,RMS} \cdot \sin\left(\theta_k - \psi_k\right) \qquad (14)$$

Budeanu proposed that apparent power is consist of two orthogonal components, active power (13) and non-active power, which is divided into reactive power (14) and distortion power:

$$D_b = \sqrt{S^2 - P^2 - Q_b^2}. \qquad (15)$$

B. IEEE Std 1459-2010 proposed definition

IEEE Std 1459-2010 proposes reactive power to be calculated as:

$$Q_{IEEE} = \sqrt{\sum_{k=1}^{+\infty} I_{k,RMS}^2 \cdot V_{k,RMS}^2 \cdot \sin^2\left(\theta_k - \psi_k\right)} \qquad (16)$$

Equation (16) eliminates the situation where the value of the total reactive power Q is less than the value of the fundamental component.

C. Kimbark's definition

Similar to Budeanu's definition, Kimbark proposed that apparent power is consist of two orthogonal components, non-active and active power, defined as average power. The non-active power is separated into two components, reactive and distortion power. The first is calculated by equation

$$Q_k = I_{1,RMS} \cdot V_{1,RMS} \cdot \sin\left(\theta_1 - \psi_1\right) \qquad (17)$$

It depends only of fundamental harmonic. The distortion power is defined as non-active power of higher harmonics:

$$D_k = \sqrt{S^2 - P^2 - Q_k^2}. \qquad (18)$$

D. Sharon's definition

This definition introduces two quantities: reactive apparent power, S_q, and complementary apparent power S_c, defined as:

$$S_q = V_{RMS} \cdot \sqrt{\sum_{k=1}^{+\infty} I_{k,RMS}^2 \sin^2(\theta_k - \psi_k)} \qquad (19)$$

and

$$S_c = \sqrt{S^2 - P^2 - S_q^2} \qquad (20)$$

where S is apparent power (9) and P active power (8).

E. Fryze's definition

Fryze's definition assumes instantaneous current separation into two components named active and reactive currents. Active current is calculated as

$$i_a(t) = \frac{P}{V_{RMS}^2} v(t) \qquad (21)$$

and reactive current as:

$$i_r(t) = i(t) - i_a(t). \qquad (22)$$

Active and reactive powers are

$$P = V_{\mathrm{RMS}} \cdot I_a$$
$$Q_f = V_{\mathrm{RMS}} \cdot I_r \qquad (23)$$

where I_a and I_r represents RMS values of instantaneous active and reactive currents.

F. Kusters and Moore's power definitions

This definition presents two different reactive power parameters, inductive reactive power

$$Q_{\mathrm{L}} = V_{\mathrm{RMS}} \cdot \frac{\sum_{k=1}^{+\infty} \frac{1}{k} \cdot V_{k,\mathrm{RMS}} \cdot I_{k,\mathrm{RMS}} \cdot \sin(\theta_k - \psi_k)}{\sqrt{\sum_{k=1}^{+\infty} \frac{V_{k,\mathrm{RMS}}^2}{k^2}}} \qquad (24)$$

and capacitive reactive power:

$$Q_{\mathrm{C}} = V_{\mathrm{RMS}} \cdot \frac{\sum_{k=1}^{+\infty} k \cdot V_{k,\mathrm{RMS}} \cdot I_{k,\mathrm{RMS}} \cdot \sin(\theta_k - \psi_k)}{\sqrt{\sum_{k=1}^{+\infty} k^2 \cdot V_{k,\mathrm{RMS}}^2}}. \qquad (25)$$

III. VIRTUAL INSTRUMENT FOR ACQUISITION AND REACTIVE POWER CALCULATIONS

The measurement and calculation of quantities are performed by measurement setup consists of signal acquisition modules – acquisition subsystem and software support– virtual instrument.

The acquisition and conditioning of the electrical quantities is performed by the acquisition subsystem. It is connected to the power grid from one side, and transfers the power to the load while sampling the values of three voltage and four current signals (Fig. 1). The modules for signal conditioning of the voltage and current waveforms provide attenuation, isolation and anti-aliasing.
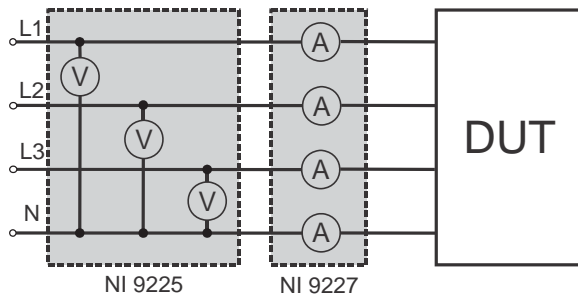

Figure 1. Connection diagram

The acquisition is performed by National Instruments cDAQ-9714 expansion chassis, providing hot-plug module connectivity [5]. The chassis is equipped with two data acquisition modules: NI9225 and NI9227. Extension chassis is connected to PC running virtual instrument via USB interface.

NI9225 has three channels of simultaneously sampled voltage inputs with 24-bit accuracy, 50 kSa/s per channel sampling rate, and 600 V_{RMS} channel-to-earth isolation, suitable for voltage measurements up to 100th harmonic (5

kHz). The 300 V_{RMS} range enables line-to-neutral measurements of 240 V power grids [6].

