

# A Majority-Based Imprecise Multiplier for Ultra-Efficient Approximate Image Multiplication

Farnaz Sabetzadeh, Mohammad Hossein Moaiyeri<sup>ID</sup>, *Senior Member, IEEE*, and Mohammad Ahmadinejad

**Abstract**—Approximate computing is an emerging approach for reducing the energy consumption and design complexity in many applications where accuracy is not a crucial necessity. In this study, ultra-efficient imprecise 4:2 compressor and multiplier circuits as the building blocks of the approximate computing systems are proposed. The proposed compressor uses only one majority gate which is different from the conventional design methods using *AND*-*OR* and *XOR* logics. Furthermore, the majority gate is the fundamental logic block in many of the emerging majority-friendly nanotechnologies such as quantum-dot cellular automata (*QCA*) and single-electron transistor (*SET*). The proposed circuits are designed using FinFET as a current industrial technology and are simulated with HSPICE at 7nm technology node. The results indicate that our imprecise compressor is superior to its previous counterparts in terms of delay, power consumption, power delay product (*PDP*) and area, and improves these parameters on average by 32%, 68%, 78%, and 66%, respectively. In addition, the proposed efficient approximate multiplier is utilized in image multiplying as an important image processing application. The HSPICE and MATLAB simulations indicate that the proposed inexact multiplier provides a significant compromise between accuracy and design efficiency for approximate computing.

**Index Terms**—Approximate computing, compressor, multiplier, majority logic, FinFET, emerging technologies.

## I. INTRODUCTION

THE increasing of the density and complexity of nanoscale digital integrated circuits have led to considerably higher power density and heat dissipation in modern VLSI chips. The high power density increases the leakage currents and reduces the reliability and lifetime of integrated circuits. Moreover, the energy consumption has become very critical especially in battery operated portable electronic devices [1]. An emerging paradigm for reducing the energy dissipation is to use approximate computing in applications where a high accuracy is not a crucial necessity.

Computation's errors and imprecisions can be tolerated in these applications, while having understandable and beneficial

outcomes that are perceptible enough for human realization. Actually, with a reasonable reduction in preciseness, many circuit parameters such as the number of devices, energy consumption, delay and area can be reduced. Accordingly, approximate computing is a successful solution for fast computation in error tolerant applications to reach simpler circuits with more energy efficiency [2], [3]. It is notable that, as a design paradigm, approximate computing can be applied in various design levels of abstraction such as transistor, logic, algorithmic, architecture and software [4].

Multiplier is one of the most widely used and power-hungry arithmetic blocks in many digital systems [5]. Because of the complex structure of hardware multipliers and as they are generally placed on the critical path of digital systems, using approximate computing can deliver considerable improvements regarding the system's performance and energy dissipation [2], [6].

Multiplication operation is typically carried out in three phases: 1) partial products generation using *AND* gates; 2) partial products reduction; 3) final products generation by means of an adder structure [7]. Considering these three steps, partial products reduction is the most important step in terms of energy consumption and area [7]. Efficient circuitry for this stage can be realized with efficient approximate 4:2 compressors. Accordingly, many efficient exact and imprecise compressors with different levels of efficiencies in hardware and accuracy related parameters have already been presented in the literature [2], [3], [8]–[14]. These approximate compressors sacrifice accuracy for energy to different extents and are useful for different imprecise applications. Moreover, extensive studies have been conducted in the area of approximate multipliers targeting approximate partial product generation, delivering promising results [6].

Scaling of the planar MOSFET has led to some critical problems such as reduced gate control, drain-induced barrier lowering (*DIBL*), threshold voltage variation, and considerably high power densities [15]. The FinFET device with a three-dimensional tri-gate structure has been evolved as a successful replace for the planar MOSFET. FinFET significantly enhances gate control, reduces short channel effects, and increases  $I_{on}/I_{off}$  ratio. Moreover, the intrinsic body of FinFET eliminates the random dopant fluctuations as an important factor for threshold voltage variations [16]. However, FinFET suffers from self-heating effect and higher power

Manuscript received January 29, 2019; revised April 21, 2019 and May 15, 2019; accepted May 19, 2019. Date of publication June 4, 2019; date of current version October 30, 2019. This paper was recommended by Associate Editor W. Zhao. (Corresponding author: Mohammad Hossein Moaiyeri.)

The authors are with the Department of Electrical Engineering, Shahid Beheshti University, Tehran 1983963113, Iran (e-mail: f.sabetzadeh@mail.sbu.ac.ir; h\_moaiyeri@sbu.ac.ir; mo.ahmadinejad@mail.sbu.ac.ir).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSI.2019.2918241

1549-8328 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See [http://www.ieee.org/publications\\_standards/publications/rights/index.html](http://www.ieee.org/publications_standards/publications/rights/index.html) for more information.

density as the confined narrow fins decrease the thermal conductivity of the channel. Therefore, low-power design methods such as imprecise computing can also help to address this issue [3].

Accordingly, proposing energy and area efficient approximate compressors and multipliers based on FinFET is of high significance. Moreover, as FinFET is the leading technology of the current electronics industry, the results will be more beneficial and of interest for the scientists and engineers working in this area.

In this article, we propose ultra-efficient approximate 4:2 compressor and multiplier designs based on majority logic. The proposed compressor is designed to make a trade-off between error distance and transistor count. While this design has only 12 transistors, it generates no false output with an error distance value of greater than 1. Furthermore, unlike the previous approximate compressors, which were designed based on AND/OR and XOR logics, our compressor is based on the majority logic, which leads to a more efficient imprecise multiplier. As the Sum output of this majority-based compressor is equal to 1, the reduction and final addition stages of the approximate multiplier realizing by this compressor become much more simpler. This leads to a considerably lower number of transistors and a shorter critical path, which result in lower power and propagation delay. It is also noteworthy that as the majority logic is the building block in many emerging nanotechnologies, our approach is also very convenient for approximate computing based on nanotechnologies.

The circuit parameters are evaluated using HSPICE and 7nm FinFET technology. Furthermore, important accuracy and quality metrics are investigated for the approximate multipliers using MATLAB. It is notable that the main essence of the approximate computing is to decrease the design complexity and energy dissipation considerably in comparison with the exact structures, while sustaining a certain quality. Accordingly, two figure of merits (*FOMs*), considering both quality and efficiency factors, are also calculated for the approximate multipliers. These metrics indicate that our design provides an effective trade-off between quality and efficiency for approximate computing.

