

SIMULTANEOUS OPTIMIZATION OF STANDBY AND ACTIVE ENERGY FOR SUB-THRESHOLD CIRCUITS

Ali T. Shaheen and Saleem M. R. Taha

Department of Electrical Engineering, University of Baghdad, Iraq

ABSTRACT

Increased downscaling of CMOS circuits with respect to feature size and threshold voltage has a result of dramatically increasing in leakage current. So, leakage power reduction is an important design issue for active and standby modes as long as the technology scaling increased. In this paper, a simultaneous active and standby energy optimization methodology is proposed for 22 nm sub-threshold CMOS circuits. In the first phase, we investigate the dual threshold voltage design for active energy per cycle minimization. A slack based genetic algorithm is proposed to find the optimal reverse body bias assignment to set of non-critical paths gates to ensure low active energy per cycle with the maximum allowable frequency at the optimal supply voltage. The second phase, determine the optimal reverse body bias that can be applied to all gates for standby power optimization at the optimal supply voltage determined from the first phase. Therefore, there exist two sets of gates and two reverse body bias values for each set. The reverse body bias is switched between these two values in response to the mode of operation. Experimental results are obtained for some ISCAS-85 benchmark circuits such as 74L85, 74283, ALU74181, and 16 bit RCA. The optimized circuits show significant energy saving ranged (from 14.5% to 42.28%) and standby power saving ranged (from 62.8% to 67%).

KEYWORDS

Dual Threshold Design, Slack Based Genetic Algorithm, Sub-threshold Circuits, Reverse Body Bias, Standby Power

1. INTRODUCTION

Modern digital CMOS circuits explore dramatically leakage power increasing with each generation. These technology scaling problems have recently made power reduction an important design aspect. The well-known CMOS sub-threshold circuits are considered candidate solution for energy constrained-low speed applications since it can significantly reduce energy per cycle (EPC) due to its very low supply voltage. However, these circuits suffer from the large delay which increases exponentially due to the exponential relation between the sub-threshold current and the supply voltage whereas the delay in above threshold operation increases according to α -power law [1].

The sub-threshold operation is a weak inversion mode depends on the sub-threshold current as the main source of current. This current is summarized in Equation (1)

$$I_{\text{leakage}} = I_0 \exp\left(\frac{V_{gs} - V_{th}}{nV_T}\right) \left[1 - \exp\left(-\frac{V_{ds}}{V_T}\right)\right] \quad (1)$$

where

$$I_0 = \mu C_{ox} \frac{W}{L} (n-1) V_T^2 \quad (2)$$

μ is effective mobility, C_{ox} is oxide capacitance, W is transistor width, L is transistor length, V_{ds} is drain-source voltage, V_{gs} is gate-source voltage, V_T is the thermal voltage, V_{th} is threshold voltage and n is the sub-threshold slope, which is a technology determined parameter [2, 3]. According to this weak inversion operation, the power reduction techniques should be used with special constraints for sub-threshold circuits due to its effects on the delay.

Unlike above threshold circuits, the EPC (power-delay product) in sub-threshold circuits may be very high even when power consumption is small because of the large delay of such circuits. Therefore the EPC is the key parameter that will determine whether this design is good or not [4]. This is demonstrated in Equation (3) which gives the approximate EPC for N-gates circuit. The 1st term of Equation (3) represents the dynamic energy whereas the 2nd term represents the static or leakage energy.

$$E = \sum_{i=1}^N 0.5 \alpha(i) \cdot C(i) \cdot V_{dd}^2 + P_{\text{leak}}(i) \cdot T \quad (3)$$

where $\alpha(i)$ is the switching activity of the i^{th} node, $C(i)$ is the capacitance of the i^{th} gate, $P_{\text{leak}}(i)$ is the leakage power of the i^{th} gate and T is the critical path delay.

Many researches were proposed for power optimization in sub-threshold circuits. Dual supply sub-threshold circuit design work was presented by a research group at Auburn University [5-9]. In [6], the authors developed a slack based algorithm to achieve maximum energy per cycle saving. Mixed integer linear programs (MILP) were developed in [8] to optimally assign dual supply voltages to sub-threshold circuits in such a way to eliminate the required level converters. Ultra-Dynamic Voltage Scaling (UDVS) was proposed for sub-threshold operation in [10]; the power saving was achieved through the combination of dithering at high performance and minimum energy operation for low performance scenarios.

On the other hand, dual- V_{th} design can be considered as a common method for reducing leakage power consumption in above-threshold and sub-threshold operations [4]. The relevance of dual threshold design for sub-threshold circuits stems from the use of high threshold voltage for some gates on the off critical paths to reduce power consumption without affecting the critical path delay.

In above threshold circuits, the dual threshold approach was proposed for power optimization with various assignment algorithms under some constraint formulas [11-14]. For example in [12] the researchers attempted to use linear programming (LP) to minimize the power under some constraints on circuit speed, gate slack, delay and cell size. In [4] the dual threshold approach was proposed for 32nm sub-threshold circuits using slack based heuristic algorithm. The author showed the effectiveness of dual threshold voltage design for 32 nm sub-threshold circuits where significant energy saving has been shown. However, these researches ignore the leakage power in the standby mode of operation. The power optimization in this mode may be effective in some applications when the device spends long periods idle. In [15], a standby energy analysis and optimization technique were proposed for low supply applications. The interaction of the optimal energy, cut-off structures, and supply voltage were investigated.

In this paper, we propose a simultaneous active and standby modes energy optimization technique for 22nm sub-threshold circuits with non-zero standby applications. Firstly, the dual threshold voltage design is investigated for active mode optimization. An optimal reverse body bias (RBB) assignment is achieved using the proposed slack based genetic algorithm (SBGA) to ensure maximum EPC saving with maximum allowable operation frequency at the optimal supply voltage (V_{ddopt}). Secondly, the optimal RBB for standby mode operation is evaluated to obtain minimum standby power at the previously determined (V_{ddopt}) for the active mode optimization. A Variable threshold technique is used to simultaneously combine the two optimization phases

2. REVERSE BODY BIAS EFFECT

In CMOS circuits it is common to adjust the threshold voltage of MOSFET transistors by applying a source-bulk bias voltage as given in Equation (4)

$$V_{th} = V_{th0} + \gamma(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{2\phi_F}) \quad (4)$$

where V_{th0} is the zero body bias threshold voltage, V_{SB} is the source-bulk bias voltage, $2\phi_F$ is the surface potential parameter, and γ is the body effect parameter. The RBB technique increases the threshold voltage of a MOSFET by applying a negative voltage across the source-to-substrate p-n junction whereas zero RBB is used in the structure of low V_{th} logic gates as demonstrated in Figure 1. In Figure 1.a, a low V_{th} two input NAND gate is constructed with zero RBB by grounding the fourth terminal of NMOS transistor and connecting the fourth terminal of PMOS transistors to V_{dd} . The same gate is constructed in Figure 1.b with non-zero RBB to increase the threshold voltage [4, 16].