NI9227 is four channels input module with 24-bit accuracy, 50 kSa/s per channel sampling rate, designed to measure 5 A_{RMS} nominal and up to 14 A peak on each channel with 250 V_{RMS} channel-to-channel isolation [7].

The virtual instrument is realized in in *National Instruments* LabVIEW developing package (Fig. 2), which provides simple creation of virtual instruments. Virtual instruments consist of interface to acquisition module and application with graphic user interface.
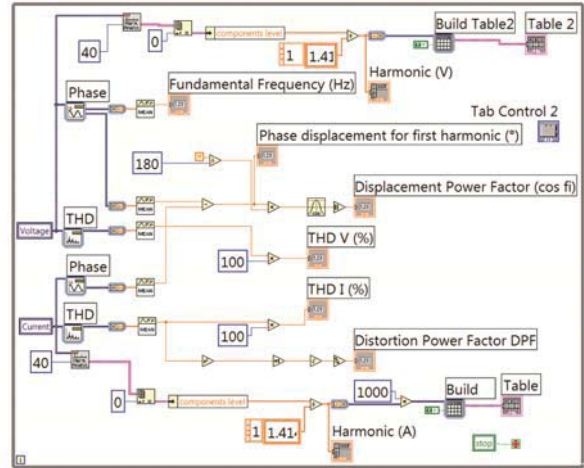

Figure 2. The G code of virtual instrument

Interface to acquisition module is implemented as device driver. cDAQ-9714 expansion chassis is supported by NIDAQmx drivers. All the measurements are performed using virtual channels. A virtual channel is collection of property settings that can include name, a physical channel, input terminal connections, the type of measurement or generation, and scaling information. A physical channel is a terminal or pin at which an analogue signal can be measured or generated. Virtual channels can be configured globally at the operating system level, or using application interface in the program. Every physical channel on a device has a unique name.

For better performance, the main application has been separated into two threads. The first thread has functions for file manipulation and saving measured values. All measured values will be saved in MS Excel file format.

The user interface of the virtual instrument consists of visual indicators. It provides basic functions for measurement. All measured values are placed in a table, and after the measurement process in appropriate file. User interface also provides controls for data manipulation and saving measured values.

IV. COMPARISON OF THE CALCULATED VALUES

We have performed measurements on small loads such as various compact fluorescent lamps (CFL, nominal power 7 W – 20 W), indoor light emitting diode lamps

TABLE I
CALCULATED VALUES FOR DIFFERENT REACTIVE POWER DEFINITIONS – CFL LAMPS

No.	Type	Power	P (W)	S (VA)	N (VAR)	Q_b (VAR)	D_b (VAR)	Q_f (VAR)	Q_{IEEE} (VAR)	S_q (VAR)	Q_k (VAR)	Q_C (VAR)	Q_L (VAR)
1	CFL Rod		11.56	17.84	13.58	-6.16	12.10	13.58	6.16	10.24	-6.16	-4.43	-6.11
2	CFL bulb E27	20	17.14	27.72	21.78	-8.43	20.08	21.78	8.43	14.48	-8.43	-6.46	-8.37
3	CFL tube E27	20	16.77	28.46	23.00	-8.44	21.39	23.00	8.45	14.55	-8.45	-6.07	-8.39
4	CFL bulb E27	15	11.59	18.91	14.94	-5.31	13.97	14.94	5.32	9.22	-5.32	-4.00	-5.28
5	Inc E27	100	86.77	86.78	0.80	-0.50	0.63	0.80	0.50	0.56	-0.50	-0.36	-0.49
6	CFL spot E14	7	5.87	9.32	7.25	-2.83	6.67	7.25	2.81	4.23	-2.81	-2.17	-2.80
7	CFL bulb E27	7	6.16	9.86	7.71	-2.64	7.24	7.71	2.65	4.83	-2.65	-2.03	-2.63
8	CFL bulb E14	9	6.46	10.78	8.63	-2.72	8.19	8.63	2.72	5.45	-2.72	-2.08	-2.70
9	CFL tube E14	11	9.89	16.11	12.72	-4.71	11.82	12.72	4.69	7.89	-4.69	-3.61	-4.66
10	CFL tube E27	18	17.10	28.86	23.24	-8.73	21.54	23.24	8.75	13.27	-8.75	-6.64	-8.68
11	CFL tube E27	11	10.63	17.67	14.12	-5.83	12.85	14.12	5.83	8.85	-5.83	-4.41	-5.79
12	CFL helix E27	11	9.58	16.27	13.16	-4.93	12.20	13.16	4.95	8.75	-4.95	-3.68	-4.90
13	Inc E14	60	55.06	55.06	0.61	-0.37	0.49	0.61	0.37	0.37	-0.37	-0.27	-0.37
14	CFL helix E27	18	17.21	28.87	23.18	-8.82	21.43	23.18	8.83	15.55	-8.82	-6.77	-8.76
15	CFL helix E27	20	18.41	30.68	24.54	-9.95	22.43	24.54	9.93	16.14	-9.93	-7.56	-9.86
16	CFL tube E27	15	12.66	21.97	17.95	-6.32	16.80	17.95	6.33	11.63	-6.33	-4.80	-6.28

