


RESEARCH ARTICLE

Ultraefficient imprecise multipliers based on innovative 4:2 approximate compressors

Mohammad Khaleqi Qaleh Jooq¹ | Mohammad Ahmadinejad² |
Mohammad Hossein Moaiyeri² 

¹Nanotechnology and Quantum
Computing Lab, Shahid Beheshti
University, Tehran, Iran

²Faculty of Electrical Engineering, Shahid
Beheshti University, Tehran, Iran

Correspondence

Mohammad Hossein Moaiyeri, Shahid
Beheshti University, Tehran, Iran.
Email: h_moaiyeri@sbu.ac.ir

Summary

This study proposes two ultraefficient imprecise multipliers based on innovative 4:2 approximate compressor designs. The first proposed multiplier employs an ultracompact 8-transistor 4:2 compressor to reduce the transistor count and energy dissipation. To improve the accuracy of the first proposed imprecise multiplier, the second multiplier also benefits from a semiaccurate 24-transistor 4:2 approximate compressor for the high-order bits. The 7-nm fin field-effect transistor (FinFET) technology, as one of the leading commercial technologies, is utilized to simulate the proposed multipliers in the HSPICE[®] environment. The simulation results indicate that the proposed imprecise multipliers show significant improvements regarding transistor count, delay, power, and power-delay product as compared to their state-of-the-art imprecise and exact counterparts. Along with the superior hardware efficiency, the MATLAB simulations demonstrate that the proposed multipliers also provide reasonable levels of accuracy. Moreover, a figure of merit (FOM) considering hardware efficiency and output quality is considered in order to evaluate the multipliers comprehensively. The FOM simulation results indicate that the proposed imprecise multipliers make a significant trade-off between hardware efficiency and quality for approximate-computing applications dealing with image multiplication.

KEYWORDS

approximate compressors, approximate computing, image multiplication, imprecise multipliers

1 | INTRODUCTION

Over the past few decades, the huge progress in shrinking the size of the complementary metal-oxide-semiconductor (CMOS) technology has successfully lead to high-performance digital circuits. However, miniaturization of the size of the CMOS transistors to compact the very large-scale integration (VLSI chips) has imposed many serious concerns. These concerns such as drain-induced barrier lowering (DIBL), high leakage currents, high power density, and reliability issues have restrained the functionality of the nanoscale circuits and systems.^{1,2} Many solutions from the device level to the system level have been proposed to cope with these challenges.

From the system-level aspect, in many areas such as artificial intelligence, machine learning, and image processing, where a certain level of inaccuracy is tolerable mainly due to the error resilience of the human brain and visual system,

approximate computing has been introduced.^{3,4} In the realm of the VLSI universe, approximate computing has become a hotspot in a diverse range of applications, where the computation preciseness takes second place after energy consumption. Therefore, to address the challenges mentioned above in deep nanoscale error-tolerant VLSI applications, approximate computing can be considered in various design levels of abstraction.⁵⁻⁷

Various silicon and nonsilicon channel-based transistors have been developed to pave the way for device scaling beyond Moore's law.⁸⁻¹² The FinFET device with higher controllability over the gate region, improved I_{ON}/I_{OFF} ratio, mitigation of the threshold voltage variations by the elimination of the random dopant fluctuations, and the possibility of successful scaling to sub-10-nm technology nodes has been raised as a groundbreaking candidate for representing high-performance ultrascaled VLSI circuits and systems.^{13,14}

Arithmetic blocks such as adders and multipliers are the most essential building blocks of the digital microprocessors and microsystems. Multipliers are one the most power-hungry units, which directly influence the overall performance of the computing systems. Therefore, utilizing approximate computing methods can significantly improve the performance of the multipliers regarding power consumption, speed, and transistor count.¹⁵⁻¹⁷

Typically, multiplication is commenced by generating partial products with an array of AND gates. This process is followed by the reduction of the partial products and then the final summation for computing the final results. Among these stages, partial product reduction is the most fundamental step, which has a significant contribution to the overall energy consumption, latency, and hardware complexity.¹⁸ In an imprecise multiplier, partial product reduction can be realized through approximate compressors. Utilizing low-complexity and energy-efficient compressors with an adequate level of accuracy can remarkably enhance the entire performance of the multipliers, especially in error-resilient applications such as image processing.¹⁹⁻²²

In this study, ultraefficient imprecise multipliers utilizing innovative approximate 4:2 compressors to reduce the design complexity and energy dissipation are proposed. The first proposed approximate 4:2 compressor uses only eight transistors and significantly reduces the energy dissipation. The second proposed approximate 4:2 compressor has 24 transistors and offers a higher level of accuracy. The combination of these approximate compressors provides superior enhancements regarding the delay, power, energy, and design complexity for the proposed multipliers. Along with the extraordinary hardware performance characteristics, the proposed imprecise multipliers show promising precision, which makes them very suitable for image-processing applications. The electrical performance metrics of the approximate multipliers are investigated through HSPICE[®] simulations using 7-nm FinFET technology, while the accuracy and quality metrics are calculated in the MATLAB environment. These metrics reveal that our proposed designs make a superior trade-off between the hardware efficiency and quality metrics for approximate computing.

The remainder of this paper is organized as follows: Section 2 briefly reviews the backgrounds of the research. The proposed designs are described and discussed in Sections 3 and 4. The comprehensive simulation results and comparisons are presented in Section 5. Finally, the main results are highlighted in Section 6.

2 | BACKGROUNDS

2.1 | FinFET technology

Over the past decades, planar MOSFETs had laid the groundwork in the design and manufacturing of the high-performance circuits and systems down to 22-nm feature size. However, with the emersion of various intensive short channel effects, the advantages of these transistors have gradually dwindled. The multigate FinFET device as a leading-edge technology has been the focal point of the semiconductor industry owing to a higher I_{ON}/I_{OFF} ratio, better sub-threshold characteristics, fascinating carrier mobility, and ease of integration with the commercial techniques.^{21,22}

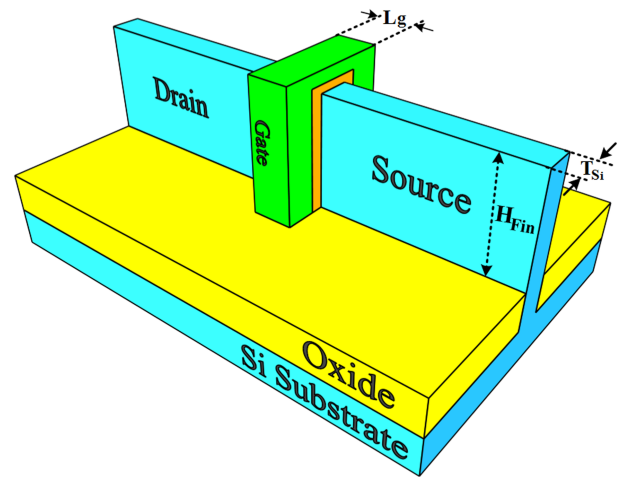
The schematic of the tri-gate FinFET is shown in Figure 1. The ON current of FinFET can be expressed by the following²³:

$$I_{ON} = kN_{Fin} \frac{2H_{Fin} + T_{Si}}{L_g} (V_{GS} - V_T)^\alpha, \quad (1)$$

where k is a fitting constant parameter, α is the velocity saturation index, V_{GS} is the gate-source voltage, V_T is the device threshold voltage, N_{Fin} is the number of fins, and H_{Fin} and T_{Si} are the height and thickness of the fins, respectively, which are assumed to be constant due to fabrication related issues.