Although increasing the RBB voltage across the source-to-substrate p-n junction of a MOSFET increases the threshold voltage, and consequently reducing the sub-threshold leakage current; the RBB also increases the tunnelling leakage current at the reverse biased source-to-body and drain-to-body p-n junctions.

When the RBB voltage is increased, both the bulk and surface band-to-band tunnelling current components increase but the sub-threshold leakage current decreases. Therefore, there exist an optimum RBB voltage (specific to a process technology) that minimizes the overall leakage power consumption [17].

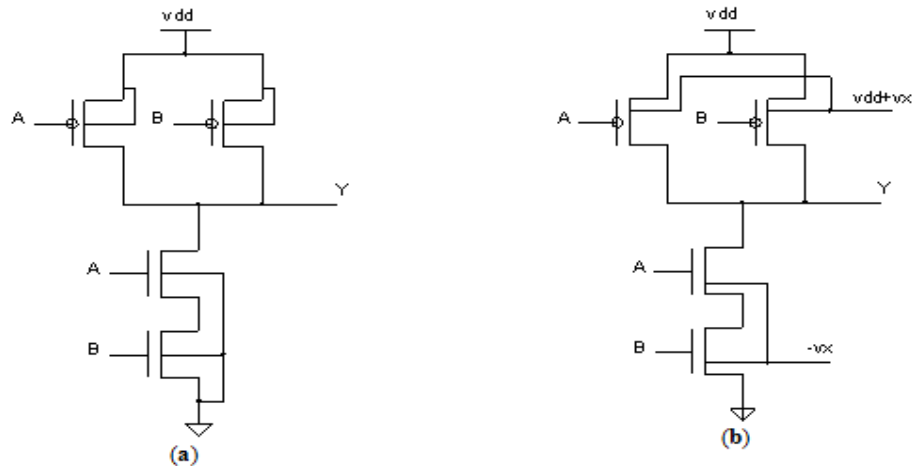


Figure 1. Two input NAND gate: (a) Low V_{th} (RBB=0); (b) High V_{th} (RBB= V_x)

3. CIRCUIT MODEL AND STATIC TIMING ANALYSIS

The first step in our design methodology is to classify the logic gates of the circuit according to gate type and no. of inputs such as (AND2, AND3, AND4, OR2, etc.) all these gates are considered with different fan-outs to cover all cases in the circuit. HSPICE is used to perform an accurate simulation for each gate type in order to obtain a power-delay-capacitance library (PDC) for all gates as shown in Figure 2. The library contains leakage power as a function of (Vdd and RBB), gate delay as a function of (Vdd, RBB and fan-out or node capacitance) and node capacitance as a function of (Vdd and Fan-out). The assignment algorithm depends on the library content at each (Vdd and RBB values) to calculate and compare the overall power and delay.

The simulations must be performed under different values of the supply voltage (Vdd). For each gate, the lower delay (DL) value is obtained with (RBB=0) design, whereas the higher delay value (DH) occurs with specific RBB for a given Vdd and fan-out values. Depending on these values our algorithm will optimally select some gates of the non-critical paths to be switched from low threshold voltage design (RBB=0) to high threshold voltage design (RBB≠0) without increasing the critical path delay of the circuit at the specific supply voltage.

Secondly, the combinational gate-level circuit is modeled as Directed Acyclic Graph (DAG), $G=(V, E)$ as shown in Figure 3. The logic gates represented as nodes or vertices within a set (V) and the connections between nodes are represented as directed edges within edges set (E) from the inputs to the outputs. Assign the first node to the Primary Inputs (PI) and the last node to the Primary Outputs (PO).

Usually, static timing analysis is used to provide timing information and to calculate the actual arrival time (AAT) and required arrival time (RAT) at the output of the gates. Also, STA is used to find the critical path delay (T) and the slack of each gate (slk). The gate slack is defined as the difference between the critical path delay and the maximum path delay through this gate; so it represents the delay margin for each gate. For correct operation of the chip with respect to setup (maximum path delay) constraints, it is required that $AAT(v) \leq RAT(v)$ [4, 18].

In this paper, we use the APAC (All paths and cycle) Algorithm proposed in [19] to get all paths between any two nodes in a given DAG. Then by using the delay contents of the PDC library and the APAC results we can find the arrival time of each node, critical path delay and the slack of each node, as in the STA algorithm given in Table 1.

All paths from node (1) (primary inputs node) to any node (v) (primary outputs node) are given in Step 4 of Table 1. These paths are represented by the indices of nodes along each path.

Consequently, after substituting the delay of each node from PDC library which may be DL or DH according to the algorithm assignment, the maximum delay from node (1) to node (v) is obtained. Similarly, the maximum delay from node (v) to node (N) (which is the primary output node) is determined in step 9. In step 10 the longest path through each gate is calculated depending on the previous results. The critical path delay of the overall circuit and the slack of each gate are calculated in steps 11 and 12 respectively.

Finally, it is required to estimate the switching activity at each node to complete the circuit model in order to evaluate and compare the EPC according to Equation (3). Probabilistic estimation is used with the help of DAG analysis to evaluate the switching activity at each node, where the roots of the inputs for each gate can be determined.