TABLE II
CALCULATED VALUES FOR DIFFERENT REACTIVE POWER DEFINITIONS – LED LAMPS

No.	Type	Power	P (W)	S (VA)	N (VAR)	Q_b (VAR)	D_b (VAR)	Q_f (VAR)	Q_{IEEE} (VAR)	S_q (VAR)	Q_k (VAR)	Q_C (VAR)	Q_L (VAR)
1	Spot White E27	15	16.92	34.24	29.77	-3.88	29.52	29.77	4.14	20.01	-4.13	-1.98	-4.06
2	Spot White E27	10	13.23	26.33	22.76	-2.97	22.56	22.76	3.17	15.45	-3.17	-1.51	-3.12
3	Bulb W White E27	8	10.00	19.53	16.77	-2.81	16.54	16.77	2.94	11.52	-2.93	-1.74	-2.89
4	Bulb W White E27	6	8.51	9.45	4.11	0.08	4.11	4.11	0.07	3.29	0.07	0.08	0.07
5	Bulb White E27	6	8.69	9.58	4.04	0.09	4.04	4.04	0.08	3.28	0.08	0.08	0.08
6	Bulb White E27	3	4.07	7.70	6.54	-0.84	6.48	6.54	0.90	4.35	-0.90	-0.45	-0.88
7	RGB Change E27	3	1.92	3.17	2.52	0.01	2.52	2.52	0.01	1.39	0.00	0.05	0.00
8	Spot White E14	3	4.00	8.05	6.99	-0.98	6.92	6.99	1.04	4.86	-1.04	-0.52	-1.02

(LED, nominal power 3 W – 15 W), and two incandescent lamp for power reference (60 W and 100 W).

Table 1 shows values for compact fluorescent lamps, as well as two incandescent lamps. Table 2 shows calculated values of reactive power for LED indoor lamps. Following values are displayed: active power (P), apparent power (S), non-active power (N), Budeanu's reactive power (Q_b), Budeanu's distortion power (D_b), Fryze's reactive power (Q_f), IEEE Std 1459-2010 proposed definition for reactive power (Q_{IEEE}), Shanon's apparent power (S_q), Kimbark's reactive power (Q_k), Kusters-Moore's capacitive (Q_C) and inductive (Q_L) reactive power.

Comparison of Budeanu's reactive and distortion power suggests that all examined CFL and LED lamps are non-linear loads ($D_b > Q_b$). Reactive power calculated from Fryze's definition (23) is equal to non-active power, $N = \sqrt{S^2 - P^2}$. Kimbark's equation (17) for reactive power, which takes only fundamental harmonic into account, gives approximately ±3% deviance from Budeanu's formula (Q_b). It suggests that the actual contribution of harmonic frequencies to reactive power is small – less than 3% of the total reactive power.

IEEE proposed definition (16) always provides value of the total reactive power greater than the value of the fundamental component.

REFERENCES

[1] H. W. Beaty, D. G. Fink, *Standard handbook for electrical engineers* (McGraw-Hill, New York, 2007).
[2] John G. Webster, *The Measurement Instrumentation and Sensors Handbook* (CRC Press, 1999).
[3] T. H. Tumiran, M. Dultudes, The Effect Of Harmonic Distortion To Power Factor, *Proceedings of the International Conference on Electrical Engineering and Informatics*, 2007, pp. 834–837 Institute Teknologi Bandung, Indonesia
[4] M. Erhan Balci, M. Hakan Hocaoglu, "Quantitative comparison of power decompositions," Electric Power Systems Research 78 (2008) 318–329

[5] National instruments cDAQ-9714 Product Data Sheet, National Instruments, http://ni.com

[6] National instruments NI-9225 Product Data Sheet, National Instruments, http://ni.com

[7] National instruments NI-9227 Product Data Sheet, National Instruments, http://ni.com

Simulation of Utility Losses Caused by Nonlinear Loads at Power Grid

Dejan Stevanović, Borisav Jovanović and Predrag Petković

Abstract: This paper quantifies losses in a utility system caused by nonlinear loads. Distortion power is considered as a quantity that reflects the best effects of the losses. These losses are caused by nonlinear loads which are connected to the grid. The paper gives a review of trends related to changes of the character of loads connected to the utility together with its effects. The major problem reflects in the form of losses that utility register due to inadequate measurement equipment. We analyze core of the problem and suggest the solution that is verified by appropriate model and simulation. The method is suitable for implementation in electronic smart meters. This is confirmed by an upgrade of DSP dedicated for power/energy calculation within an original solid-state power meter. The enhanced version of DSP is designed in CMOS 0.35μm technology, using Cadence design tools for designed ASIC circuit.
Key Words: Distortion power, utility, power losses

I. INTRODUCTION

Classic approach to power metering and billing in households relays on registration of active power. It was sufficient in systems with dominant linear resistive loads (electric stove, water heater, electric furnace, incandescent bulbs). The active power is by definition:

$$P = V_{\mathrm{RMS}} I_{\mathrm{RMS}} \cos(\theta) \qquad (1)$$

It has been supposed that large reactive loads exist in industry (predominantly the inductive motors). Therefore power meters aimed for industrial applications were capable to measure active and reactive power. The reactive power by definition is:

$$Q = V_{\mathrm{RMS}} I_{\mathrm{RMS}} \sin(\theta) \qquad (2)$$

Electromechanical meters have bandwidth limitation. Therefore they cannot take into account harmonics [1].