TABLE 1 Some of key parameters for 7-nm FinFET technology node²⁴

| Device parameter | Value |
|--|------------------------------------|
| Fin thickness (T_{Si}) | 6.5 nm |
| Fin height (H_{Fin}) | 32 nm |
| Physical gate length (L_g) | 21 nm |
| Fin pitch (P_{Fin}) | 27 nm |
| Equivalent oxide thickness (EOT) | 1 nm |
| Body doping (N_{body}) | 10^{16} cm^{-3} |
| Doping of the source/drain region (N_{sd}) | $2 \times 10^{20} \text{ cm}^{-3}$ |
| Low field mobility of n-type (p-type) device | 250(210) $\text{cm}^2/\text{V.s}$ |

FIGURE 1 Schematic of the FinFET transistor [Colour figure can be viewed at wileyonlinelibrary.com]

In this work, the 7-nm FinFET technology is utilized as the platform for our demonstrations in the Synopsys HSPICE[®] tool. Table 1 indicates some of the critical parameters for the 7-nm FinFET transistors. More details about this model can be found in previous studies.^{24,25}

2.2 | 4:2 compressor

Multiplication is one of the most vital operations in VLSI arithmetic circuits and systems. Partial product reduction is a critical phase in a multiplier and has a significant impact on its performance. Compressors can be considered as efficient building blocks for partial product reduction. The fundamental schematic of a 4:2 precise compressor is shown in Figure 2, which can be realized by two full adders. In a 4:2 compressor, there are four main inputs (x_1 , x_2 , x_3 , and x_4) and two main outputs (carry and sum). Moreover, an extra carry input (C_{in}) comes from the previous stage with a lower order, and a carry output (C_{out}) is fed to the next stage with a higher order.

The Boolean function of a conventional 4:2 exact compressor can be described as follows¹⁹:

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

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

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

2.3 | Related works

Approximate computing has revolutionized the performance of the arithmetic blocks in different design levels of abstraction. Approximate compressors can be employed to enhance the efficiency of multipliers, which are among the

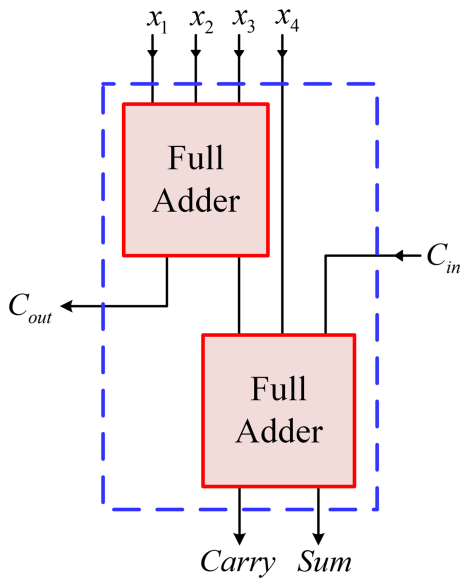


FIGURE 2 The general schematic of a 4:2 exact compressor [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/cua.2876)]

most power-hungry units in VLSI chips. Various imprecise multipliers have been suggested to restrain the energy consumption and maintain a reasonable level of accuracy.

Momeni et al. demonstrated two scenarios for designing imprecise 8-bit Dadda multipliers.¹⁹ These imprecise multipliers were developed based on two approximate 4:2 compressors realized with NOR and XNOR logics. These 4:2 compressors can be effectively demonstrated by 28 and 26 transistors. However, the second structure, which neglects the input and output carry (C_{in} and C_{out}), is more efficient regarding the accuracy and performance metrics.

To make a trade-off between preciseness and performance metrics, Akbari et al. proposed a dual-quality architecture for 4:2 compressors that can work in both exact and approximate fashions during execution.²⁶ At the transistor level, the DQ4:2C4 design as the most effective component of this work can be realized by 26 transistors.

To reduce the error rate (probability) of the approximate multipliers, Gorantla et al. investigated five scenarios to demonstrate efficient imprecise multipliers based on approximate compressors.²⁷ Despite reducing the error rate, these structures require high numbers of transistors, which degrades the electrical characteristics. The most efficient approximate 4:2 compressor proposed in Gorantla²⁷ requires 30 transistors to be realized at the transistor level.

Venkatachalam et al. presented imprecise multipliers based on approximate compressors using the OR logic along with the XOR gates to lower the switching activity factor and, thereby, the total power consumption.²⁸ At the transistor level, Venkatachalam's compressor can be designed by 36 transistors.

By modifying the truth table of the approximate 4:2 compressor presented by Yang et al.,²⁹ Ha and Lee proposed an imprecise multiplier based on a 30-transistor approximate 4:2 compressor realized with XOR and AND-OR logics.³⁰ Ansari et al. proposed efficient imprecise multipliers using an AND-OR logic-based imprecise 4:2 compressor, which consists of 20 transistors.³¹ Intending to utilize higher-order approximate compressors, Esposito et al. demonstrated an innovative imprecise multiplier to make an effective trade-off between the performance and accuracy metrics.³²

Reddy et al. proposed a modified Dadda imprecise multiplier.³³ This effective imprecise multiplier utilizes a multiplexer-based approximate 4:2 compressor, which can be realized with 30 transistors.

Ahmadinejad et al. proposed two imprecise multipliers based on approximate 4:2 and 5:2 compressors, which can be implemented with 16 and 20 transistors.²¹ Moreover, Sabetzadeh et al. suggested an imprecise multiplier for image multiplication.² The majority logic-based approximate compressor used in this multiplier is realized with 12 transistors. Edavoor et al. demonstrated an effective approximate multiplier design based on 30-transistor dual-stage 4:2 compressors.³⁴

3 | PROPOSED 4:2 APPROXIMATE COMPRESSORS

3.1 | The first proposed 4:2 compressor (8T4:2)

The truth table of the first proposed 4:2 approximate compressor is given in Table 2. In the proposed design, the C_{in} and C_{out} signals are neglected as a practical approach for reducing design complexity.¹⁹ As shown in Table 2, the sum

TABLE 2 The truth table of the first proposed approximate 4:2 compressor

| x_4 | x_3 | x_2 | x_1 | Carry | Sum | ED |
|-------|-------|-------|-------|-------|-----|----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | -1 |
| 0 | 0 | 1 | 0 | 1 | 0 | +1 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 1 | 0 | +1 |
| 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | -1 |
| 1 | 0 | 0 | 0 | 1 | 0 | +1 |
| 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 1 | 0 | 1 | 1 | 1 | 0 | -1 |
| 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| 1 | 1 | 0 | 1 | 1 | 0 | -1 |
| 1 | 1 | 1 | 0 | 1 | 0 | -1 |
| 1 | 1 | 1 | 1 | 1 | 0 | -2 |

output is considered to be “0,” and the carry output utilizes only three of four inputs (x_2 , x_3 , and x_4). In the approximate compressors, error distance (ED) is defined as the difference between the wrong and correct outputs.