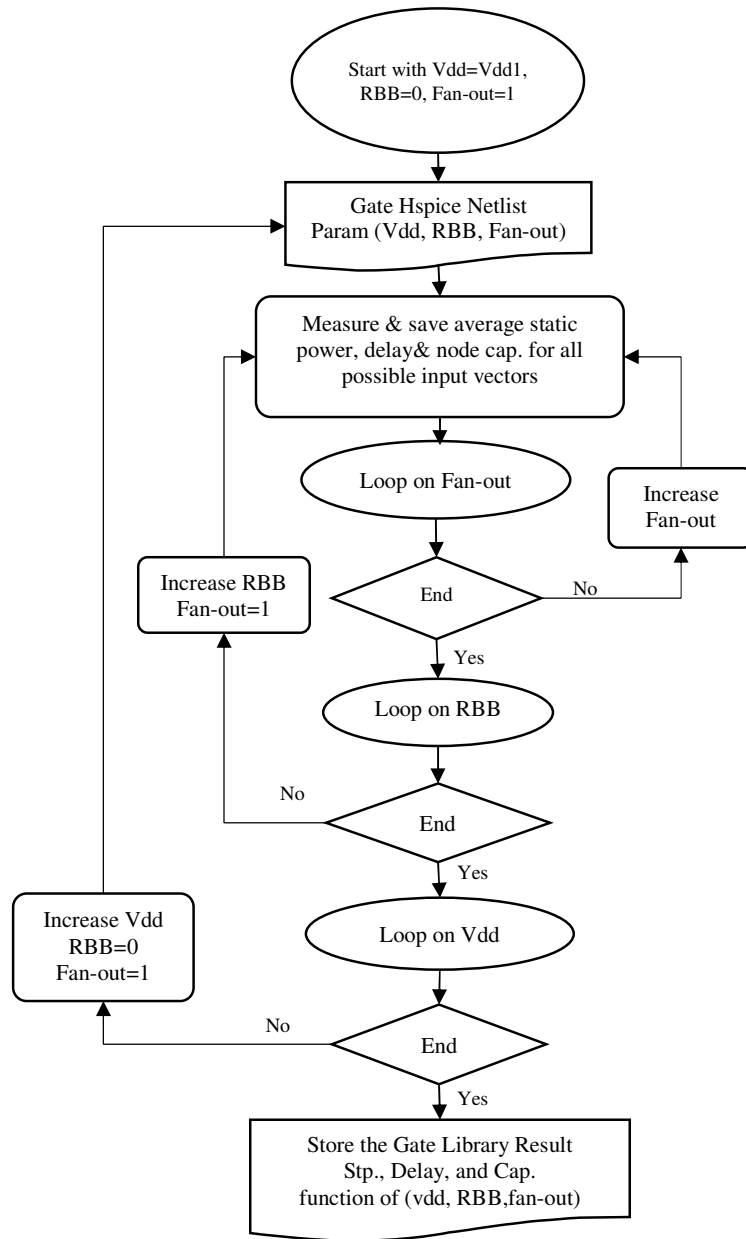


Figure 2. Flow chart of PDC library simulation for each gate

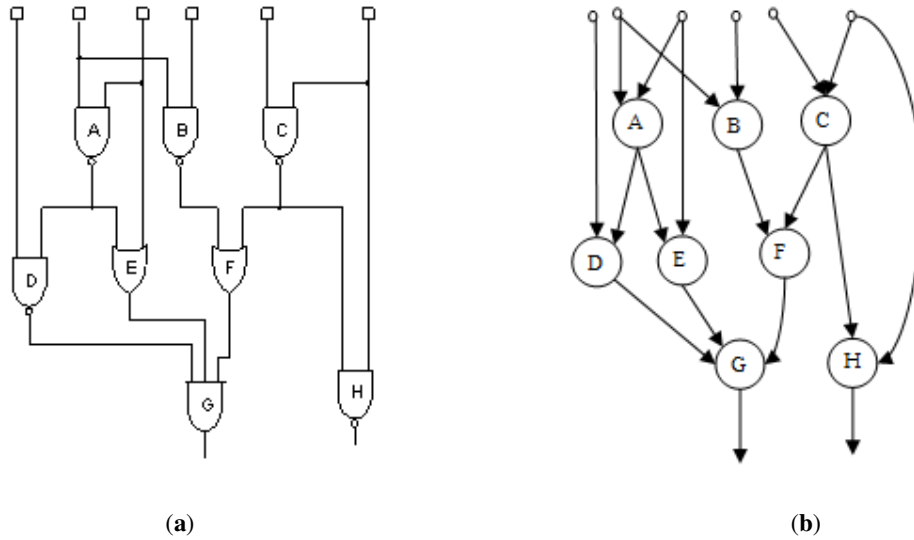


Figure 3. DAG of simple circuit: (a) Gate level representation; (b) DAG

Table 1. STA algorithm

1.	Input: directed graph $G=(V, E)$ and PDC library: where V is set of circuit nodes $V=\{v_1, v_2, \dots, v_N\}$
2.	Output: T : Critical Path delay and slk: slack of each gate
3.	\forall node $v \in V$ find:
4.	$gpi\{v\} = APAC(G, 1, v)$; all paths from node 1 to node v
5.	$gpo\{v\} = APAC(G, v, N)$; all paths from node v to node N
6.	use $gpi\{v\}, gpo\{v\}$ and PDC Library to find:
7.	$D(v)$; delay of node v
8.	$TPI(v)$; the longest time from node 1 to node v ,
9.	$TPO(v)$; the longest time from node v to node N
10.	$DP(v) = DPI(v) + TPO(v) + D(v)$; the longest path through node v
11.	$T = \max\{DP(v)\}$ where $v = 1 \sim N$; critical path delay
12.	$slk(v) = T - DP(v)$; slack of each gate

4. PROPOSED METHODOLOGY

4.1. Phase 1- Active Mode Energy Optimization

In this phase, a dual threshold voltage design is applied to reduce the active mode EPC of sub-threshold circuits in the 22nm scale. A framework is presented to find the optimal supply voltage and the optimal dual threshold voltages using the proposed SBGA given in Table 2. This energy minimization methodology depends on the optimal assignment of high threshold voltage (RBB $\neq 0$) for some gates on the off-critical paths. The genetic algorithm is modified through the use of STA given in Table 1 in such a way to assign dual threshold voltages to gates in sub-threshold circuits without increasing the critical path delay at V_{ddopt} . As seen from Table 2, this algorithm combines the original structure of the genetic algorithm with the STA algorithm given in Table 1.

As given in the previous section, each gate is designed with low threshold voltage transistors (with $RBB=0$) and with high threshold voltage transistors (with $RBB \neq 0$) (as given in Figure 1 for NAND2). The PDC library is evaluated for all gates under different values of fan-out, RBB and supply voltage as given in Figure 2. Then the SBGA is automatically assign V_{th} -low design to critical paths gates to maintain the performance. Consequently, SBGA also finds an optimal set of gates on the off-critical paths that can be assigned to V_{th} -high design without any performance degradation.