The rapid development of electronics has changed the profile of the common customer's load. Electronic equipment has become dominant consumers. Their characteristic is to operate on small (<5V) DC voltage while supplied from AC 240V RMS. In order to increase efficiency of rectifiers and voltage regulators their operation frequency is moved from 50Hz to several kHz.

Dejan Stevanović is with Innovation Center, School of Electrical Engineering in Belgrade d.o.o. (ICEF) Bul. kralja Aleksandra 73, 11120 Belgrade, Serbia, E-mail: dejan.stevanovic@venus.elfak.ni.ac.rs.

Borisav Jovanović and Predrag Petković are with the Department of Electronics, Faculty of Electronic Engineering, Univer-sity of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: {predrag.petkovic and borisav.jovanovic}@elfak.ni.ac.rs.

Consequently this caused decreased dimensions of passive reactive components. Moreover, in order to diminish losses on active elements, transistors operate in the switch mode. All desired effects were obtained: effectiveness of rectifiers and regulators were considerably improved. As result more power goes to the loads (electronic equipment) and less dissipate on AC/DC convertors. However a small problem aroused. All such loads introduce large nonlinear distortion current that produce higher voltage drop in power line. The number of nonlinear loads has increased according to the tremendous rise of electronics appliances. Therefore, the level of power consumption at nonlinear loads becomes comparable in value with the linear.

The inert power system could not follow the development of electronics and even did not pay attention to possible consequences. The largest two blackouts in North American history (1965 and 2003) demonstrate sensitivity of the power system to small unjustly neglected problems. According to [2] "Both blackouts were the result of cascading failures of the power system, in which seemingly small and localized problems caused the system to become unstable and subsequently affect a much wider area." The current status in Serbia is that utility still uses electrical power meters capable to register only active energy consumption. Moreover, recently published tenders for new electronic power meters does not requests measurements of non-active components of power. The advantage options are required only from industrial power meters.

This paper is aimed to show the real consequences of using obsolete meters to register contemporary consumption. The subsequent section gives definitions of power components that appear in the grid in presence of the nonlinear loads. As one will see instead of active and reactive components, the apparent power contents additional component that is caused only by harmonics in nonlinear loads. This component is known as *distortion power*. It is important to stress that it could have value comparable to the active power and even to exceed it. Therefore, if the utility does not register this component it will have high level of losses.

The third section explains the effects of harmonics to the power system. The subsequent section describes the power meter model capable to deal with harmonic distortion. The fifth section presents simulation results. The architecture of DSP dedicated for energy metering in power meter is presented in sixth section before conclusion.

II. COMPONENTS OF POWER IN DISTORTED SYSTEMS

Traditional power system quantities such as RMS values of current and voltage, power (active, reactive, apparent) are defined for pure sinusoidal condition. Due to harmonic distortion of current and/or voltage the definition of all power components have to be modified. The effect of harmonics must be taken into account. In case when harmonics exist in the power supply system, the instantaneous values of voltage or current can be express as:

$$x(t) = \sum_{h=1}^{M} X_h \sin(\omega_h t + \alpha_h) \qquad (3)$$

where h is the number of harmonic, M denotes the highest harmonic, while X_h, α_h, represent amplitudes and phase of signal. Frequency of the hth harmonic is ω_h. RMS value of signal expressed by Eq. (3) is defined as:

$$X_{\text{RMS}} = \sqrt{\sum_{h=1}^{M} X_{\text{RMSh}}^2} \ , \qquad (4)$$

where X_{RMSh} is the RMS values of the hth harmonic of the voltage or current. Product of the voltage and current at the same harmonic frequency gives the harmonic power. Total active power is defined as:

$$P = \sum_{h=1}^{M} V_{\text{RMS}_h} I_{\text{RMS}_h} \cos(\theta_h). \qquad (5)$$

It could be presented as a sum of components related to the fundamental and other harmonics:

$$P = P_1 + P_H, \qquad (6)$$

where P_1 denotes contribution of the fundamental harmonic *(h=1)* and therefore named *fundamental active power* component; P_H comprises the sum of all higher components *(h=2,...M)* and is referred to as *harmonic active power*.

According to Budeanu the reactive power is defined as:

$$Q_B = \sum_{h=1}^{M} V_{\text{RMS}_h} I_{\text{RMS}_h} \sin(\theta_h) = Q_1 + Q_H \qquad (7)$$

where, similarly to Eq. (5), Q_1 and Q_H denote *fundamental reactive power* and *harmonic reactive power*, respectively.