As an example, in Table 2, when the input is “1,110” the decimal summation of these inputs is 3. However, the resulted approximate output for this input combination in the proposed 4:2 compressor is “10,” which corresponds to the decimal value of 2. Therefore, the error distance for this input combinations is $ED = 2 - 3 = -1$. In the proposed approximate compressor, the small EDs with opposite signs (+1 and -1) can decrease each other's impact to some extent in the structure of the imprecise multiplier.³²

In the binary multipliers, AND gates are utilized to generate partial products. As a result, assuming that the inputs of the multiplier are distributed uniformly,²⁹ the probability that a partial product equals “0” (“1”) is 75% (25%). Therefore, the error contribution for the “1,111” input combination is $(1/256 = 0.39\%)$, and the ED of -2 does not produce a considerable error in the output of our imprecise multiplier.

Based on Table 2, the Boolean functions of the proposed approximate 4:2 compressor can be expressed as follows:

$$Sum = 0, \quad (5)$$

$$Carry = x_2 + x_3 + x_4. \quad (6)$$

The schematic of the first proposed approximate 4:2 compressor is shown in Figure 3. According to Equations (5) and (6), the proposed design is a 3-input OR gate, which can be designed with only eight transistors. This ultracompact structure can remarkably reduce the power consumption, transistor count, and the propagation delay of the imprecise multiplier, while providing an acceptable output quality as compared to the previous related works.

3.2 | The second proposed 4:2 compressor (24T4:2)

The aim of the first proposed ultracompact 8-transistor approximate 4:2 compressor is to provide a superior simplicity and ultraenergy-efficient operation along with a reasonable accuracy for some imprecise applications. In designing imprecise multipliers, an effective approach to offer higher accuracy is to utilize compressors with higher accuracy for the bits of higher significance.^{30,31} Exact compressors can be served to compensate the error in approximate multipliers. However, they cost more hardware overhead and higher energy dissipation. Therefore, using semiaccurate approximate 4:2 compressors can be considered as a more efficient solution to enhance the accuracy along with having a lower impact on the electrical characteristics of the imprecise multipliers. Accordingly, we propose another approximate 4:2

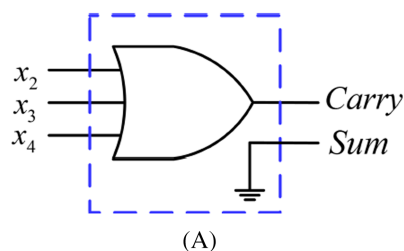


FIGURE 3 The first proposed approximate compressor. (A) The block diagram and (B) transistor-level schematic [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/cia.2876)]

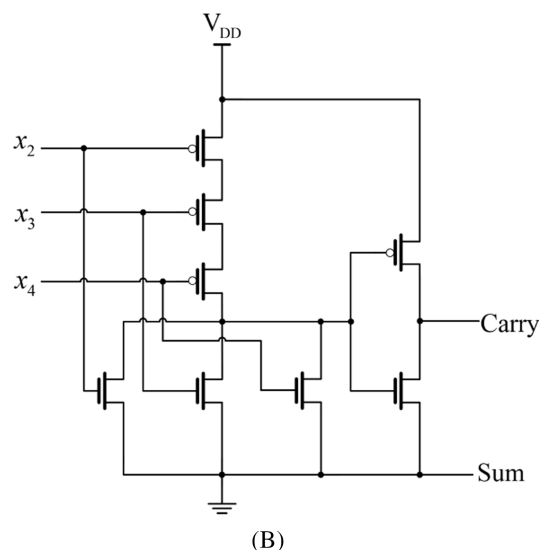


TABLE 3 The truth table of the second proposed approximate 4:2 compressor

| x_4 | x_3 | x_2 | x_1 | Carry | Sum | ED |
|-------|-------|-------|-------|-------|-----|----|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 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 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 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 | 1 | 1 | +1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 1 | -1 |
| 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | -1 |

compressor with a higher accuracy to be utilized in the columns of higher significance in the imprecise multipliers for the approximate applications, where a higher level of accuracy is desired.

The truth table of the second proposed approximate 4:2 compressor is shown in Table 3. As shown in Table 3, the second proposed approximate compressor has three “+1” and three “-1” error distances. It is noteworthy that the erroneous outputs with opposite error distance signs (such as -1 and +1) can decrease the effect of each other in the multiplier structure and can lead to lower output quality degradation.

Based on the truth table, the output equations for the second approximate 4:2 compressor are given by the following:

$$Sum = (x_2 + x_4) + (x_1 \oplus x_3) = \overline{(x_2 + x_4)} \cdot \overline{(x_1 \oplus x_3)}, \quad (7)$$

$$Carry = (x_1 + x_2) \cdot (x_3 + x_4). \quad (8)$$

Figure 4 illustrates the circuit schematic of the second proposed approximate 4:2 compressor, which has 24 transistors. One can observe that the higher accuracy of the second approximate compressor is achieved by the penalty of a higher number of transistors (24 transistors vs. eight transistors). However, it is notable that the number of transistors is still lower enough than the exact compressor (38 transistors). To implement the XNOR function, the 6-transistor full-swing transmission gate-based XNOR gates³⁵ are used to reduce the number of transistors. It is notable that to have fair comparisons, this efficient structure is also used for the previous designs, including XOR and XNOR functions. It is also worth mentioning that as the output of the XNOR gates are connected to a CMOS gate in the proposed structure, the transmission gates will not be cascaded in the reduction tree of the imprecise multiplier.

4 | PROPOSED IMPRECISE MULTIPLIERS

In this section, we propose two imprecise Dadda multipliers based on the proposed approximate compressors. The partial product reduction circuitries of the proposed inexact multipliers are shown in Figure 5. The partial products are generated using an array of CMOS AND gates. Then, the partial products are reduced using conventional CMOS half

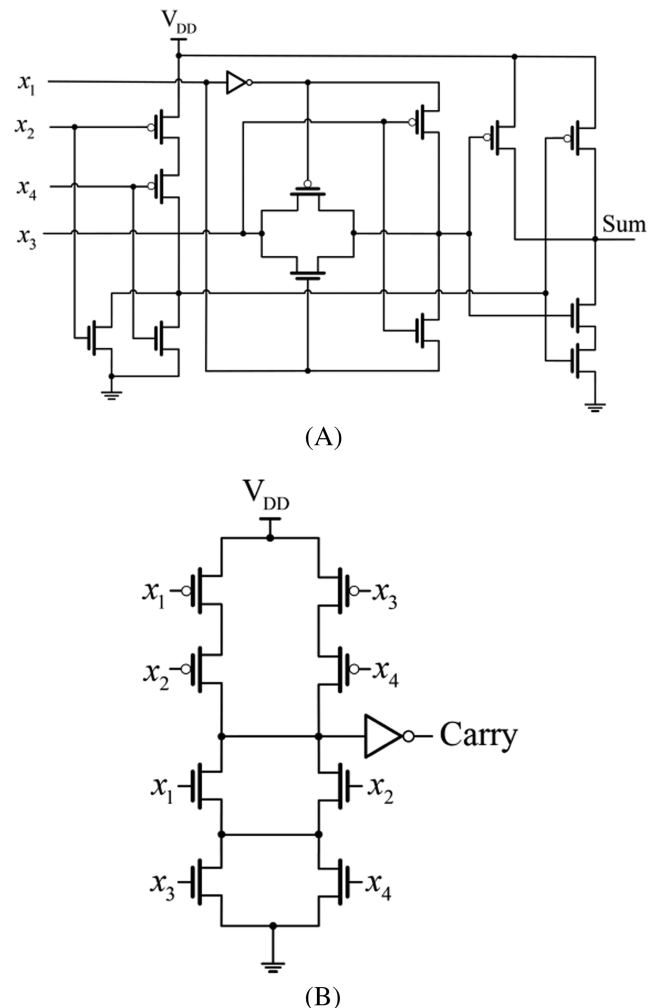


FIGURE 4 The second proposed approximate compressor: (A) sum circuit (B) carry circuit