The first step in the proposed assignment algorithm (shown in Table 2) is to assign V_{th} -low design for all gates so that the delay of any node $D(v)$ will be $DL(v)$. Through step 4 of Table2, STA algorithm (given in Table 1) is applied to find the critical path delay (T) and the slack of each gate $slk(v)$. The critical path delay of this design at the specified V_{dd} is used as the timing requirement for the dual threshold voltage design. In step 5, the V_{th} -high design is assigned for all gates so that the delay of any node $D(v)$ will be $DH(v)$. STA algorithm is applied again to find the new critical path delay (T_{high}) and delta value for each gate $\Delta(v)$ which equal to $DH(v) - DL(v)$. The upper and lower slack bounds used in the optimization are calculated in step 7. A set of nodes that must be kept at V_{th} -low design is determined in step 8. The characteristic of these nodes is to have small slack time less than the slack lower bound (SL). When any gate contained in $\{Low_Set\}$ is switched from V_{th} -low to V_{th} -high design, the delay of the gate is converted from DL to DH and therefore gate slack will become negative as illustrated below in Equation (4). However, our algorithm does not allow negative slack because negative gate slack means a performance degradation occurred, where longest path delay exceeds the original circuit delay T .

Table 2. Slack Based Genetic Algorithm (SBGA)

-
1. $\forall V_{dd} \in \{V_{dd_set}\}$ do
 2. $\forall RBB_j \in \{RBB_set\}$ do
 3. Assign $RBB=0$ design for all gates
 4. Apply STA algorithm and find: T , delay and slack of each gate $DL(v)$ and $slk(v)$:
 5. Assign $RBB=RBB_j$ for all gates
 6. Apply STA algorithm an find: T_{high} , and delay of each gate $DH(v)$
 7. Calculate: $k = T_{high}/T$; $\Delta(v) = DH(v) - DL(v)$;
slack upper bound $SU = (k-1).T/k$ and slack lower bound $SL = \min \{ \Delta(v) \}$
 8. If $slk(v) \leq SL$ then $v \rightarrow \{Low_Set\}$ the indices set is $\{idx1\} = 0$'s
and no. of nodes in the set = $N1$
 9. If $slk(v) \geq SU$ then $v \rightarrow \{High_Set\}$ the indices set is $\{idx2\} = 1$'s
and no. of nodes in the set = $N2$
 10. The remaining nodes grouped in a set $\{Rem_set\}$ with indices set is $\{idx3\}$
and no. of nodes in the set = $N3 = N - N1 - N2$
 11. Apply GA to $\{Rem_set\}$: to select which nodes assigned to RBB_j (V_{th} -H)
 12. Initialization: Select iterations no. , cross over prob., mutation prob. and population size
 13. Pop: Initial population with random zeros and ones ($M, N3$) ; M : no. of chromosomes
 14. \forall chromosome $M_x \in pop$
 15. Mix the three sets : $\{Low_Set\}$, $\{High_Set\}$ and $\{Rem_set\}$
 16. Apply STA algorithm: find inst. critical path delay T_x ,and energy E_x
 17. Calculate fitness: $E_x, T_x = \left(\sum_{i=1}^N 0.5 \cdot \alpha(i) \cdot C(i) \cdot V_{dd}^2 + P_{leakage}(i) \cdot T_x \right) \cdot T_x$
 18. Apply selection: select the chromosomes of minimum fitness
 19. New generation pop_new : Apply cross-over and mutation
 20. $\forall M_x \in pop_new$: repeat steps (14-18) until stop
 21. End RBB loop
 22. End - V_{dd} loop

Another set $\{High_Set\}$ is determined in step 9 to contain all nodes that can be switched directly to (V_{th} -high) design without affecting the critical path delay. The characteristic of these nodes is to have large slack time greater than the slack upper bound (SU) which determined in step 7. So,

this assignment will still keep its slack non-negative, as illustrated below in Equations (5) and (6). This step will reduce the complexity of the genetic algorithm where the unique solution (i.e. same critical path delay of the original circuit) can be easily obtained.

Let us consider, $slk^*(v)$ is the updated gate slack and $DP^*(v)$ is the updated longest path delay of gate v after it is switched from V_{th} -low design to V_{th} -high design. For the gates with gate slack $slk(v) < SL$,

$$slk^*(v) = T - DP^*(v) = T - [DP(v) + \Delta(v)] = slk(v) - \Delta(v) \quad (4)$$

Equation (4) will be negative since we define SL to be $slk(v) < SL \leq \Delta(v)$.

As for the gates whose gate slack $slk(v) \geq SU$,

$$slk^*(v) = T - DP^*(v) = T - k \cdot DP(v) = T - k \cdot [T - slk(v)] \quad (5)$$

From the definition of SU , we know that $slk(v) \geq SU = (k-1) \cdot T/k$,

$$T - k \cdot [T - slk(v)] \geq T - k \cdot T + k \cdot SU = 0 \quad (6)$$

Therefore, Equation (5) remains non-negative after V_{th} -high assignment. Here an approximation is used that $k = T_{high}/T \approx DP^*(v)/DP(v)$.

The remaining nodes are grouped in a new set $\{Rem_set\}$, with nodes indices $\{idx3\}$. Some of the nodes contained in this set are optimally assigned to V_{th} -high design through the application of the modified genetic algorithm. The chromosome length should be the same as the number of nodes contained in $\{Rem_set\}$. Binary encoding is used to represent the nodes in each individual. (0) is used for V_{th} -low design assignment and (1) is used for V_{th} -high design assignment.

For each individual which represents the nodes of $\{Rem_set\}$, all above three sets are mixed according to their node indices in order to generate a full vector that represents nodes indices of the overall circuit. Depending on this vector and through the use of PDC library the instantaneous EPC (Ex) is calculated. Also, the instantaneous critical path delay (Tx) is evaluated in step 16, where STA algorithm is applied again and the delay of each node is $D(v)$ which may be $DL(v)$ or $DH(v)$ according to the gate index (v) in the final vector. These two objectives (Ex and Tx) are multiplied to generate the fitness value as in step 17. A new generation is obtained after the selection, cross-over and mutation steps. The genetic algorithm steps are repeated until the minimum fitness value is obtained. As a result, the maximum delay after the dual threshold voltage assignment does not exceed T (the critical path delay of the original circuit at the specific supply voltage) and the optimized EPC is determined. This process repeated for all RBB values and the presented framework find the optimal RBB value and optimal supply voltage operation.

4.2. Phase 2- Standby Leakage Power Minimization

As reported in the literature, the leakage power can be significantly reduced with the application of reverse body bias voltage. Hence, a high reverse body bias can be applied to all gates when a standby signal is detected. As a result, the threshold voltages of all gates are increased, and therefore a significant leakage power reduction can be obtained. However, as given in section 2, not all RBB values are allowed. The RBB voltage increasing results in sub-threshold leakage current decreasing but both the bulk and surface band-to-band tunneling current components increase. So, there exist an optimal RBB voltage for maximum standby leakage power saving. The PDC library contains the static power for each gate type at different RBB and V_{dd} values.