Many authors claimed that the Budeanu's definition is not correct and cannot be used for calculating reactive power. However, this definition still occupies a significant number of pages on *The IEEE Standard Dictionary,* [3]. Its heritage is hard to dispute. Almost all contemporary textbooks written by appreciated scientists are to present Budeanu's definition of apparent power as the right canonical expression. More about calculating reactive power can be found in [3]

The vector sum of active and reactive power represents phasor power:

$$S = \sqrt{P^2 + Q^2} \ . \qquad (8)$$

However, this stands only for sinusoidal conditions. In presence of harmonics it is applicable to each harmonic component of active and reactive power separately [4]. Therefore, it will not be equal to the apparent power what applies in the case of sinusoidal condition. The difference reflects through the distortion power D. Consequently, the apparent power U (physically known as the product of RMS values of voltage and current) represents a vector sum of phasor power and distortion power [5]:

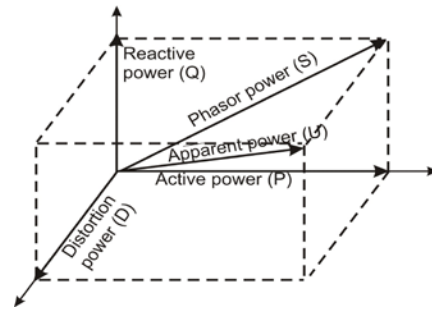$$U = I_{\text{RMS}} * V_{\text{RMS}} = \sqrt{S^2 + D^2} \qquad (9)$$

Fig. 1. Geometrical representation of relationship between active, reactive, phasor, distortion and apparent power, [5].

Fig. 1 illustrates relationship between active P, reactive Q, phasor S, distortion D and apparent power U, in monophase system with harmonic pollution. Obviously, Fig. 1 together with Eq. (8) and Eq. (9) express the fact that in unpolluted condition the distortion power will be equal to zero and apparent power U will be equal to phasor power S.

III. EFFECTS OF HARMONICS TO THE POWER SYSTEM

Electrical equipments, depending on the function they perform, react differently to harmonic distortions of the supply voltage. The distorted voltage has no effect to light bulbs but there is a large group of equipment which operating function relays on sine-wave voltage supply. Their best representatives are induction motors. Any deformation in voltage waveform introduces loss in form of increased coil temperature. Undoubtedly this reduces the life of the motor [6].

Besides, wide class of equipment that utilizes thyristor based control requires very precise supply voltage. Harmonic distortion may cause a malfunction of the apparatus.

The neutral line current in a three-phase power system may exceed the value of active power line. In single phase system the harmonic distortions raises a risk of overloading the neutral line. This usually causes:

- Overheating of neutral line, with reducing the life span of the conductor and with possibility to cause fire.
- High voltage between neutral line and ground can affect the operation of digital equipment and local area network (LAN), if the grounding is bad [9].

Harmonic distortions degrade power system characteristics and jeopardize all its components.

The distortion current causes additional heating of transformer and therefore reduces their lifespan. On the other hand, when distortion voltages are present in supply voltage for capacitor batteries, dielectric is overheating and threats to explode. Detail information about problems caused by harmonics can be found in [6], [7], [8].

IV. BEHAVIORAL POWER METER MODEL

Integrated power meters relay on digital signal processing of voltage and current samples. Therefore accurate modelling requires discrete-time models of all power components. Instantaneous value of current or voltage in time domain describe equation:

$$x(t) = \sqrt{2} X_{RMS} \cdot \cos(2\pi ft + \varphi) \ . \qquad (10)$$

After the discretization in equidistant time intervals it transforms to:

$$x(nT) = \sqrt{2} X_{RMS} \cdot \cos(2\pi \frac{f}{f_{sempl}} n + \varphi) \ , \qquad (11)$$

where f and f_{sempl}, are frequency of the signal and the sampling frequency. By definition the RMS value is:

$$X_{RMS} = \sqrt{\frac{\sum\limits_{n=1}^{N} x(nT)^2}{N}} \ . \qquad (12)$$

The active power is obtained as average of the instantaneous multiplication of instantaneous values for current and voltage, and average active power one gets in form:

$$P = \frac{\sum\limits_{n=1}^{N} v(nT)i(nT)}{N} = \frac{\sum\limits_{n=1}^{N} p(nT)}{N}. \qquad (13)$$

The same model is used for reactive power after voltage samples are shifted for π/2. Possible sources of error in active and reactive power calculation are caused with the phase difference between voltage and current and the fact that the power line frequency is slightly changed around the nominal (50Hz). These errors can be eliminated/diminished by additional calibration and correction within appropriate filters.

Once when P is calculated according to Eq. (13), Q calculated on similar way using shifted voltage samples, and U obtained as the product of RMS of voltage and current Eq. (9), one easily can compute distortion power as:

$$D = \sqrt{U^2 - P^2 - Q^2} \ . \qquad (14)$$

The previous equations represent bases for RTL model development of a power meter.

Firstly it was implemented in Matlab. Simple numerical integration of *P, Q, U, D* and *S* in time gives appropriate energies.