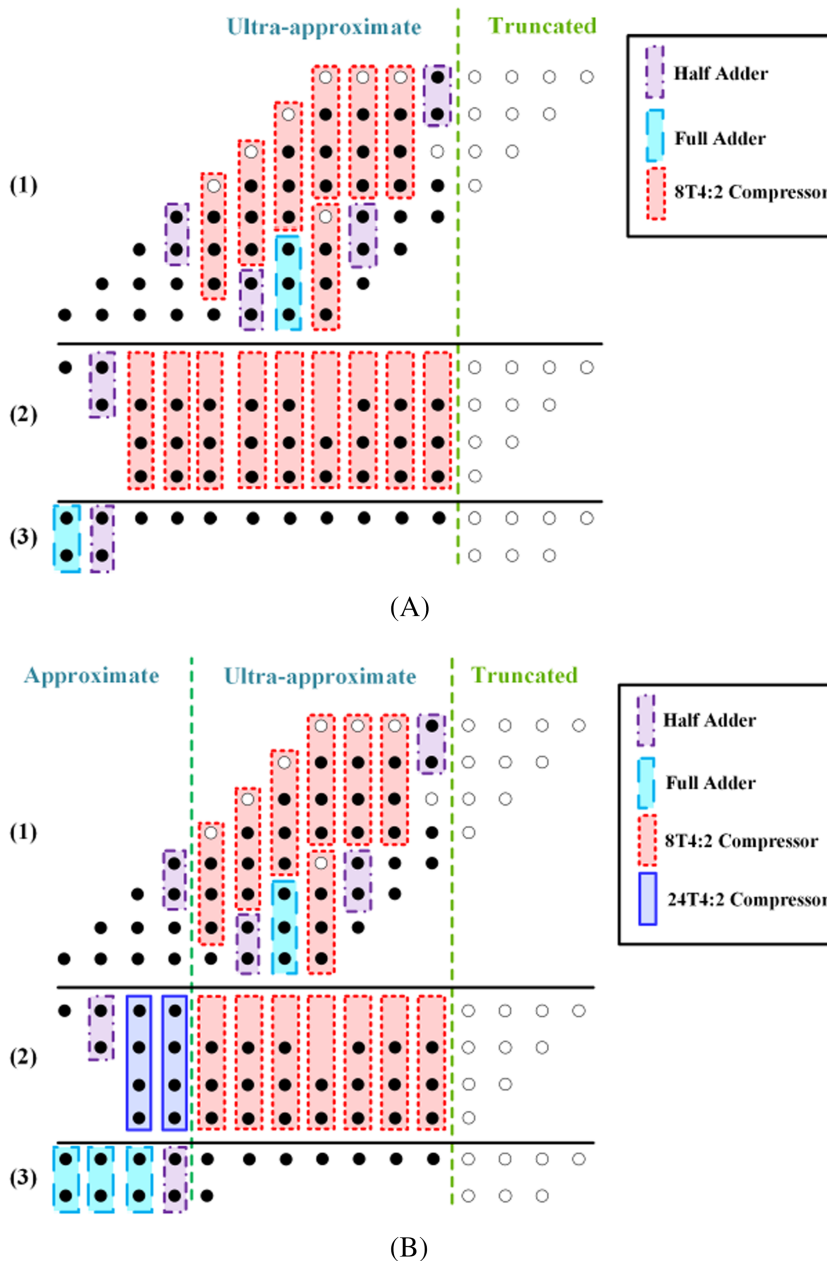


FIGURE 5 The proposed 8-bit imprecise Dadda multipliers. (A) The first proposed design (PM1). (B) The second proposed design (PM2) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/cra.2876)]

adder and mirror full adder cells³⁵ as well as the proposed compressors in two stages. Subsequently, in the final stage, a ripple carry adder (RCA) stage based on the mirror CMOS full adder³⁵ generates the final products. The four least-significant columns of the partial products are truncated to simplify the hardware implementation and reduce energy consumption with a negligible effect on the preciseness.²¹

To offer a significant hardware efficiency, the first proposed imprecise multiplier (PM1) (see Figure 5A) utilizes only the 8T4:2 compressor. However, to compensate efficiency for accuracy, the second proposed imprecise multiplier (PM2) (see Figure 5B) utilizes the 24T4:2 compressor in some of the highly significant columns.

Due to the elimination of one of the inputs (x_1) in the proposed 8T4:2 approximate compressor, some of the partial products in the ultra-approximate columns (denoted with white circles in Figure 5A,B) are not involved in the multiplication process. Therefore, there is no need to generate these partial products, which leads to the elimination of eight and seven AND gates from the first and second proposed imprecise multipliers, respectively. Moreover, due to considering "Sum = 0" in the proposed 8T4:2 approximate compressor, some of the inputs of the second stage are not generated (equals to "0"). According to this extraordinary feature, some of the full adder blocks in the RCA stage are removed, which has a significant impact on the total delay and power consumption of the multiplication process.

5 | PERFORMANCE EVALUATION

In this section, by performing comprehensive simulations, first, we benchmark the electrical performance of the proposed imprecise Dadda multipliers through evaluating various design metrics such as critical path delay, power consumption, transistor count, power-delay-product (PDP), and energy-delay product (EDP). It is noteworthy that because of the almost equal electron and hole mobilities in FinFETs as well as the circuit complexity and power dissipation concerns, we use single-fin transistors for the circuit-level demonstrations. The mean error distance (MED) and normalized mean error distance (NED) metrics are calculated by the MATLAB simulations. Finally, to investigate the functionality of the proposed imprecise multipliers in the related practical applications, we consider image multiplication and sharpening as two critical operations in the image processing applications. The quality metrics are evaluated and compared for these applications.

5.1 | Hardware analysis

In this section, we benchmark the circuit-level performance of the proposed imprecise 8-bit Dadda multipliers. For this purpose, the Synopsys HSPICE[®] tool and the 7-nm FinFET technology described in Section 2 are used. Moreover, the supply voltage of 0.7 V and the operational frequency of 2 GHz are considered for the simulations. To have fair comparisons, the previous related works are designed efficiently based on the details given in the corresponding references at the transistor level using the 7-nm FinFET technology. Furthermore, we utilize the efficient 38-transistor exact 4:2 compressor presented in Arasteh et al³⁶ as the building block of the precise Dadda multiplier. The four least-significant columns are truncated for all of the approximate multipliers. For more clarity, Table 4 gives the specifications of the previous imprecise multipliers under investigation. We use input buffers and fan-out-of-four (FO4) inverters as loads to evaluate the worst-case critical path propagation delay. Moreover, long streams of random input combinations with full switching activity are generated and used as stimuli to assess the power dissipation of the multipliers.