Hence, the static power of the overall circuit is evaluated at the optimal supply voltage determined from phase1 as given in Equation (7). The calculations are performed for all RBB values and the optimal RBB value for standby mode is the one corresponding to the minimum standby power. This RBB should be applied for all gates during the standby interval.

$$P_{standby} = \sum_{i=1}^N P_{leak}(i) \quad (7)$$

where $P_{leak}(i)$ is the leakage power of the i^{th} gate.

From the active mode optimization phase, three design parameters are determined; the optimal set of gates are to be modified (set A), the RBB bias value for these gates (V_x) and V_{ddopt} . The other gates (set B) should be assigned to $RBB=0$. On the other hand, the optimal RBB voltage for minimum standby power consumption (V_y) is found through phase 2 at the previously determined V_{ddopt} . This RBB value (V_y) is assigned for all gates when a standby signal is detected. Therefore, there are two sets of gates and two RBB values are determined for each set of gates one for active mode and the other for standby mode of operation as shown in Figure 4 which demonstrate the variable threshold voltage design of an inverter CMOS gate.

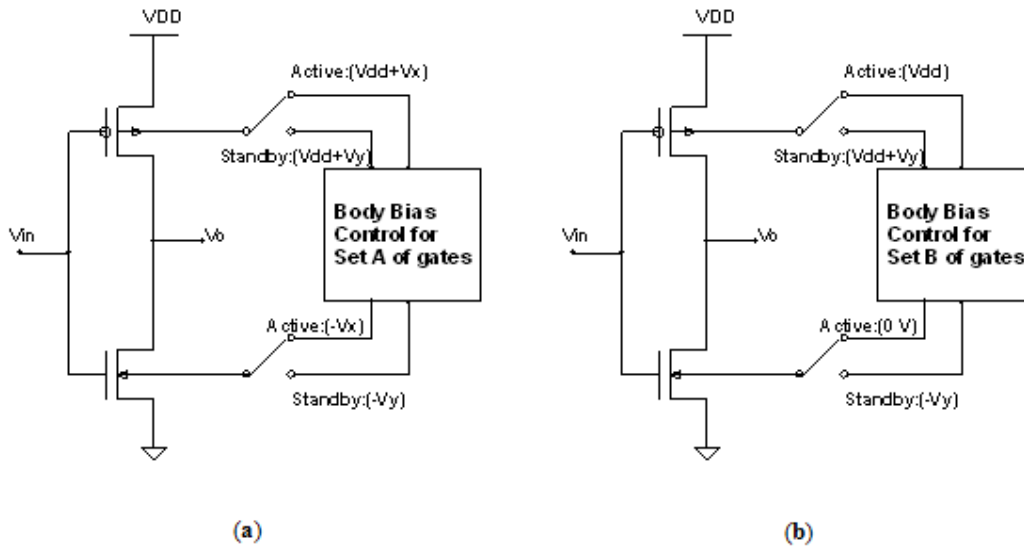


Figure 4. Variable threshold CMOS inverter: (a) From set A of gates; (b) From set B of gates

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, a simulation details and experimental results are given. The transistor model used is PTM 22nm bulk CMOS technology with $W=5L$ for NMOS transistors and $W=12L$ for PMOS transistors. HSPICE simulator is used to simulate all gates with and without RBB under different values of V_{dd} and fan-outs. Hence, the gate data library for each gate in the simulated circuits is obtained. The framework finds the optimal supply voltage and optimal reverse body bias (RBB) for minimum EPC. Optimal dual threshold assignment is achieved through the use of the proposed SBGA to ensure operation at the maximum allowable frequency at (V_{ddopt}). The supply voltage used in our simulations is ranged from 0.25 V to 0.4 V in step of 0.01 V (i.e. 16 values of supply voltage are used), and RBB voltage is ranged from 0.1 V to 1 V.

Experimental results of dual Vth design and single Vth-low design are obtained and compared for some benchmark circuits (16bit RCA, 74L85, 74283 and ALU7418) where the amount of energy saving that achieved by dual threshold assignment is influenced by the circuit topology. The EPC simulations of these circuits are shown in Figure 5, Figure 7, Figure 9, and Figure 10 respectively. Table 3 gives more detailed analysis, where (optimal EPC, Vddopt, RBB, operation frequency, and percentage of Vth-high gates) are given for all circuits. For each (Vdd and RBB combination) the algorithm is applied to select some gates of the non-critical paths to be switched from Vth-low to Vth-high through the application of RBB to these gates. EPC is calculated for the new design at the specific RBB for all Vdd values. Each RBB results in one curve as shown in EPC Figures and there exist a minimum EPC at an optimal Vdd. The lowest energy occurs at Vddopt of the lowest energy curve, i.e. the optimal Vdd-optimal RBB.

Maximum active mode energy saving (42.28%) is achieved for the 16 bit RCA at optimal RBB (0.7 V) and Vddopt (0.3 V) which corresponds to (10.91 MHz) clock frequency. The 16 bit RCA circuit explores a large number of short off-critical paths and the path length difference between the critical path and off-critical paths is sufficiently large. Therefore, it allows more gates (60.2% of total gates) to be assigned to 0.7V RBB. Gate slack distribution of 16 bit RCA circuit with and without optimization is shown in Figure 6. The non-optimized 16 bit RCA circuit has an optimal EPC (4.29 fJ) occurs at Vddopt (0.32V) which corresponds to clock frequency (12.041 MHz). From Figure 5 we can find that EPC of the optimal dual threshold design is (2.515 fJ) at Vdd (0.32V), so there exist a significant EPC saving (41.37%) even for the same operation frequency of the non-optimized circuit (12.041 MHz). Similar aspect can be noted for all simulated circuits. Alternatively, we can design the circuits to operate with lower energy saving at Vdd greater than Vddopt and improve the performance where this aspect is required for some applications especially for large circuits or circuits that have long critical paths. So, the critical path delay is large and resulting in low operation frequency.