The part that calculates I_{RMS} (and V_{RMS}) has already been developed in LEDA laboratory for previous versions of our solid state power meter named IMPEG [10], [11].. It is presented in Fig. 2. Blocks denoted as I_{offset} and I_{gain} are used to compensate the errors described above.

In order to calculate *D*, the model has been modified as Fig. 3 presents. It consists of the multiplayer, the accumulator, the square-root and the finite-state-machine block.
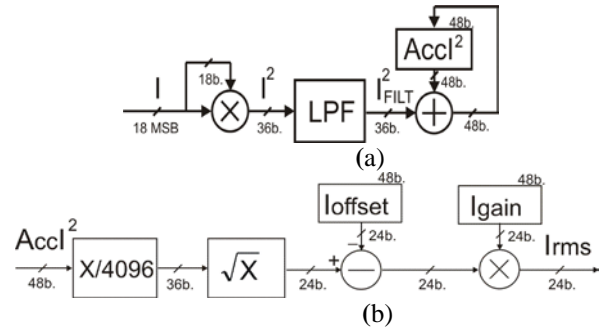

(a)

(b)

Fig. 2.Block diagram of the model for I_{RMS} (and V_{RMS}) calculation using Eq. (12)
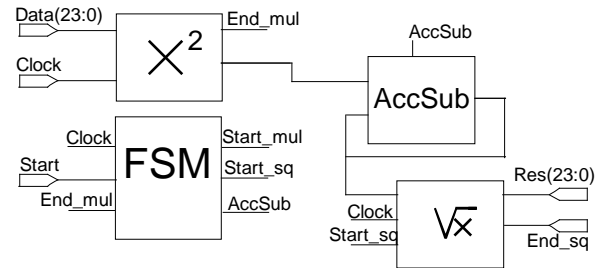

Fig. 3. Block diagram of model for calculating the distortion power

The model is based on multiply implementation of block diagram in Fig. 3. Practically model calculates I_{RMS} and V_{RMS} when $i(nT)$ or $v(nT)$ are supplied to Data input. Model for *P* differs only in feeding multiplier with both $i(nT)$ and $v(nT)$.

Model for *Q* has identical structure but it is fed with voltage samples displaced for π/2.

Apparent power samples are calculated directly applying Eq.(9).

The multiplier accepts samples of the apparent power *U* trough Data port. After 24 clock cycles required for 24-bit signal using Booth's algorithm the squared value of the apparent power U^2 appears and being stored. Thereafter the value of the active power is squared and the new value is subtracted from the value of apparent power. The same process is repeated for reactive power *Q*. Finally, the obtained value is sent to the input of the square root block

that provides distortion power D. FSM block provides control signals that schedule correct operation.

The model was confirmed by simulation. Moreover it has been verified on prototyped power meter realised by EWG electronics [12]The prototype has been developed using the electric power meter that was based on standard IC 71M6533 manufactured by MAXIM. Consequently it already provided I_{RMS}, V_{RMS}, P and Q. on the same manner as Eq. (12) and Eq. (13) describe. Thereafter U and D were calculated according to Eq. (9) and Eq.(14).

V. SIMULATION RESULTS

We used the developed model to simulate different types of non-linear loads. In order to approve both the method for distortion power calculation and the model we used measured data for currents published in [8], [13]. Fig. 4 illustrates the appropriate waveforms.

In order to simulate a possible realistic case, we supposed that the voltage is polluted, as well. Namely it has 3rd harmonic with amount of 3% in respect to the fundamental component. We considered eight different types of loads connected to the grid. There are:

  a) Incandescent light bulb (ILB)
  b) Heater (HR)
  c) Fluorescent lamp (FL)
  d) EcoBulb Compact Fluorescent Lamp (ECFL)
  e) Phillips Compact Fluorescent Lamp (PCFL)
  f) 6-pulse 3-ϕ diode rectifier dc power supply (3-DR)
  g) 6-pulse switched-mode power supply (SMPS)
  h) 6-pulse PWM controlled variable speed drive (PWM VSD)
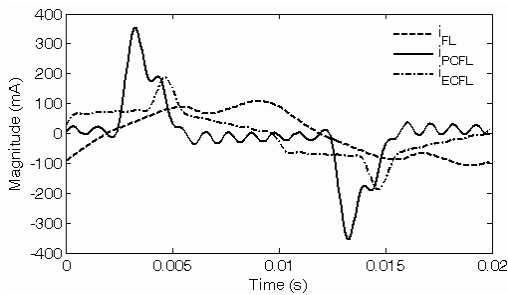

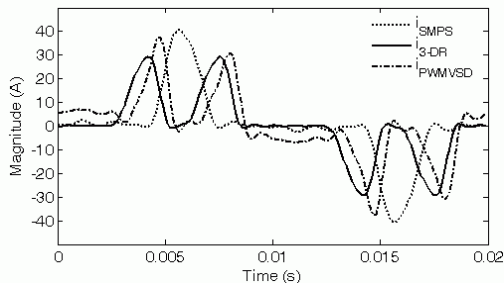Fig. 4.a. Current waveforms for Fluorescent lamps: FL, ECFL and PCFL

Fig. 4.b. Current waveforms for Rectifiers: 3-DR, SMPS and PWM VSD

The first two cases represent linear loads. Therefore the obtained current follows the voltage waveform. All other loads are nonlinear. Consequently they draw distorted current. Fig. 4.a illustrates currents of FL, ECFL and PCFL. Fig. 4.b presents waveforms of currents through 3-DR, 6-SMPS, and PWM VSD. Table I summarizes the obtained results.