The rest of this study is organized as follows: Section II briefly reviews the backgrounds of this research. The proposed designs are presented in Section III. The comprehensive simulation results and comparisons are given in Section IV, and finally, Section V concludes the study.

## II. BACKGROUNDS

### A. FinFET

FinFET is a quasi-planar multi-gate transistor with an ultra-thin body. In a FinFET, the gate dominantly controls the thin channel from multiple sides, which leads to a smaller subthreshold swing, a lower DIBL and a higher  $I_{on}/I_{off}$  ratio. Moreover, the fully depleted undoped body of FinFET resolves random dopant fluctuations and leads to a higher carrier mobility. Accordingly, FinFET has emerged as a feasible replace for the planar MOSFET at deep nanoscale technologies [17], [18]. Figure 1 shows the structure of a tri-gate FinFET. The driving

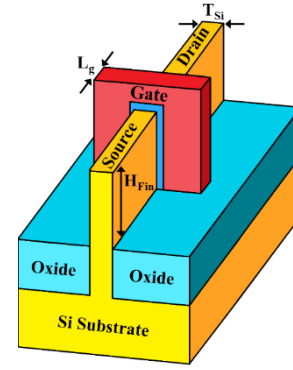


Fig. 1. Structure of a FinFET device.

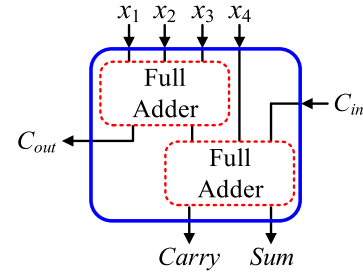


Fig. 2. Schematic of the 4:2 compressor.

current of a FinFET can be expressed as [17]

$$I_D = \beta N_{Fin} \frac{2H_{Fin} + T_{Si}}{L} (V_{GS} - V_{th})^\alpha \quad (1)$$

where,  $N_{Fin}$  is the number of fins and  $T_{Si}$  and  $H_{Fin}$  denote the thickness and height of the fins, which are usually fixed in a specific technology for fabrication related reasons. Moreover,  $\alpha$  and  $\beta$  are fitting constants.

While FinFET has excellent electrical characteristics, it suffers from self-heating effect and higher power density as the thin fins decrease the thermal conductivity of the channel [19]. Self-heating can increase leakage currents and degrade the performance and lifetime of digital circuits. Approximate computing as an effective approach for lowering the energy dissipation is also useful for handling these device-related negative issues at higher design levels of abstraction [3].

### B. Compressors

Compressors are used for implementing the partial products reduction stage in high-performance and energy-efficient multipliers [7]. The general schematic of an exact 4:2 compressor is shown in Fig. 2. The exact 4:2 compressor has four main inputs ( $x_1, x_2, x_3, x_4$ ) and two main outputs (*Sum* and *Carry*). Furthermore, an input carry ( $C_{in}$ ) comes from the preceding block of lower significance and an output carry ( $C_{out}$ ) goes to the next block of higher significance. The output equations of the conventional 4:2 compressor can be expressed as [7]

$$Sum = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus C_{in} \quad (2)$$

$$Carry = (x_1 \oplus x_2 \oplus x_3 \oplus x_4) \cdot C_{in} + \overline{(x_1 \oplus x_2 \oplus x_3 \oplus x_4)} \cdot x_4 \quad (3)$$

$$C_{out} = (x_1 \oplus x_2) \cdot x_3 + \overline{(x_1 \oplus x_2)} \cdot x_1 \quad (4)$$

By increasing the importance of the energy consumption and heat dissipation in recent integrated circuits, approximate circuits have become more interesting. Accordingly, many efficient 4:2 approximate compressors have already been presented in the literature [2], [3], [9]–[14]. Most of these imprecise structures have ignored  $C_{in}$  and  $C_{out}$  as an effective approach for enhancing design efficiency.

In [2], two state-of-the-art designs for approximate 4:2 compressors have been proposed and used in an 8-bit Dadda multiplier. The second design, which ignores  $C_{in}$  and  $C_{out}$  signals, is more efficient and makes an excellent trade-off between preciseness and performance factors. This structure is based on the XNOR and NOR logics and can be designed efficiently at the transistor level with 26 transistors.

In [9], a dual-quality structure for 4:2 compressor with the capability of switching between the exact and approximate operating modes during runtime has been proposed. The overall structure of this compressor includes two approximate and supplementary parts. The most effective approximate component of [9] (DQ4:2C4), which operates based on the XOR and AND-OR logics can be implemented efficiently at the transistor level with 26 transistors.

Since the XOR logic usually increases the area and power consumption, a different approach has been presented in [10], which replaces XOR gates with the OR gates in some cases. This structure can be implemented at the transistor level with 36 transistors.

The approximate 4:2 compressor presented in [12] has been designed by modifying the truth table of the design presented in [11]. This approach, which considerably improves the design performance and efficiency compared to [11], is based on the XOR and AND-OR logics and can be implemented efficiently with 30 transistors.

In [13], an approximate 4:2 compressor has been suggested based on some revisions on the truth table of the exact compressor. The Boolean logic functions of the most efficient 4:2 compressor presented in this work can be implemented with 30 transistors.

Another approximate 4:2 compressor based on the AND-OR logics has been presented in [14], which can be implemented efficiently with 20 transistors. However, existence of a four-input NOR gate in this design degrade its performance and efficiency to some extent.

In some of these approaches, higher accuracies have been achieved with the cost of more transistors and higher delay and power dissipation. However, as can be observed in some effective approaches like [2], significant reduction of the area and energy consumption, while providing an acceptable level of preciseness, is the main goal in approximate computing.

In the next section, the proposed approximate compressor will be discussed. It is noteworthy that the proposed compressor significantly reduces the number of transistors, while trying to restrict the error distance. Moreover, unlike the discussed previous circuits, which were mainly designed based on the AND/OR and XOR logics, it is designed based on majority logic, which leads to a more efficient imprecise multiplier and makes it compatible with the emerging majority-based nanotechnologies.