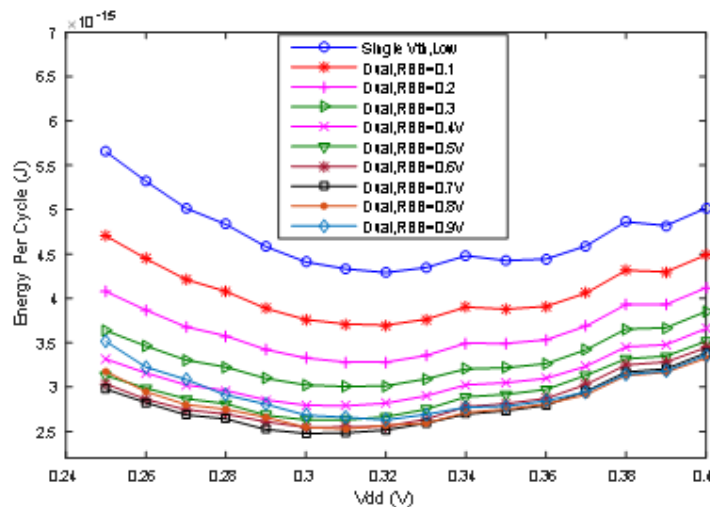


Figure 5. EPC of 16 bit RCA circuit for single and dual Vth design

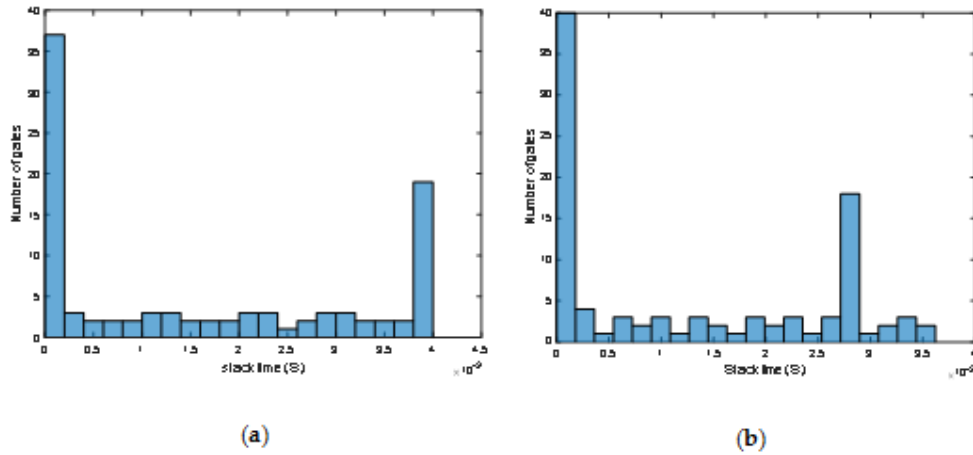


Figure 6. Gate slack distribution of 16 bit RCA: (a) Single Vth-Low (RBB=0) at Vdd= 0.3V; (b) Dual- Vth at Vdd= 0.3V and RBB=0.7V

The lower bound of active energy saving (14.58%) is achieved for the 74L85 circuit at Vddopt (0.31V) since its balanced structure result in a large number of candidate gates with small slack time and few gates with relatively large slack time as shown in Figure 8. So, the algorithm selected the RBB (0.2V) as an optimum case which considered suitable for all candidate gates. This is due to the proportional relation between the gate delay and the applied RBB value. However, RBB of 0.2V result will optimally exploit the available slack time but the energy saving will be small as compared with other circuits.

The other two circuits show EPC saving of (21% and 15.27%) for 74283 and ALU74181, respectively, as given in Table 3 where more details can be found.

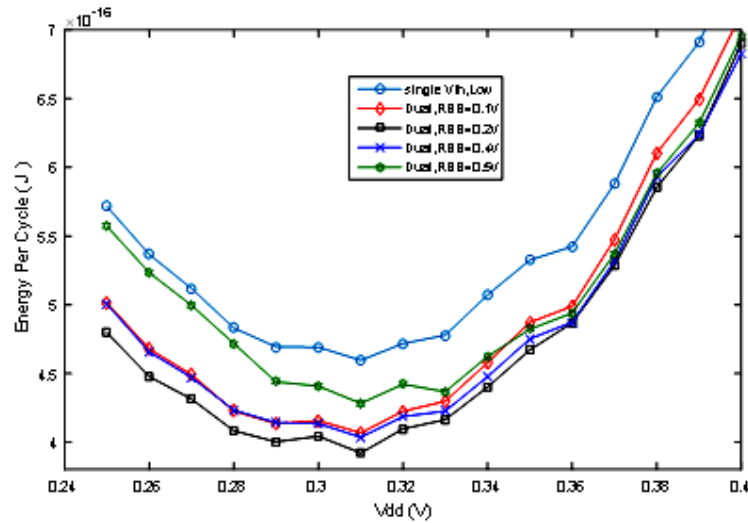


Figure 7. EPC of 74L85 circuit for single and dual Vth design

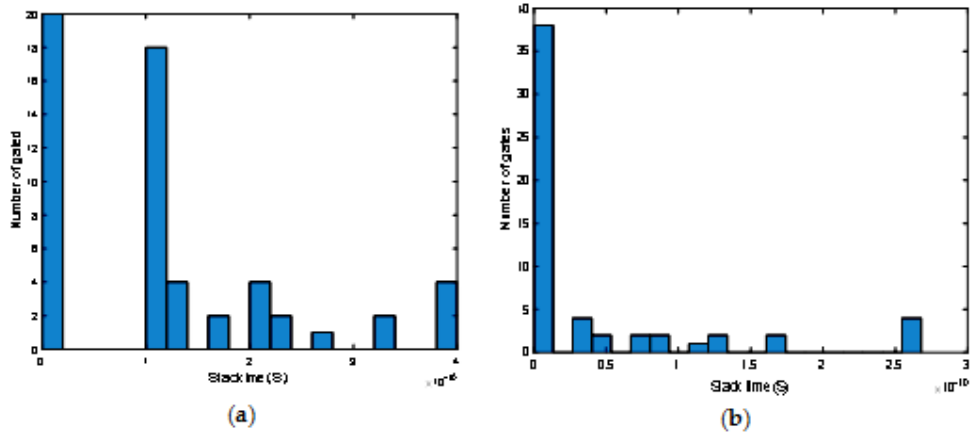


Figure 8. Gate slack distribution of 74L85: (a) Single Vth-low (RBB=0) at Vdd= 0.31V; (b) Dual Vth at Vdd= 0.31V and RBB=0.2V