The performance metrics of the proposed designs and their state-of-the-art counterparts are given in Table 5. Based on the results, the first proposed imprecise multiplier shows superior improvements regarding transistors count (51%), delay (61%), power (59%), PDP (83%), and EDP (93%) as compared to its counterparts. Despite the relative increase in the electrical metrics due to the use of 24T4:2 compressors for accuracy compensation, the second design demonstrates on average 44%, 49%, 54%, 76%, and 88% improvements regarding transistor count, delay, power, PDP, and EDP, respectively, as compared to the M1-M12, M14, and exact multipliers. It can be concluded from the results that utilizing the first ultracompact approximate compressor in the proposed multipliers significantly reduces the number of transistors and provides a remarkable simplicity at the partial product reduction stage. Accordingly, by eliminating some of the full

TABLE 4 Descriptions for the previous multipliers under investigation

| Multiplier | Specifications |
|------------|--|
| M1 | Dadda imprecise multiplier based on the compressor of Momeni et al ¹⁹ |
| M2 | Dadda imprecise multiplier based on the DQ4:2C4 Akbari et al ²⁶ |
| M3 | Dadda imprecise multiplier based on the compressor of Gorantla ²⁷ |
| M4 | Dadda imprecise multiplier based on the compressor of Venkatachalam and Ko ²⁸ |
| M5 | The imprecise multiplier proposed in Venkatachalam and Ko ²⁸ |
| M6 | Dadda imprecise multiplier based on the compressor of Ha and Lee ³⁰ |
| M7 | M8_1 multiplier proposed in Ansari et al ³¹ |
| M8 | M8_2 multiplier proposed in Ansari et al ³¹ |
| M9 | Dadda imprecise multiplier based on the compressor of Ansari et al ³¹ |
| M10 | 2StepTrunc multiplier proposed in Esposito et al ³² |
| M11 | Modified Dadda imprecise multiplier proposed in Reddy et al ³³ |
| M12 | Mul_2 multiplier proposed in Ahmadinejad et al ²¹ |
| M13 | Dadda imprecise multiplier based on the compressor of Sabetzadeh et al ² |
| M14 | Dadda imprecise multiplier based on the compressor of Edavoor et al ³⁴ |

TABLE 5 Performance comparison of the multipliers

| Multipliers | Transistor count | Critical path delay (ps) | Power (μ W) | PDP (fJ) | EDP (fJ \times ps) |
|---------------------|------------------|--------------------------|------------------|----------|----------------------|
| PM1 | 532 | 88.1 | 4.08 | 0.36 | 32 |
| PM2 | 646 | 120.7 | 5.01 | 0.6 | 73 |
| M1 | 1,104 | 222.4 | 9.32 | 2.07 | 460 |
| M2 | 1,104 | 222.6 | 10.11 | 2.25 | 501 |
| M3 | 1,168 | 232.9 | 11.82 | 2.75 | 640 |
| M4 | 1,264 | 236.5 | 10.81 | 2.56 | 605 |
| M5 | 1,112 | 235.1 | 10.02 | 2.36 | 554 |
| M6 | 1,168 | 244.1 | 11.78 | 2.88 | 703 |
| M7 | 1,248 | 239.7 | 13.50 | 3.24 | 776 |
| M8 | 1,206 | 239.7 | 12.07 | 2.89 | 693 |
| M9 | 1,008 | 239.8 | 8.80 | 2.11 | 506 |
| M10 | 1,180 | 236.5 | 10.74 | 2.54 | 601 |
| M11 | 1,248 | 246.9 | 12.92 | 3.19 | 788 |
| M12 | 920 | 221.8 | 7.40 | 1.64 | 364 |
| M13 | 642 | 137.2 | 4.91 | 0.67 | 92 |
| M14 | 1,168 | 229.4 | 11.18 | 2.56 | 587 |
| Exact ³⁶ | 1,530 | 252.8 | 18.80 | 4.75 | 1,201 |

adder blocks in the third stage, a significant reduction in the delay, power, and energy dissipation is achieved. Moreover, utilizing the second proposed approximate 4:2 compressor for only the highest significant bits to improve the output accuracy does not lead to a considerable degradation in the delay and power consumption thanks to the remarkable simplicity provided by the first proposed 4:2 compressor.

5.2 | Accuracy simulations

This section investigates some of the critical error metrics to benchmark the accuracy of the imprecise multipliers. The accuracy simulations are carried out using the MATLAB software. The accuracy of the final results can be evaluated by introducing fundamental error metrics such as mean error distance (MED) and normalized mean error distance (NED), error rate (ER), and mean relative error distance (MRED). The MED metric is defined as the average error distance over all of the possible input combinations of an imprecise multiplier. The NED metric is defined as the mean error distance normalized by the maximum error. Error rate is calculated as the probability of producing an incorrect output. Furthermore, MRED is the average value of all possible relative error distances. More details about these accuracy metrics can be found in previous studies.^{2,31,37}

Table 6 shows a comparison between the accuracy metrics of the imprecise multipliers under investigation. The error metrics are obtained by considering all the possible input combinations for the 8-bit multiplier ($2^8 \times 2^8 = 65,536$). Utilizing the 24T4:2 compressors for the high-significant bits of the second proposed multiplier significantly reduces the MED and NED (about 1.77 times) metrics as compared to the first proposed design. Moreover, the M1 and M13 multipliers show the greatest MED and NED values in comparison with the other designs due to the existence of wrong outputs corresponding to the “0000” input combination in their approximate 4:2 compressors. Besides presenting fabulous hardware characteristics, the second proposed imprecise multiplier also shows better MED and NED values as compared to M1, M5, M8, M13, and M14. In terms of MRED metric, M1 and M13 show the worst performance, followed by the first and second proposed designs. However, it should be noted that the approximate applications that deal with the human senses MED and NED metrics, which are defined as the difference between the exact and inaccurate results, are more important than MRED metric that is based on the relative error distance.² Regarding the ER metric, as the four least-significant columns of the partial products are truncated, all the approximate designs show relatively high

TABLE 6 Comparison of the error metrics of the imprecise multipliers

| Multipliers | MED | NED | ER | MRED |
|-------------|-------|-------|--------|--------|
| PM1 | 2,760 | 0.042 | 0.9921 | 0.3436 |
| PM2 | 1,555 | 0.024 | 0.9915 | 0.2332 |
| M1 | 3,477 | 0.054 | 0.9996 | 4.2270 |
| M2 | 1,320 | 0.020 | 0.9178 | 0.7990 |
| M3 | 892 | 0.014 | 0.9023 | 0.0703 |
| M4 | 1,231 | 0.019 | 0.9178 | 0.0777 |
| M5 | 1,561 | 0.024 | 0.9423 | 0.0830 |
| M6 | 710 | 0.011 | 0.8832 | 0.0458 |
| M7 | 1,086 | 0.017 | 0.9085 | 0.0644 |
| M8 | 1,803 | 0.028 | 0.9085 | 0.0878 |
| M9 | 1,368 | 0.021 | 0.9351 | 0.0938 |
| M10 | 787 | 0.012 | 0.8757 | 0.0383 |
| M11 | 496 | 0.008 | 0.9309 | 0.0788 |
| M12 | 1,190 | 0.018 | 0.9169 | 0.0820 |
| M13 | 3,457 | 0.055 | 0.9998 | 4.1712 |
| M14 | 2,190 | 0.034 | 0.9805 | 0.3505 |

error rates. It is worth mentioning that in many error-resilient applications such as image processing, increasing the accuracy of the results higher than a certain value is not a suitable approach on the grounds of imposing unreasonable energy consumption.²⁰