TABLE I  
THE TRUTH TABLE OF PROPOSED APPROXIMATE 4:2 COMPRESSOR

$x_4$	$x_3$	$x_2$	$x_1$	Carry	Sum	Error distance
0	0	0	0	0	1	+1
0	0	0	1	0	1	0
0	0	1	0	0	1	0
0	0	1	1	0	1	-1
0	1	0	0	0	1	0
0	1	0	1	1	1	+1
0	1	1	0	0	1	-1
0	1	1	1	1	1	0
1	0	0	0	0	1	0
1	0	0	1	1	1	+1
1	0	1	0	0	1	-1
1	0	1	1	1	1	0
1	1	0	0	1	1	+1
1	1	0	1	1	1	0
1	1	1	0	1	1	0
1	1	1	1	1	1	-1

### III. PROPOSED APPROXIMATE DESIGNS

The recent imprecise compressors have usually been designed based on the AND-OR and XOR logics. Using the XOR logic increases the overall switching activity [20] and consequently the dynamic power. On the other hand, recent studies have shown that using majority logic can lead to a higher design efficiency in comparison with the other common implementations in emerging nanotechnologies [21]–[24]. In this section, our proposed majority-based imprecise 4:2 compressor and multiplier structures are presented.

#### A. Imprecise 4:2 Compressor

The proposed approximate 4:2 compressor operates according to the equations given in (5)–(7). In our design, similar to other designs like [2], the  $C_{in}$  and  $C_{out}$  signals are ignored for design efficiency reasons. In addition, the inputs  $x_1$ ,  $x_3$  and  $x_4$  are given to a majority gate that produces the *Carry* output. Furthermore, *Sum* is considered constant and equal to ‘1’, and no additional hardware is required for calculating the *Sum* value. This great simplification causes a significant reduction in the overall energy consumption and propagation delay of the design.

$$C_{out} = C_{in} = 0 \quad (5)$$

$$Carry = Majority(x_1, x_3, x_4) = x_4(x_1 + x_3) + x_1x_3 \quad (6)$$

$$Sum = V_{DD} \quad (7)$$

The truth table of the proposed design is shown in Table I. A basic concern in designing the approximate blocks is to minimize the error distance (*ED*) between the outcomes of the exact and inexact designs.

The error distance is defined as the arithmetic distance between a false output and its correct value [25]. As an example in Table I, when all inputs are ‘1’, the decimal value of the sum of the inputs is 4. However, the approximate compressor produces ‘1’ for both *Carry* and *Sum* outputs, which results in a decimal value of 3 and consequently an error distance of –1. It is notable that there is no false output with an *ED* value of 2 and greater in the proposed approach. These erroneous outputs can lead to unacceptably great errors when



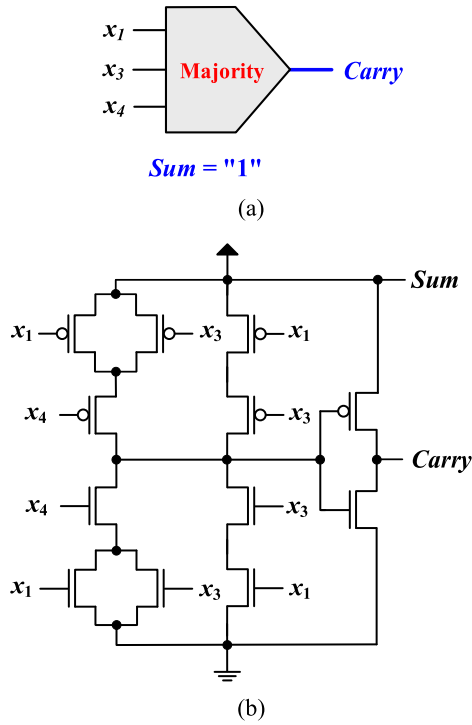


Fig. 3. The proposed approximate 4:2 compressor (a) Block diagram. (b) Circuit design.

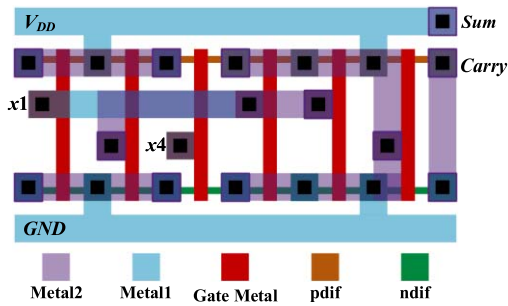


Fig. 4. Layout view of the proposed approximate 4:2 compressor.

the compressor is utilized in the structure of an approximate multiplier. Moreover, the error distances with opposite signs such as  $-1$  and  $+1$  can mutually diminish each other's impact in the structure of a multiplier.

The proposed approximate 4:2 compressor is shown in Fig. 3. The proposed imprecise design has a very simple structure and includes only one majority gate. The majority gate is designed with 12 transistors based on the complementary logic style. The proposed design superiorly reduces the number of transistors and critical path length and consequently the power consumption and delay as compared to its previous counterparts.

The layout of the proposed approximate compressor, designed based on the FinFET layout design rules reported in [26] and on the fin grid of 7nm, is shown in Fig. 4.

According to (1), the driving current of a FinFET can be enhanced by using multiple parallel fins as channel. However, this also significantly increases the total switching capacitance due to the three-dimensional structure of FinFET and enlarges the cell area due to the considerable fin pitch overhead.

Accordingly, as the energy consumption and area are two significant factors in approximate computing, and given that the electrons and holes can have near mobilities in the FinFETs [17], single-fin devices are used.

The majority gate is the basic logic block in many emerging technologies such as quantum-dot cellular automata (QCA), single electron tunneling (SET), magnetic tunnel junction (MTJ), nanomagnetic logic (NML), tunneling phase logic (TPL) and memristor [27]. Moreover, realization of the majority logic based on DNA has been demonstrated in [28]. The obtainability of efficient majority gate implementations in these emerging devices makes our proposed design also very convenient for approximate computing based on the emerging nanotechnologies.

### B. Approximate Multiplier

In high-performance multipliers the reduction stage, as the most critical and power-hungry stage, is implemented using compressors [7]. Using approximate compressors in this stage results in an approximate multiplier [2]. The general structure of an approximate 8-bit Dadda multiplier, based on the 4:2 compressors ignoring  $C_{in}$  and  $C_{out}$ , has been described in [2]. In this structure, the partial products are generated using an array of the AND gates and are reduced mainly based on the approximate compressors. Finally, in the last stage, a ripple carry adder (RCA) generates the final products.