TABLE I
SIMULATION RESULTS FOR DIFFERENT TYPES OF LOADS

	ILB	HR	FL	ECFL	PCFL	3-DR	SMPS	PWM VSD
I_{RMS}[A]	0.434	10.01	0.103	0.091	0.129	13.5	14.8	14.2
V_{RMS}[V]	230.3	230.3	230.3	230.3	230.3	230.3	230.3	230.3
P[W]	99.8	2305.0	17.32	18.49	15.85	2251.4	2183.9	2305.2
Q_B[VAR]	0	0	15.37	-5.99	-9.3	470.3	412.3	-8.96
U[VA]	99.8	2305.8	23.6	20.9	29.7	3115.4	3416.9	3277.1
D_B[VAR]	0	0	4.64	7.73	23.3	2101.4	2595.3	2329.3
D_B/P[%]	0	0	26.79	41.81	147.00	93.34	118.84	101.05
$(U-P)/P$ [%]	0.00	0.03	36.26	13.03	87.38	38.38	56.46	42.16
S [VA]	99.80	2305.0	23.16	19.44	18.38	2300.0	2222.4	2305.22
$(S-P)/P$ [%]	0.00	0.00	33.70	5.12	15.94	2.16	1.77	0.00

As expected, for both linear resistive loads, the active power, P, equals to the apparent power, U. Therefore the distortion power calculated using Eq. (14) equals zero. The other cases with non-linear loads should result with non-zero distortion power. The currents of each load are very rich with harmonics. Therefore, I_{RMS} increases proportionally to harmonics and consequently U and D_B, rise, as well

Cases FL, ECFL and PCFL represent small loads, with $P<20$W. Fig.4.a suggests that the current waveform of the PCFL is the most distorted. Hence, we expect to get greater D_B then in case of FL and ECFL. The result of simulation listed in the row D_B in Table I confirms the expectation.

Large loads (P greater of 2kW) have much higher impact to the grid and deserve more attention. Fig.4.b indicates that the SMPS is the biggest source of harmonic pollution. Simulation results confirm this anticipation.

Obviously, the measure of the distortion power is in direct relation with the nonlinearity of a particular load. Losses caused by nonlinear loads expressed as distortion power could range up to 147% relative to the active power. This was the case for PCFL. However because of the low nominal active power of 16W this is not a treat for the utility at the household level. One should concern more about larger loads like 3-DR, SMPS and PWM VSD. Table I presents that the amount of the distortion power is comparable with the active power. Moreover it is greater than 2kW.

It is interesting to estimate losses that utility has only due to the lack of registration of distortion and reactive power. Therefore we present the ratio of the difference between apparent and active power and the active power. Obviously PCFL represents the worst case with losses of 87%. However due to relatively small nominal active power this has neglecting effect to the power system. In contrary larger loads produce greater losses. Table I

indicates that the largest losses provide the device with the most high distortion level. Actually SMPS produces losses of 56% on the load of 2183W. On other two loads of comparable nominal power the utility has losses of 40%. These definitely are not negligible.

VI. DSP DEDICATED FOR DISTORTION POWER METERING

Block DSP represents a part of integrated power meter (IMPEG). Instantaneous values of current and voltage are obtained from digital filters and based on them on every second DSP calculates RMS value of current I_{RMS} and voltage V_{RMS}, active P, reactive Q, distortion D and apparent U power, power factor and frequency [10, 11]. Using value of active and reactive power DSP generate impulse for every Wh measured energy. This impulse increment register of DSP, that save information of active and reactive power (generated or consumed).DSP block work at 4.194 MHz and with accuracy less than 0.1% calculates all mentioned parameters. It accepts 16–bit wide inputs representing voltage, current and phase-shifted voltage samples from digital filters. Thereafter, it calculates already mentioned final power line parameters. Three sets of power line measurement results are obtained for different power line phases called R, S and T. The current input dynamic range is from 10 mA RMS to 100 RMS, while for voltage input it is up to 300V RMS. Results are represented within DSP by 24-bit 2's complement values.

DSP utilizes controller/datapath architecture which consist of several blocks: finite state machines, three static single port 64x24 bit Random Access Memories, datapath registers, arithmetical units for addition, subtraction, division, square rooting, multiplication and other digital blocks. Digital blocks can be divided into five main groups (Fig. 1):

1. Frequency measurement circuit
2. RAM memory block
3. Part for I2, V2, P, Q accumulating and energy calculation
4. Part for current and voltage RMS, active, reactive , apparent and distortion power and power factor calculation);
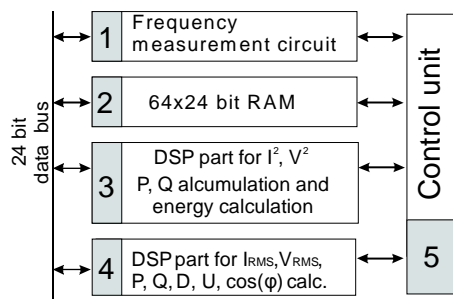5. Control unit that manages all other parts of DSP.