5.3 | Application in image processing

In daily human life and many fields of science and technology, digital image processing plays an increasingly essential role with numerous applications from robotics, television, satellite, biomedical engineering, machine vision, pattern recognition, and industrial systems.³⁸ The image multiplication and sharpening as two of the most commonly used algorithms in image processing are considered for the assessments in this section. To perform image multiplication, a MATLAB program is developed to multiply two images pixel by pixel, employing the proposed imprecise multipliers. Notably, the Berkeley Segmentation Dataset (BSD) is utilized for the image multiplication and sharpening.³⁹

For the image sharpening application, the sharpened output image is determined by the following²⁶:

$$Y(x,y) = 2X(x,y) - \frac{1}{273} \sum_{i=-2}^2 \sum_{j=-2}^2 M_{\text{sharpening}}(i+3,j+3)X(x-i,y-j), \quad (9)$$

where X and Y are the input and output image, respectively, and $M_{\text{sharpening}}$ is the sharpening mask, which is defined as follows:

$$M_{\text{sharpening}} = \begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}. \quad (10)$$

In the image sharpening application, except for the multiplication operation, the rest of operations such as add, subtract, and division are assumed to be exact. To assess the quality of the resulted images, peak signal-to-noise ratio

(PSNR) and structural similarity index (SSIM) metrics were introduced. The details about these critical parameters can be found in Wang et al.⁴⁰

To have a qualitative evaluation, Figure 6 illustrates an example of a visual comparison between the image multiplication results for the exact and imprecise multipliers. According to the results, from the human vision point of view, the proposed imprecise multipliers provide appropriate and acceptable results.

Tables 7 and 8 indicate the PSNR and SSIM values for the image multiplication and sharpening applications realized by the imprecise multipliers.

Based on the comparison results, the second proposed imprecise multiplier presents better PSNR and SSIM as compared to the first one in both image sharpening and multiplication. Despite showing lower PSNRs and SSIMs than some of the previous designs (M6, M10, and M11), the proposed imprecise multipliers have superior efficiency at the hardware level, which makes them suitable for many fast and energy-efficient image-processing applications.

To have a comprehensive assessment of different imprecise multipliers and to evaluate how each of them compromises the quality for hardware efficiency, we use a figure of merit metric, based on the image processing applications as follows:

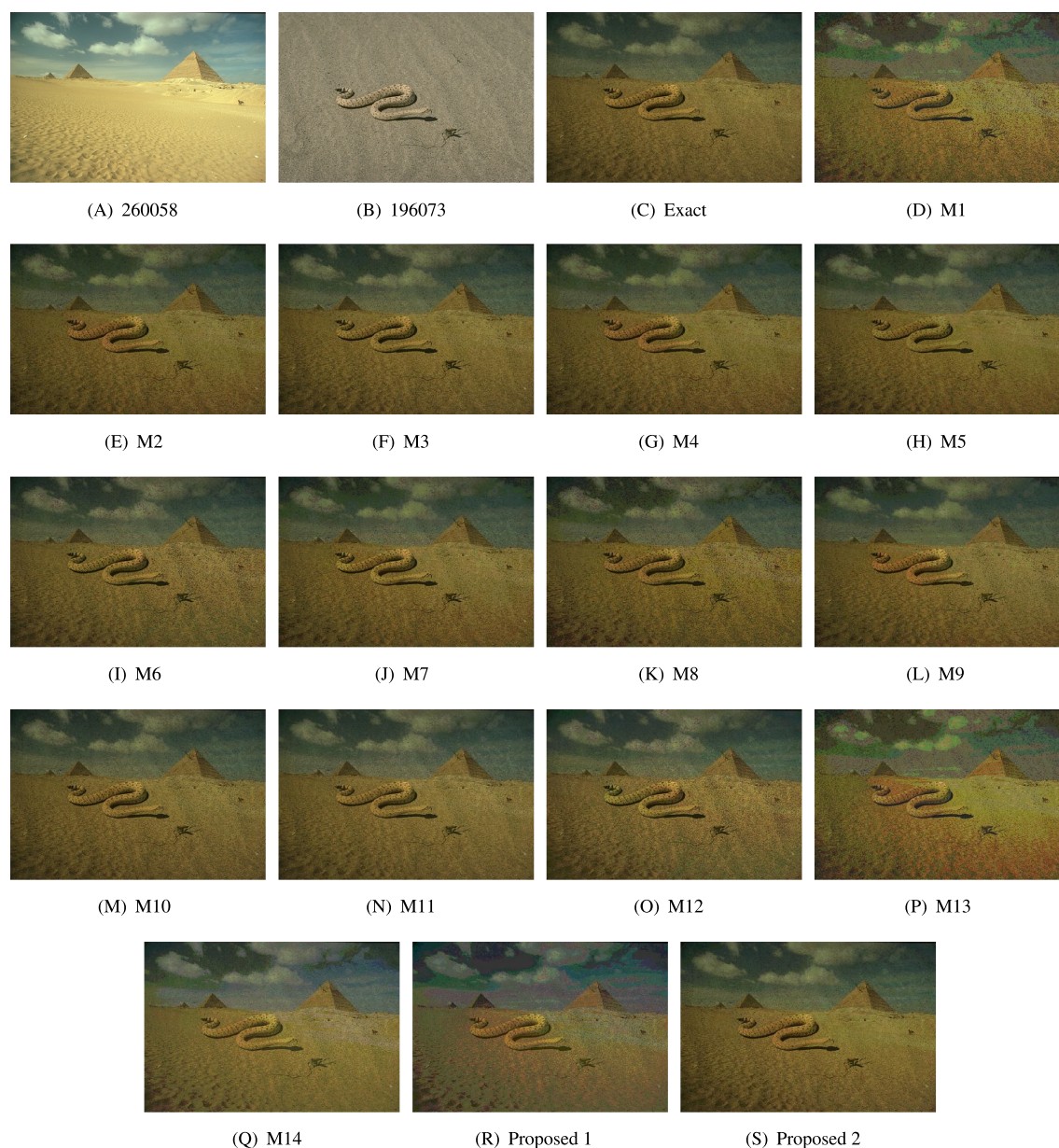


FIGURE 6 An example of the image multiplication results of different imprecise multipliers [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 7 The PSNRs (dB) and SSIMs for some of the image multiplications