The reduction circuitry of the modified version of the approximate 8-bit Dadda multiplier, realized by the proposed imprecise compressor is shown in Fig. 5. In this structure, to improve the performance and efficiency of the approximate multiplier with a negligible impact on its preciseness, the first four least significant columns of the partial products can be truncated [12]. Moreover, using exact computing for the high significant bits can lead to a considerably higher accuracy [12]–[14]. However, using exact compressors reduces the efficiency of the approximate multiplier. Furthermore, obtaining a preciseness more than a certain level but with a higher energy consumption is not desired in approximate applications. Therefore, to make an effective trade-off between the preciseness and performance parameters for approximate computing, we use exact compressors for the last five high significant bits at the output. Accordingly, as shown in Fig. 5, only four of all 16 compressors are exact, which significantly enhances the accuracy with a low hardware cost.

In the first step, three half-adders, two full-adder and six proposed imprecise compressors are utilized to reduce the partial products into at most four rows. It is notable that the half adder and full adder cells with the conventional complementary logic style [31] are used in this stage.

In the second stage, a full adder, three exact compressors [8] and six proposed imprecise compressors are used. However, five of these compressors (specified with dashed lines) have one '1' input logic, as the Sum outputs of the compressors in the previous step are equal to '1'. As a three-input majority gate with one '1' input operates as a two-input OR gate, these 12-transistor compressors are replaced with conventional OR gates with only six transistors. As the proposed compressor

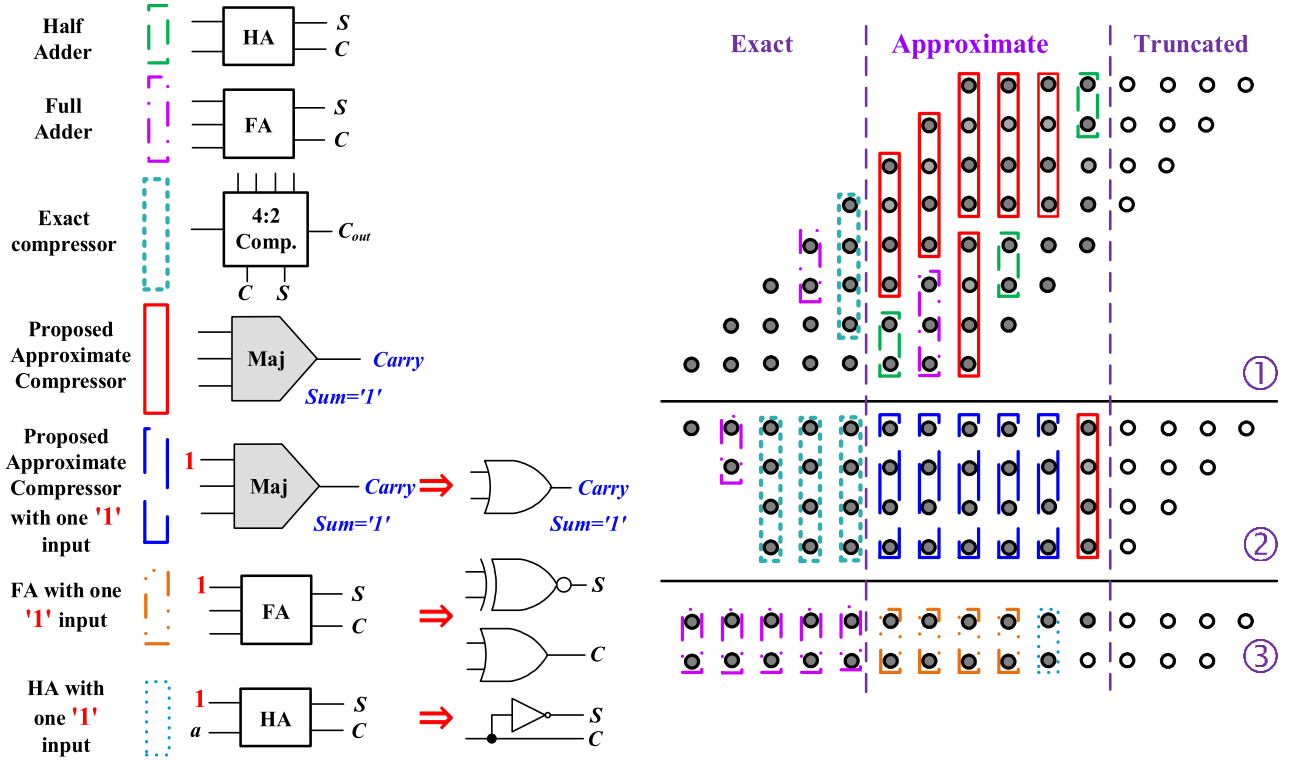


Fig. 5. Partial product reduction circuitry of the proposed approximate multiplier.

ignores the  $x_2$  input, one of the partial products is not required in each of the approximate compressors in the first stage and the first one in the second stage. As a result, seven AND gates are reduced from the partial product generator stage of the proposed multiplier.

In the RCA step, the first module is a half adder with one '1' input, which can be replaced with just an inverter gate as illustrated in Fig. 5. Moreover, as the three-input XOR and majority gates with one '1' input operate as two-input XNOR and OR gates, respectively, each of the next four full adders with one '1' input (*Sum* of the previous step compressor) are replaced with a two-input XNOR (*Sum*) and a two-input OR ( $C_{out}$ ). It is notable that the six-transistor CMOS+ design presented in [31] are utilized for two-input XNORs.

Accordingly, by utilizing the proposed imprecise compressor, the delay, number of the transistors (as a criterion for the circuit complexity) and consequently the total energy consumption of the whole multiplier are significantly reduced.

#### IV. SIMULATION RESULTS AND COMPARISONS

The compressors and multipliers are simulated using HSPICE with 7nm FinFET technology [32], [33] at 0.7V supply voltage and 2 GHz frequency. Some important parameters of the FinFET model are listed in Tables II.