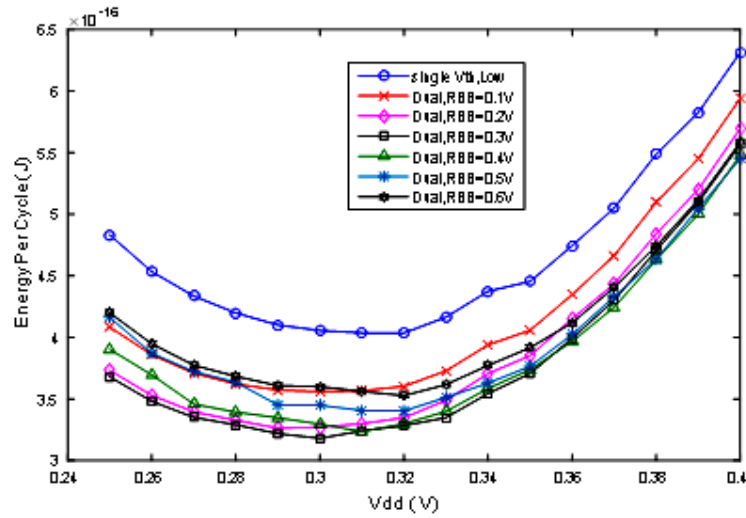


Figure 9. EPC of 74283 circuit for single and dual Vth design

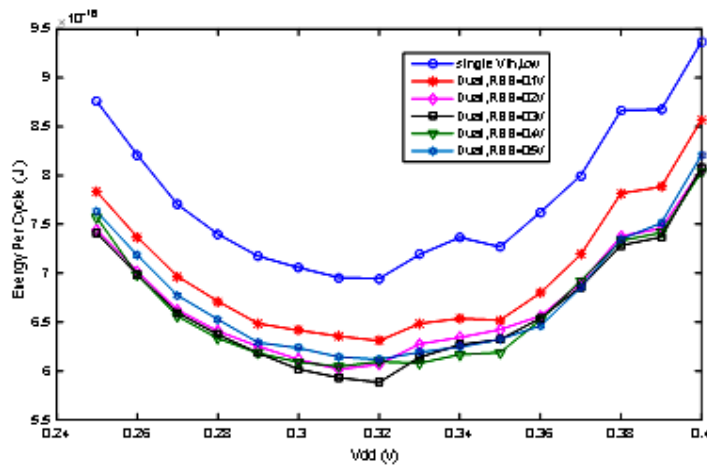


Figure 10. EPC of ALU74181 circuit for single and dual Vth design

Table 3. Experimental results of simulated circuits.

Circuit	Design	EPC (f J)	Freq. (MHz)	Vdd _{opt}	RBB	EPC Saving	Vth-H gates %
16 bit RCA	Single-Vth	4.290	12.041	0.32V	0	-	-
	Dual_Vth	2.476	10.91	0.3V	0.7V	42.28%	60.2 %
74283	Single-Vth	0.403	17.685	0.32 V	0	-	-
	Dual_Vth	0.318	17.01	0.3 V	0.3V	21 %	61.5%
ALU 74181	Single-Vth	0.694	17.216	0.32V	0	-	-
	Dual_Vth	0.588	17.216	0.32V	0.3V	15.27%	57.6%
74L85	Single-Vth	0.459	17.142	0.31V	0	-	-
	Dual_Vth	0.392	17.142	0.31V	0.2V	14.58%	54.3%

The other part of the results is the standby mode simulation results. Figure 11 shows the standby leakage power of the 16 bit RCA for different values of RBB and Vdd. As seen from this figure, the leakage power is reduced with the increasing of the RBB but and the maximum power saving (67%) is occur at (RBB= 0.8V). For more RBB increasing, the case is reversed. The circuit supply voltage should be equal (Vdd_{opt}) obtained from active mode optimization (0.3V). Similar simulations are performed for the other circuits as given in Table 4.

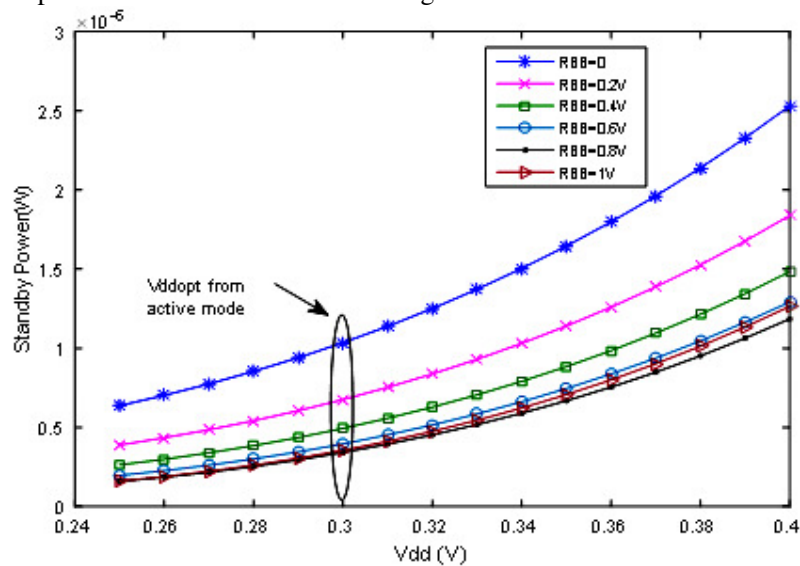


Figure 11. Standby leakage power for 16 bit RCA

Table 4. Experimental results of simulated circuits in standby mode.

Circuit	non_optm P _{stby}	Optm. P _{stby}	Vdd _{opt}	Optimal RBB	Saving
16 bit-RCA	1.035 μ w	0.341 μ w	0.3V	0.8V	67%
74L85	0.456 μ w	0.165 μ w	0.31V	0.8V	63.8%
ALU74181	0.706 μ w	0.262 μ w	0.32V	0.8V	62.88%
74283	0.358 μ w	0.12 μ w	0.3V	0.8V	66.48%

6. CONCLUSIONS

Some applications in the market may require minimum energy consumption in deep submicron technologies for both active and standby modes. Ignoring standby leakage current in sub-threshold can significantly impact the energy efficiency of the design. In this paper, a combined standby and active modes energy optimization methodology is proposed for non-zero standby 22nm sub-threshold circuits. In the first part of the paper, the validation of dual threshold voltage design for runtime energy minimization was investigated. A slack based genetic algorithm is used to find the optimal RBB or (optimal V_{th}) and the optimal supply voltage to ensure low EPC operation with the maximum allowable frequency at (V_{ddopt}). According to the obtained results, it is clear that the dual threshold voltage design still has an important role in energy optimization for 22nm sub-threshold circuits. The 16 bit RCA shows a maximum slack time utilization, where (42.28%) EPC reduction is obtained as compared with the non-optimized circuit. Whereas the minimum EPC reduction (14.58%) is obtained for the 74L85 circuit. The other two benchmark circuits; 74283 and ALU74181 result in (21%) and (15.27%) EPC reduction, respectively. This variation in EPC reduction is due to the topology of the circuit, and the available slack time from the off-critical paths to be utilized.