| Multipliers | PSNR | | | | SSIM | | | |
|-------------|----------------------|----------------------|------------------|----------------------|----------------------|----------------------|------------------|----------------------|
| | 253,055 × 260,058 | 189,011 × 385,028 | 2,092 × 3,096 | 260,058 × 19,6073 | 253,055 × 260,058 | 189,011 × 385,028 | 2,092 × 3,096 | 260,058 × 19,6073 |
| PM1 | 24.04 | 25.12 | 25.23 | 26.12 | 0.7761 | 0.8553 | 0.7707 | 0.7598 |
| PM2 | 28.45 | 30.66 | 32.65 | 32.19 | 0.9142 | 0.9613 | 0.9267 | 0.9468 |
| M1 | 27.04 | 24.02 | 24.08 | 25.68 | 0.8861 | 0.7795 | 0.7328 | 0.8137 |
| M2 | 26.74 | 29.74 | 33.84 | 31.1 | 0.9107 | 0.9779 | 0.9576 | 0.9616 |
| M3 | 30.39 | 33.95 | 37.71 | 36.86 | 0.9579 | 0.9879 | 0.9761 | 0.9861 |
| M4 | 27.13 | 30.12 | 33.87 | 31.22 | 0.9167 | 0.9790 | 0.9578 | 0.9622 |
| M5 | 26.18 | 28.95 | 34.63 | 31.56 | 0.9115 | 0.9783 | 0.9754 | 0.9660 |
| M6 | 30.20 | 35.51 | 40.34 | 33.18 | 0.9590 | 0.9919 | 0.9863 | 0.9801 |
| M7 | 28.12 | 31.42 | 34.75 | 32.07 | 0.8889 | 0.9784 | 0.9592 | 0.9527 |
| M8 | 22.68 | 27.47 | 32.15 | 28.98 | 0.8235 | 0.9672 | 0.9490 | 0.9360 |
| M9 | 26.91 | 30.19 | 34.56 | 32.02 | 0.8741 | 0.9730 | 0.9600 | 0.9596 |
| M10 | 28.53 | 33.33 | 43.11 | 37.92 | 0.9415 | 0.9919 | 0.9931 | 0.9916 |
| M11 | 37.18 | 37.00 | 36.06 | 37.71 | 0.9846 | 0.9806 | 0.9694 | 0.9848 |
| M12 | 28.04 | 31.34 | 37.57 | 31.88 | 0.8975 | 0.9773 | 0.9755 | 0.9611 |
| M13 | 25.87 | 24.13 | 24.32 | 25.76 | 0.8558 | 0.7822 | 0.6923 | 0.8462 |
| M14 | 27.30 | 27.18 | 28.21 | 25.17 | 0.9128 | 0.9434 | 0.9015 | 0.8691 |

TABLE 8 The PSNRs (dB) and SSIMs for some of the image sharpening

| Multipliers | PSNR | | | | SSIM | | | |
|-------------|-------|--------|---------|---------|-------|--------|---------|---------|
| | 3,096 | 28,075 | 135,069 | 385,028 | 3,096 | 28,075 | 135,069 | 385,028 |
| PM1 | 25.6 | 27.70 | 30.8 | 25.4 | 0.944 | 0.967 | 0.979 | 0.967 |
| PM2 | 30.0 | 31.5 | 32.3 | 27.7 | 0.981 | 0.991 | 0.988 | 0.983 |
| M1 | 12.6 | 12.7 | 13.1 | 13.2 | 0.695 | 0.518 | 0.725 | 0.704 |
| M2 | 28.0 | 28.9 | 26.2 | 27.2 | 0.966 | 0.990 | 0.983 | 0.984 |
| M3 | 27.3 | 29.5 | 25.8 | 29.6 | 0.967 | 0.991 | 0.984 | 0.987 |
| M4 | 28.6 | 29.3 | 26.6 | 27.5 | 0.972 | 0.992 | 0.984 | 0.986 |
| M5 | 28.4 | 30.4 | 27.4 | 29.8 | 0.966 | 0.994 | 0.985 | 0.990 |
| M6 | 31.5 | 33.3 | 30.8 | 34.7 | 0.976 | 0.993 | 0.990 | 0.993 |
| M7 | 30.5 | 31.9 | 28.6 | 31.0 | 0.977 | 0.991 | 0.985 | 0.987 |
| M8 | 27.2 | 29.1 | 25.9 | 28.1 | 0.955 | 0.985 | 0.981 | 0.981 |
| M9 | 34.0 | 35.2 | 30.9 | 33.2 | 0.980 | 0.994 | 0.985 | 0.991 |
| M10 | 33.5 | 36.0 | 33.0 | 34.3 | 0.977 | 0.995 | 0.990 | 0.992 |
| M11 | 32.8 | 33.5 | 33.4 | 34.0 | 0.976 | 0.994 | 0.991 | 0.994 |
| M12 | 36.8 | 38.3 | 35.9 | 36.2 | 0.983 | 0.993 | 0.987 | 0.992 |
| M13 | 12.7 | 12.8 | 13.1 | 13.2 | 0.719 | 0.531 | 0.743 | 0.722 |
| M14 | 15.9 | 17.0 | 16.9 | 16.0 | 0.849 | 0.874 | 0.899 | 0.883 |

$$FOM = PDP_{\text{saving}} \times Transistors_{\text{saving}} \times PSNR \times MSSIM. \quad (11)$$

This figure of merit considers both hardware and quality metrics in imageprocessing applications. It also emphasizes more on the quality metrics as it is of importance to the end user. It is worth noting that the PSNR and SSIM are the average values from the image sharpening and multiplication results.

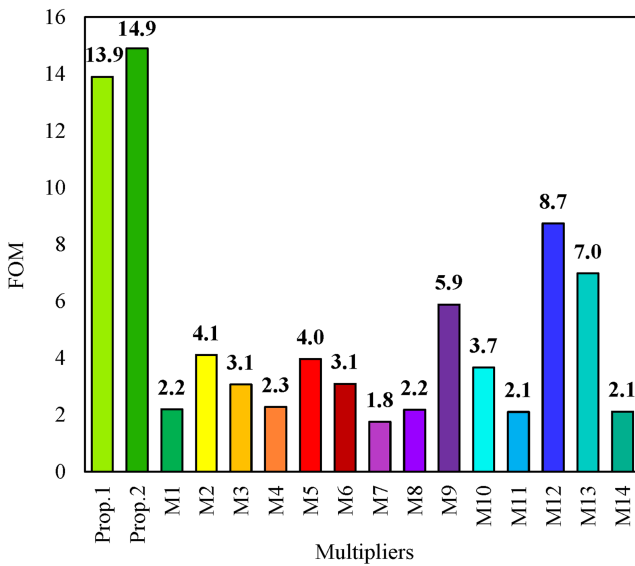


FIGURE 7 A comparison between the FOM of the imprecise multipliers [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/cra.2876)]

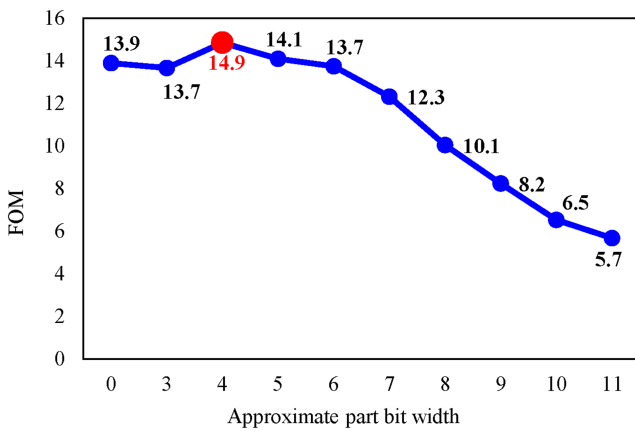


FIGURE 8 FOM versus the number columns included in the approximate part [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/cra.2876)]

Figure 7 illustrates a comparison between the FOM of the imprecise multipliers. The comparison results demonstrate that the second proposed imprecise multiplier shows higher FOM as compared to the other designs and make an effective trade-off between design efficiency and quality as an imprecise multiplier. This is on account of superior electrical characteristics along with reasonable level of accuracy of the second proposed imprecise multiplier. Significant transistor count reduction, considerable power saving, and fast delay response while providing an acceptable output quality for approximate image multiplication are the main achievements of the proposed imprecise multipliers. Based on the results shown in Figure 7, the FOM of the first proposed imprecise multiplier is 4.7 times the average FOM of the other multipliers, while this value for the second proposed design increases to five times. Therefore, the second proposed multiplier can be considered as the efficient design for image-processing applications.