##### A. Compressors

To have a fair comparison, the optimized Boolean functions of the previous imprecise compressors are implemented at the transistor level using FinFETs. Moreover, the imprecise compressors, which have ignored  $C_{in}$  and  $C_{out}$  as an effective

Parameters	n-type	p-type
Physical fin thickness	6.5 nm	6.5 nm
Fin height	32 nm	32 nm
Gate length	21 nm	21 nm
Equivalent oxide thickness	1.25 nm	1.25 nm
Body doping	$10^{16} \text{ cm}^{-3}$	$10^{16} \text{ cm}^{-3}$
Source/Drain doping	$2 \times 10^{20} \text{ cm}^{-3}$	$2 \times 10^{20} \text{ cm}^{-3}$
Low field mobility ( $\mu_0$ )	$252 \text{ cm}^2/\text{V.s}$	$210 \text{ cm}^2/\text{V.s}$
Gate Work Function ( $\Phi_M$ )	4.372 eV	4.8108 eV

approach for enhancing design efficiency, are considered for comparison.

The input signals are entered to the compressors via input buffers (two cascaded inverters) and fan-out-of-four (FO4) inverter loads are inserted at the outputs. The longest interval between the instant that the earliest input reaches  $V_{DD}/2$  and the moment that the last output reaches  $V_{DD}/2$  is calculated as the propagation delay. To evaluate the power consumption of the compressors, exhaustive input patterns, considering all of the possible input transitions have been generated using a C++ code and have been fed into each design as the input stimuli through HSPICE simulation. Furthermore, long streams of random input combinations are generated to estimate the power consumption of the multipliers using HSPICE simulation. The power-delay product (PDP) is also calculated to make a clearer assessment of the energy efficiency of the designs. Moreover, to compare the areas, the layouts of the FinFET-based compressors have been designed efficiently on the fin grid of 7nm and based on the FinFET layout design rules reported in [26].

TABLE III  
SIMULATION RESULTS OF 4:2 COMPRESSORS

Compressor	Delay (ps)	Power (nW)	PDP (aJ)	Transistors	Area ( $\mu\text{m}^2$ )
Proposed	17.21	70	1.20	12	0.105
[2]	19.55	188	3.68	26	0.307
[9]	20.01	190	3.80	26	0.307
[10]	25.99	270	7.02	36	0.443
[12]	29.91	214	6.41	30	0.329
[13]	24.65	262	6.46	30	0.355
[14]	28.61	137	3.92	20	0.176
Exact [8]	34.38	464	15.95	38	0.484

TABLE IV  
COMPARISON OF THE MULTIPLIERS BASED ON DIFFERENT 4:2 COMPRESSORS

Multipliers	Critical path Delay (ps)	Power (uW)	PDP (fJ)	Transistors
Proposed	143.4	7.27	1.04	852
[2]	222.4	10.94	2.43	1180
[9]	222.6	11.14	2.48	1180
[10]	236.5	11.73	2.77	1300
[12]	244.5	12.73	3.11	1228
[13]	233.1	12.85	3.00	1228
[14]	239.8	10.06	2.41	1108
Exact [8]	252.8	18.79	4.75	1530

The most efficient exact 4:2 compressor presented in [8] and the imprecise 4:2 compressors are compared in Table III. According to the results, the proposed design has the lowest number of transistors, delay, power consumption and *PDP* as compared to the other state-of-the-art 4:2 compressors.

This is due to the simple structure of the proposed compressor consisting of just one majority gate (generating the *Carry* output) and the constant value of ‘1’ assigned to the *Sum* output. In addition to the significantly smaller area, this also reduces the propagation delay and switching activity and consequently the power dissipation of our design. The proposed circuit improves the delay, power, *PDP* and area on average by 32%, 68%, 78% and 66%, respectively as compared to the previous designs listed in Table III.

### B. Approximate Multipliers

In this section, the hardware efficiency and accuracy parameters of the approximate multipliers are evaluated and analyzed.

1) *Hardware Analysis*: The exact compressor-based 8-bit Dadda multiplier [8], the approximate 8-bit Dadda multiplier (with truncated four LSB columns and exact four MSB columns) based on different imprecise compressors and the proposed design are simulated using HSPICE with the 7nm FinFET technology. The worst-case critical path delay, average power consumption, *PDP* and the total number of transistors for different multipliers under study are given in Table IV. The proposed multiplier has the lowest propagation delay, power dissipation and *PDP* compared to the other designs due to the quite simpler structure of the proposed compressor as well as the very simplified structure of the multiplier as described in Section III. Our proposed multiplier improves the delay,

TABLE V  
ACCURACY METRICS OF THE APPROXIMATE MULTIPLIERS

Multiplier	ER (%)	MRED	NMED
Proposed	99.82	0.447	0.007
[2]	99.81	0.497	0.007
[9]	89.88	0.041	0.006
[10]	89.88	0.041	0.006
[12]	87.08	0.025	0.003
[13]	88.31	0.031	0.003
[14]	91.93	0.043	0.005

power, *PDP* and the number of transistors on average by 39%, 40%, 61% and 31%, respectively, as compared to other multipliers listed in Table IV.

2) *Accuracy Analysis*: The accuracy factors of the considered multipliers are investigated in this section based on the results obtained by the MATLAB simulations.

In addition to the error distance parameter as a metric for determining the output quality, some other more illustrative metrics, such as error rate (*ER*), mean relative error distance (*MRED*) and normalized mean error distance (*NMED*) can be considered to assess the output quality of the approximate circuits in error resilient applications [34], [35].

Error rate is calculated as the probability of producing an incorrect result. Furthermore, *MRED*, which is the average value of all possible relative error distances, is expressed by

$$MRED = \frac{1}{2^{2N}} \sum_{i=1}^{2^{2N}} \frac{|ED_i|}{M_i} \quad (8)$$

where,  $N$  is the bit length of multiplier, and  $ED_i$  and  $M_i$  are the error distance (distance between the exact and approximate outputs) and exact result for each combination of input operands, respectively.