Fig.5 DSP block diagram

There is a single 24-bit data bus connecting these subblocks of DSP. The control path of DSP unit is implemented as a finite state machine and it generates a number of control signals that determine what component can write to 24-bit data, what registers are loaded from the bus and what arithmetical operation is performed. Controller performs the periodically repeated sequence that lasts exactly 1024 clock periods which is divided into four 256 clock period subsequences. The first three FSM subsequences are called R, S and T and they control the calculations made for each phase of the three-phase energy system. During R, S and T subsequences intensive calculations are performed only within subpart 3 (Fig.4). More detail about architecture of DSP can be found in [10] and [11].

VII. CONCLUSION

This paper presented a model for distortion power calculation. Simulation of six nonlinear and two linear loads verified the model. Moreover the results indicated that utility suffers large losses due to the lack of registering distortion power. Actually the utility in Serbia and in the greater part of the world relays billing only on active power measurements. However, for real cases of large nonlinear loads the losses overcome 50% in comparison to the active power. Having in mind that the number of nonlinear loads rapidly increases this amount rises with no visible ending in the near future. Therefore this seems to be a tremendous problem. It is interesting that some of developed countries recently has invested a lit of money to replace old power meters capable to measure only active power with the new that measure reactive power as well. According to [1] Italy distributor has decided to install more than 20 million household energy meters with active and reactive power measurement. However without measuring distortion power all this all of this misses the point. Two bottom rows in Table I represent the phasor power S (Eq. (8)) and the according losses calculated in respect to active power. The characteristic case is 6-pulse PWM controlled variable speed drive with the nominal active power of 2300W with small reactive power but with large distortion power. The relative loss regarding the phasor power is almost zero, but regarding the apparent power is 42%.

Obviously the utility will be faced very soon with a tremendous problem with losses if does not start to measure all power components. Some may argue that it is sufficient to register only the apparent power. However, this would be step back because contemporary power meters are capable to measure all components. As we have recently published [14] measuring distortion power at PCC helps the utility to determine the source of nonlinear pollution at the grid. Therefore it would be capable to bill separately every component of power. In our opinion this will cut the losses but will serve as mighty tool to manage the loading profile of the consumers.

ACKNOWLEDGEMENTS

7. REFERENCES

[1] E. Moulin, "*Measuring Reactive Power in Energy Meters*", Metering International, 2002

[2] "The Future of the Electric Grid," Massachusetts Institute of Technology,2011

[3] A. E. Emanuel, "Power Definitions and the Physical Mechanism of Power Flow" ", *Wiley*, 2010

[4] "IEEE Standard Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Nonsinusoidal, Balanced, Or Unbalanced Conditions", *IEEE Std 1459-2010.*

[5] J. G. Webster, "*The measurement, instrumentation, and sensors handbook*", IEEE Press, 1999.

[6] Y. Alhazmi, "*Allocating power quality monitors in electrical distribution systems to measure and detect harmonics pollution*", Electronic Theses and Dissertations, University of Waterloo, Ontario, Canada, 2010

[7] Singh G.K.: "*Power system harmonics research a survey*", European Transactions on Electrical Power, Vol.19, 2007, pp. 151–172.

[8] Wakileh. J. G., "*Power Systems Harmonics*", Springer, 2001

[9] T. Shaughnessy, *"Clearing Up Neutral-to-Ground Voltage Confusion*", Electrical Construction & Maintenance, February 1, 2007.

[10] B. Jovanovic, M. Damnjanović, P. Petković, "*Digital Signal Processing for an Integrated Power Meter*"**,** Proceedings of 49. Internationales Wissenschaftliches Kolloquium, Technische Universirtat Ilmenau, Ilmenau, Germany, vol. 2, pp. 190-195, September 2004

[11] B. Jovanović, M. Damnjanović, "Digital Signal Processing in three-phase Integrated Power Meter"**,** *Proc. of the 52th ETRAN conference*, Palić, June 2008, EL2.3-1-4.

[12] EWG - multi metering solutions, www.ewg.rs

[13] Z. Wei, "*Compact Fluorescent Lamps phase dependency modelling and harmonic assessment of their widespread use in distribution systems*", Electronic Theses and Dissertations, University of Canterbury, Christchurch, New Zealand, 209

[14] Stevanović, D., Petković, P.: "*A New Method for Detecting Source of Harmonic Polution at Grid*", Proc. of 16th International Symposium Power Electronics Ee2011, Novi Sad, Serbia, 26.10.-28.10., 2011, T6-2.9 pp. 1-4, ISBN 978-86-7892-356-2

AUTHOR INDEX

* First author