In the second part of the paper, a variable threshold standby power minimization approach is integrated with the active mode energy optimization given in part1. A significant standby leakage power saving on average of (65%) is obtained. This shows the impact of leakage reduction in standby mode and how can this affect the energy efficiency of the design, especially for non-zero standby circuits.

As future works, other techniques can be integrated with this approach such as sizing technique transistor stacking, gate replacement techniques, and Multi- V_{th} technique. For efficient slack time exploiting and greater energy reduction, different gate designs can be used. Moreover, aggressive downscaling may display higher leakage current in the sub-threshold region, so it is worth to explore the validation of dual threshold design in the future CMOS technologies.

REFERENCES

- [1] Rani NG, Kumar NP, Charles BS, Reddy PC, Ali SM. Design of Near-Threshold CMOS Logic Gates. International Journal of VLSI Design & Communication Systems. 2012 Apr 1;3(2):193.
- [2] A. Wang, B. H. Calhoun; Chandrakasan, A. P. Sub-threshold Design for Ultra Low-Power Systems, 1st ed.; Springer: New York, NY 10013, USA, 2006.
- [3] Ahmed R. ADAPTIVE SUPPLY VOLTAGE MANAGEMENT FOR LOW POWER LOGIC CIRCUITRY OPERATING AT SUBTHRESHOLD. International Journal of VLSI Design & Communication Systems. 2015 Apr 1;6(2):1.
- [4] Yao, J. Dual-Threshold Voltage Design of Sub-Threshold Circuits, Doctoral dissertation-Auburn University, USA, 2014.
- [5] Kim, K. Ultra Low Power CMOS Design, Doctoral dissertation-Auburn University, USA, 2011.
- [6] Kim, K.; Agrawal, V. D. Dual voltage design for minimum energy using gate slack. In Proceedings of the IEEE Intl. Conf. on Industrial Technology, Auburn University, Auburn, AL, USA, 14 Mar - 16 Mar 2011; pp. 419-424.
- [7] Kim, K.; Agrawal, V. D. Minimum energy CMOS design with dual subthreshold supply and multiple logic-level gates. In Proceedings of the IEEE International Symposium on Quality Electronic Design, Santa Clara, CA USA, March 14-16, 2011; pp. 1-6.

- [8] Kim, K.; Agrawal, V. D. True minimum energy design using dual below-threshold supply voltages. In Proceedings of the IEEE Intl. Conf. on VLSI Design, IIT Madras, Chennai, India ,2-7 January 2011; pp. 292-297.
- [9] Feeney, L. M.; Nilson, M. Ultra low energy CMOS logic using below-threshold dual-voltage supply. *J. Low Power Electron.* 2011, 7, pp.460-470.
- [10] Benton H. C.; Anantha, P. C. Ultra-dynamic voltage scaling (UDVS) using sub-threshold operation and local voltage dithering. *IEEE Journal of Solid-State Circuits* 2006, 41, pp. 238-245.
- [11] Gao, F.; Hayes, J. P. Total power reduction in CMOS circuits via gate sizing and multiple threshold voltages. In Proceedings of Design Automation Conference, Anaheim, CA, USA, 13-17 June 2005; pp. 31-36.
- [12] Lu, Y. Power and Performance Optimization of Static CMOS Circuits with Process Variation, Doctoral dissertation-Auburn University, 2007.
- [13] Lu, Y. ; Agrawal, V. D. Leakage and dynamic glitch power minimization using integer linear programming for Vth assignment and path balancing. In Proceedings of the Intl. Workshop on Power and Timing Modeling, Optimization and Simulation, Leuven, Belgium, 20-23 September 2005; pp. 217-226.
- [14] Nguye, D.; Davare, A.; Orshansky, M.; Chinney, D.; Thompson, B.; Keutzer, K. Minimization of dynamic and static power through joint assignment of threshold voltages and sizing optimization. In Proceedings of International Symposium on Low Power Electronics and Design, 25-27 August 2003, Seoul, Korea; pp. 158-163.
- [15] Seok, M.; Hanson, S.; Sylvester, D.; Blaauw, D. Analysis and optimization of sleep modes in subthreshold circuit design. In Proceedings of the 44th annual Design Automation Conference, 4-8 June 2007, San Diego, California, USA; pp. 694-699,ACM
- [16] Deepika G, Krishna PR, Rao KS. A SUB THRESHOLD SOURCE COUPLED LOGIC BASED DESIGN OF LOW POWER CMOS ANALOG MULTIPLEXER. *International Journal of VLSI Design & Communication Systems.* 2014 Oct 1;5(5):45.
- [17] Volkan, K.; Eby, G. F. Multi-voltage CMOS circuit design, 1st ed.; John Wiley & Sons: Southern Gate, Chichester, England, 2006.
- [18] Andrew, B.K.; Jens, L.; Igor, L. M.; Jin, H. VLSI physical design: from graph partitioning to timing closure, 1st ed.; Springer Science & Business Media: Dordrecht Heidelberg, London, 2011.
- [19] Ricardo,S. APAC: An exact algorithm for retrieving cycles and paths in all kinds of graphs. *Tékhné-Revista de Estudos Politécnicos* 2009, Vol. VII, N° 12, pp. 39-55.

AUTHORS

Ali T. Shaheen received his B.S. degree in electrical engineering from University of Baghdad in July 2000, and his M.S. degree in electronic and communication from University of Baghdad, Iraq in 2007. He is working as a lecturer in the Electrical Engineering Department, University of Baghdad and currently he is a Ph.D. research student in the same university. His research interests include the fields of digital signal processing, communication and currently he is working on the field of low power electronic circuits.

Saleem M. R. Taha received the B.S. degree in electrical engineering with elective courses in computer science, the Higher Diploma degree in electronics and communications, the M.S. degree in digital systems design, and the Ph.D. degree in reversible and quantum computations from the University of Baghdad, Baghdad, Iraq, in 1978, 1980, 1982 and 2011, respectively. In May 1983 he joined the Department of Electrical Engineering, College of Engineering, University of Baghdad, and since July 1998 he has been a Professor at that university. He has supervised more than fifty M.S. and Ph.D. research students. In

addition, he has contributed to and authored more than sixty five papers in the areas of hybrid circuit design, digital signal processing, digital instrumentation and measurements, microcomputer applications, biomedical engineering, computer-aided design, reversible and quantum computations.