The *NMED* metric, which is the mean error distance normalized by the maximum value of error that an inaccurate multiplier can have, is calculated as

$$NMED = \frac{1}{2^{2N} (2^N - 1)^2} \sum_{i=1}^{2^{2N}} |ED_i| \quad (9)$$

The accuracy metrics of the multipliers are presented in Table V. The results were achieved by applying all 65536 numbers for the 8-bit multipliers. The results indicate high error rate for all of the imprecise multipliers (relatively higher for our design). This is mainly due to the truncation of the four low significant columns. However, in imprecise applications the error distance related parameters are of more significance. The proposed multiplier has a slightly less *MRED* than that obtained by the multiplier realized by the compressor presented in [2], while, it has a greater *MRED* compared to the other multipliers.

However, it is noteworthy that in many imprecise applications such as those, which deal with the human senses, the difference between the exact and inaccurate results (for instance, the difference between the exact and inexact colors in a pixel) is more important than their relative difference. Moreover, in these applications obtaining an accuracy more

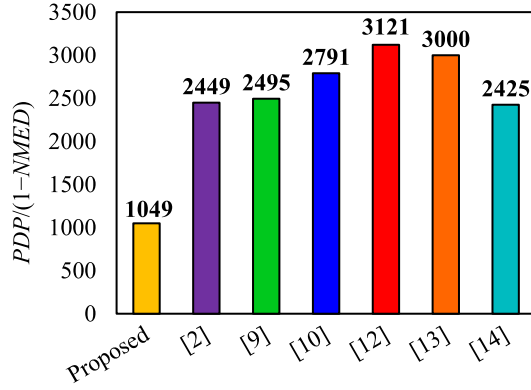


Fig. 6. The *FOM1* factor for the considered approximate multipliers.

than a certain level but with a higher energy consumption is not desired. Accordingly, in these cases the *NMED* metric would be more illustrative. The proposed design has an acceptable *NMED* as an approximate multiplier, which is not much higher than most of the previous designs. It is notable that the multipliers of [12] and [13], obtains lower *NMEDs* with high energy and transistor count costs (see Table IV).

To make a trade-off between preciseness and energy-efficiency, an illustrative figure of merit was suggested in [9] as  $(PDP \times \text{delay} \times \text{area} / (1 - NMED))$ . While our multiplier has a considerably lower delay and transistor count compared to the other designs, to have a fairer comparison, we have modified this factor as

$$FOM1 = \frac{PDP}{1 - NMED} \quad (10)$$

The *FOM1* factor for each of the approximate multipliers is shown in Fig. 6. In this figure, a design with a smaller *FOM1* value reaches a better hardware-accuracy trade-off.

The results indicate that the proposed design has on average 61% smaller *FOM1* factor compared to the other designs. Accordingly, the proposed multiplier obtains a better trade-off between accuracy and energy-efficiency, as the main objective of the approximate computing, especially in applications like imprecise image multiplication.

### C. Image Multiplication

In this section, to evaluate the effectiveness of the considered imprecise multipliers in real applications, they are utilized in image multiplication, as one of the most essential operations in image processing. A program has been developed using MATLAB to multiply two images pixel by pixel using the discussed approximate multipliers. The peak signal to noise ratio (*PSNR*) as a common metric for image quality assessment is calculated to evaluate the preciseness of the output images. The *PSNR* metric is defined as [36]

$$PSNR = 10 \log_{10} \left( \frac{m \times p \times MAX_I^2}{\sum_{i=0}^{m-1} \sum_{j=0}^{p-1} [I(i, j) - K(i, j)]^2} \right) \quad (11)$$

TABLE VI  
THE *PSNRs* FOR THE IMAGE MULTIPLICATIONS USING APPROXIMATE MULTIPLIERS REALIZED BY DIFFERENT COMPRESSORS

Multiplier	sky × boat	peppers_col. × peppers_gray	sky × cameraman	sky × moon	Lena_col. × Lena_gray
Proposed	41.10	40.96	41.64	41.40	41.98
[2]	41.26	40.99	41.62	41.21	42.01
[9]	41.03	41.84	41.89	41.68	41.55
[10]	41.14	41.97	41.99	41.85	41.83
[12]	45.29	45.52	47.65	45.87	45.30
[13]	45.25	46.05	46.30	46.10	45.30
[14]	41.41	42.21	42.27	42.43	42.27

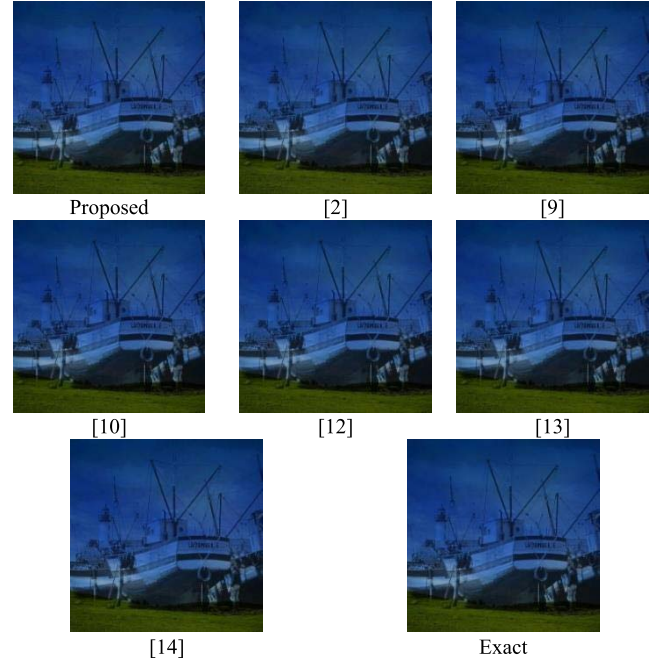


Fig. 7. Multiplied images of the sky and boat images using different 8-bit multiplier based on different inexact 4:2 compressors.

where,  $MAX_I$  is the maximum value of each pixel,  $m$  and  $p$  are the image dimensions and  $I(i, j)$  and  $K(i, j)$  are in turn the exact and obtained values for each pixel.

Our comprehensive simulation results have authenticated that the proposed design provides *PSNRs* greater than 40dB in various image multiplications, while a *PSNR* of 30dB can be considered as good enough [14]. The *PSNR* values for five image multiplication examples are presented in Table VI. According to the results, the proposed multiplier provides good *PSNRs*, while it leads considerably better performance and energy parameters.

In order to provide an observable impression regarding the impact of the approximate multiplications on the image qualities, as an example, the multiplied images of the sky and boat images using different approximate multipliers are illustrated in Fig. 7. It can be concluded from the results that the difference between the outputs of the exact and the approximate multipliers is not significant.

To have an illustrative evaluation about the appropriateness of different approximate designs for image multiplication, both quality and performance metrics should be taken into consideration. A figure of merit metric has been defined in [37]



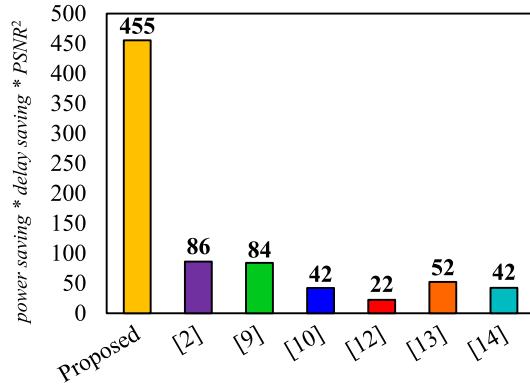


Fig. 8. The *FOM2* factor for the considered approximate multipliers.

as “ $PSNR^2 \times area\ saving \times power\ saving$ ”. In this metric, the delay parameter, which is one of the most important performance parameters in arithmetic circuits, is ignored. Moreover, it considers both power and area savings, simultaneously, which does not seem to be fair as the effects of area saving is somehow considered in the power saving. Accordingly, to have a fairer comparison, we use a modified *FOM* as

$$FOM2 = Power\ saving \times Delay\ saving \times PSNR^2 \quad (12)$$

Figure 8 shows the comparison of the *FOM2* factors of the approximate multipliers. According to the results, the proposed multiplier has a significantly greater *FOM2* as compared to the other designs (10.2x). The results indicate that the proposed design achieves the best trade-off between the quality and efficiency as compared to the other designs in approximate multiplication. It is noteworthy that the in addition to the considerably greater *FOM2*, the proposed approximate multiplier also significantly reduces the number of transistors (see Table IV) that can result in a higher yield and consequently lower fabrication costs.

## V. CONCLUSION

As the inexact computing is a brilliant method for energy efficient computation at nanoscale, in this study an ultra-efficient 4:2 approximate compressor based on the majority-gate logic has been presented. The proposed design has a simple structure and shows significant improvements in terms of transistor count, power consumption and delay as compared to its counterparts. This structure is proper for the current and future VLSI technologies, especially for the majority-friendly emerging nanotechnologies like QCA and SET. Moreover, the approximate compressors have been utilized in the reduction circuitry of the Dadda multiplier and the application of these approximate multipliers in image processing, especially in image multiplication, has been investigated. Some crucial metrics for measuring the accuracy of the outputs, such as *MRED*, *NMED* and *PSNR* have been evaluated in MATLAB for the approximate multipliers. In addition, comprehensive figure of merits, considering both quality and design efficiency, have been considered. The proposed multiplier has significantly improved the number of transistors, critical path delay and energy dissipation as compared to the other designs. In addition, the *FOMs* of the proposed multiplier

are significantly better than those of the other approximate multipliers. Accordingly, the proposed imprecise multiplier provides a significant compromise between the performance characteristics and preciseness for approximate computing.

## REFERENCES

- [1] A. Wang, B. H. Calhoun, and A. P. Chandrakasan, *Sub-Threshold Design for Ultra Low-Power Systems*, vol. 95. New York, NY, USA: Springer, 2006.
- [2] A. Momeni, J. Han, P. Montuschi, and F. Lombardi, “Design and analysis of approximate compressors for multiplication,” *IEEE Trans. Comput.*, vol. 64, no. 4, pp. 984–994, Apr. 2015.
- [3] M. H. Moaiyeri, F. Sabetzadeh, and S. Angizi, “An efficient majority-based compressor for approximate computing in the nano era,” *Microsyst. Technol.*, vol. 24, no. 3, pp. 1589–1601, Mar. 2018.
- [4] S. Mittal, “A survey of techniques for approximate computing,” *ACM Comput. Surv.*, vol. 48, no. 4, p. 62, May 2016.
- [5] V. Leon, S. Xydis, D. Soudris, and K. Pekmestzi, “Energy-efficient VLSI implementation of multipliers with double LSB operands,” *IET Circuits Devices Syst.*, to be published. [Online]. Available: <https://digital-library.theiet.org/content/journals/10.1049/iet-cds.2018.5039>. doi: 10.1049/iet-cds.2018.5039.
- [6] V. Leon, G. Zervakis, S. Xydis, D. Soudris, and K. Pekmestzi, “Walking through the energy-error Pareto frontier of approximate multipliers,” *IEEE Micro*, vol. 38, no. 4, pp. 40–49, Jul./Aug. 2018.
- [7] C.-H. Chang, J. Gu, and M. Zhang, “Ultra low-voltage low-power CMOS 4-2 and 5-2 compressors for fast arithmetic circuits,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 10, pp. 1985–1997, Oct. 2004.
- [8] A. Arasteh, M. H. Moaiyeri, M. Taheri, K. Navi, and N. Bagherzadeh, “An energy and area efficient 4:2 Compressor based on FinFETs,” *Integration*, vol. 60, pp. 224–231, Jan. 2018.
- [9] O. Akbari, M. Kamal, A. Afzali-Kusha, and M. Pedram, “Dual-quality 4:2 Compressors for utilizing in dynamic accuracy configurable multipliers,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 4, pp. 1352–1361, Apr. 2017.
- [10] S. Venkatachalam and S.-B. Ko, “Design of power and area efficient approximate multipliers,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 5, pp. 1782–1786, May 2017.
- [11] Z. Yang, J. Han, and F. Lombardi, “Approximate compressors for error-resilient multiplier design,” in *Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Nanotechnol. Syst. (DFTS)*, Oct. 2015, pp. 183–186.
- [12] M. Ha and S. Lee, “Multipliers with approximate 4-2 compressors and error recovery modules,” *IEEE Embedded Syst. Lett.*, vol. 10, no. 1, pp. 6–9, Mar. 2018.
- [13] A. Gorantla and P. Deepa, “Design of approximate compressors for multiplication,” *ACM J. Emerg. Technol. Comput. Syst.*, vol. 13, no. 3, p. 44, May 2017.
- [14] M. S. Ansari, H. Jiang, B. F. Cockburn, and J. Han, “Low-power approximate multipliers using encoded partial products and approximate compressors,” *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 3, pp. 404–416, Sep. 2018.
- [15] M. R. Khezeli, M. H. Moaiyeri, and A. Jalali, “Comparative analysis of simultaneous switching noise effects in MWCNT bundle and Cu power interconnects in CNTFET-based ternary circuits,” *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 1, pp. 37–46, Jan. 2019.
- [16] C. Xu *et al.*, “Impact of write pulse and process variation on 22 nm FinFET-based STT-RAM design: A device-architecture co-optimization approach,” *IEEE Trans. Multi-Scale Comput. Syst.*, vol. 1, no. 4, pp. 195–206, Dec. 2015.
- [17] S. K. Gupta and K. Roy, “Low power robust FinFET-based SRAM design in scaled technologies,” in *Circuit Design for Reliability*. New York, NY, USA: Springer, 2015, pp. 223–253.
- [18] S. S. Ensan, M. H. Moaiyeri, M. Moghaddam, and S. Hessabi, “A low-power single-ended SRAM in FinFET technology,” *AEU—Int. J. Electron. Commun.*, vol. 99, pp. 361–368, Feb. 2019.
- [19] B. Swahn and S. Hassoun, “Gate sizing: FinFETs vs 32nm bulk MOSFETs,” in *Proc. 43rd Annu. Des. Automat. Conf.*, Jul. 2006, pp. 528–531.
- [20] J. Rabaey, *Low Power Design Essentials*. New York, NY, USA: Springer, 2009.
- [21] W. Ibrahim, V. Beiu, and M. H. Sulieman, “On the reliability of majority gates full adders,” *IEEE Trans. Nanotechnol.*, vol. 7, no. 1, pp. 56–67, Jan. 2008.
- [22] V. Pudi, K. Sridharan, and F. Lombardi, “Majority logic formulations for parallel adder designs at reduced delay and circuit complexity,” *IEEE Trans. Comput.*, vol. 66, no. 10, pp. 1824–1830, Oct. 2017.



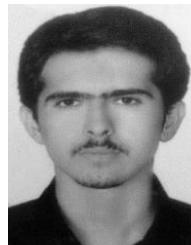
- [23] T. Zhang, W. Liu, E. McLarnon, and M. O'Neill, F. Lombardi, "Design of majority logic (ML) based approximate full adders," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.
- [24] S. Angizi, H. Jiang, R. F. DeMara, J. Han, and D. Fan, "Majority-based spin-CMOS primitives for approximate computing," *IEEE Trans. Nanotechnol.*, vol. 17, no. 4, pp. 795–806, Jul. 2018.
- [25] J. Liang, J. Han, and F. Lombardi, "New metrics for the reliability of approximate and probabilistic adders," *IEEE Trans. Comput.*, vol. 62, no. 9, pp. 1760–1771, Sep. 2013.
- [26] S. Salahuddin, H. Jiao, and V. Kursun, "A novel 6T SRAM cell with asymmetrically gate underlap engineered FinFETs for enhanced read data stability and write ability," in *Proc. Int. Symp. Qual. Electron. Design*, Mar. 2013, pp. 353–358.
- [27] G. Jaberipur, B. Parhami, and D. Abedi, "Adapting computer arithmetic structures to sustainable supercomputing in low-power, majority-logic nanotechnologies," *IEEE Trans. Sustain. Comput.*, vol. 3, no. 4, pp. 262–273, Oct./Dec. 2018.
- [28] W. Li, Y. Yang, H. Yan, and Y. Liu, "Three-input majority logic gate and multiple input logic circuit based on DNA strand displacement," *Nano Lett.*, vol. 13, no. 6, pp. 2980–2988, 2013.
- [29] S. Azimi, S. Angizi, and M. H. Moaiyeri, "Efficient and robust SRAM cell design based on quantum-dot cellular automata," *ECS J. Solid State Sci. Technol.*, vol. 7, no. 3, pp. Q38–Q45, 2018.
- [30] H. Iwamura, M. Akazawa, and Y. Amemiya, "Single-electron majority logic circuits," *IEICE Trans. Electron.*, vols. E81–C, no. 1, pp. 42–48, Jan. 1998.
- [31] R. Zimmermann and W. Fichtner, "Low-power logic styles: CMOS versus pass-transistor logic," *IEEE J. Solid-State Circuits*, vol. 32, no. 7, pp. 1079–1090, Jul. 1997.
- [32] L. Clark *et al.*, "ASAP7: A 7-nm finFET predictive process design kit," *Microelectron. J.*, vol. 53, pp. 105–115, Jul. 2016.
- [33] Accessed: 2016. [Online]. Available: <http://asap.asu.edu/asap>
- [34] H. Jiang, J. Han, and F. Lombardi, "A comparative review and evaluation of approximate adders," in *Proc. 25th Ed. Great Lakes*, May 2015, pp. 343–348.
- [35] V. Leon, G. Zervakis, D. Soudris, and K. Pekmestzi, "Approximate hybrid high radix encoding for energy-efficient inexact multipliers," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 3, pp. 421–430, Mar. 2018.
- [36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [37] V. Gupta, D. Mohapatra, S. P. Park, A. Raghunathan, and K. Roy, "IMPACT: Imprecise adders for low-power approximate computing," in *Proc. 17th IEEE/ACM Int. Symp. Low-Power Electron. Design*, Aug. 2011, pp. 409–414.



**Farnaz Sabetzadeh** received the B.Sc. and M.Sc. degrees in electrical engineering from Shahid Beheshti University, Tehran, Iran, where she is currently pursuing the Ph.D. degree in electrical engineering. Her research interests include VLSI circuit design, image processing, and approximate computing.



**Mohammad Hossein Moaiyeri** (M'12–SM'18) is currently an Assistant Professor with the Faculty of Electrical Engineering, Shahid Beheshti University. His research interests include nanoelectronic circuit design, low-power VLSI design, VLSI implementation of MVL and fuzzy logic, and mixed-signal integrated circuits design.



**Mohammad Ahmadinejad** received the B.Sc. degree in electrical engineering from Shahid Beheshti University, Tehran, Iran, where he is currently pursuing the M.Sc. degree in electrical engineering. His research interests include VLSI circuit design, nanoelectronics, and approximate computing.