In order to investigate the best combination of the approximate and ultra-approximate columns in the second proposed multiplier, the bit width of the approximate part increases from one to 11, and the FOM for each design is calculated. The results of this assessment is illustrated in Figure 8. The results demonstrate that the approximate multiplier that uses approximate compressors in four significant columns (see Figure 5) shows the best FOM.

6 | CONCLUSIONS

In this study, two ultraefficient imprecise multipliers have been proposed based on innovative approximate 4:2 compressors in 7-nm FinFET technology. Intending to reduce the number of transistors and energy dissipation, the first multiplier has utilized ultraefficient 8-transistor 4:2 approximate compressors. This superior compactness has significantly improved the transistor count, delay, power, and PDP as compared to the other state-of-the-art multipliers. The

second proposed imprecise multiplier has been designed in a way to enhance the accuracy of the output results while providing a significant energy-efficiency. A semiaccurate 24-transistor 4:2 compressor has been proposed to be used for the high-order columns of the second proposed multiplier. Along with improving the accuracy by the semiaccurate 4:2 compressors, the second proposed multiplier has shown superior performance parameters. In order to comprehensively compare the multipliers and to evaluate how each of them compromises the accuracy and quality for energy efficiency, a figure of merit based on PSNR, SSIM, NED, and PDP has been considered. According to the comparison results, the proposed multipliers have made an effective trade-off between hardware efficiency and quality for approximate computing. Our comprehensive results conclude that the proposed imprecise multipliers are promising candidates as the building blocks for ultracompact and energy-efficient approximate arithmetic circuits and systems.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Mohammad Hossein Moaiyeri  <https://orcid.org/0000-0001-9711-7923>

REFERENCES

- Esposito D, De Caro D, Di Meo G, Napoli E, Strollo AGM. Low-power hardware implementation of least-mean-square adaptive filters using approximate arithmetic. *Circ Syst Sig Process*. 2019;38(12):5606–5622.
- Sabetzadeh F, Moaiyeri MH, Ahmadinejad M. A majority-based imprecise multiplier for ultra-efficient approximate image multiplication. *IEEE Trans Circ Syst I: Reg Papers*. 2019;66(11):4200–4208.
- Priyadarshini M, Kumaravel S. Low power and area efficient error tolerant adder for image processing application. *Int J Circ Theory Appl*. 2020;48(5):696–708.
- Cilardo A, De Caro D, Petra N, Caserta F, Mazzocca N, Napoli E, Strollo AGM. High speed speculative multipliers based on speculative carry-save tree. *IEEE Trans Circ Syst I: Reg Papers*. 2014;61(12):3426–3435.
- Abed S, Khalil Y, Modhaffar M, Ahmad I. High-performance low-power approximate Wallace tree multiplier. *Int J Circ Theory Appl*. 2018;46(12):2334–2348.
- Leon V, Zervakis G, Soudris D, Pekmetzi K. Approximate hybrid high radix encoding for energy-efficient inexact multipliers. *IEEE Trans Very Large Scale Integr (VLSI) Syst*. 2017;26(3):421–430.
- Ahmadinejad M, Moaiyeri MH. Energy-efficient magnetic 5:2 compressors based on she-assisted hybrid MTJ/FINFET logic. *J Comput Electron*. 2020;19(1):206–221.
- Jooq MKQ, Mir A, Mirzakuchaki S, Farmani A. Semi-analytical modeling of high performance nano-scale complementary logic gates utilizing ballistic carbon nanotube transistors. *Phys E: Low-dim Syst Nanostruct*. 2018;104:286–296.
- Goyal C, Ubhi JS, Raj B. A low leakage TG-CNTFET-based inexact full adder for low power image processing applications. *Int J Circ Theory Appl*. 2019;47(9):1446–1458.
- Jooq MKQ, Mir A, Mirzakuchaki S, Farmani A. A robust and energy-efficient near-threshold SRAM cell utilizing ballistic carbon nanotube wrap-gate transistors. *AEU-Int J Electron Commun*. 2019;110:152874.
- Razi F, Moaiyeri MH, Rajaei R, Mohammadi S. A variation-aware ternary spin-hall assisted STT-RAM based on hybrid MTJ/GAA-CNTFET logic. *IEEE Trans Nanotechnol*. 2019;18:598–605.
- Hamedani SG, Moaiyeri MH. Comparative analysis of the crosstalk effects in multilayer graphene nanoribbon and MWCNT interconnects in sub-10 nm technologies. *IEEE Trans Electromagn Compat*. 2019;62(2):561–570.
- Xie Q, Lin X, Wang Y, Chen S, Dousti MJ, Pedram M. Performance comparisons between 7-nm FinFET and conventional bulk CMOS standard cell libraries. *IEEE Trans Circ Syst II: Express Briefs*. 2015;62(8):761–765.
- Song T, Rim W, Park S, et al. A 10 nm FinFET 128 Mb SRAM with assist adjustment system for power, performance, and area optimization. *IEEE J Solid-State Circ*. 2016;52(1):240–249.
- Mazahir S, Hasan O, Hafiz R, Shafique M. Probabilistic error analysis of approximate recursive multipliers. *IEEE Trans Comput*. 2017;66(11):1982–1990.
- Van Toan N, Lee J-G. FPGA-based multi-level approximate multipliers for high-performance error-resilient applications. *IEEE Access*. 2020;8:25481–25497.
- Pilipović R, Bulić P. On the design of logarithmic multiplier using radix-4 booth encoding. *IEEE Access*. 2020;8:64578–64590.
- Jiang H, Han J, Qiao F, Lombardi F. Approximate radix-8 booth multipliers for low-power and high-performance operation. *IEEE Trans Comput*. 2015;65(8):2638–2644.
- Momeni A, Han J, Montuschi P, Lombardi F. Design and analysis of approximate compressors for multiplication. *IEEE Trans Comput*. 2014;64(4):984–994.

20. Anusha G, Deepa P. Design of approximate adders and multipliers for error tolerant image processing. *Microprocess Microsyst.* 2020;72:102940.
21. Ahmadinejad M, Moaiyeri MH, Sabetzadeh F. Energy and area efficient imprecise compressors for approximate multiplication at nano-scale. *AEU-Int J Electron Commun.* 2019;110:152859.
22. Moaiyeri MH, Sabetzadeh F, Angizi S. An efficient majority-based compressor for approximate computing in the nano era. *Microsyst Technol.* 2018;24(3):1589–1601.
23. Gupta SK, Roy K. Low power robust FinFET-based SRAM design in scaled technologies. *Circuit design for reliability.* New York, NY: Springer; 2015:223–253.
24. Clark LT, Vashishtha V, Shifren L, et al. ASAP7: a 7-nm FinFET predictive process design kit. *Microelectron J.* 2016;53:105–115.
25. Accessed: 2016. [online]. <http://asap.asu.edu/asap>
26. Akbari O, Kamal M, Afzali-Kusha A, Pedram M. Dual-quality 4:2 compressors for utilizing in dynamic accuracy configurable multipliers. *IEEE Trans Very Large Scale Integ (VLSI) Syst.* 2017;25(4):1352–1361.
27. Gorantla A. Design of approximate compressors for multiplication. *ACM J Emerg Technol Comput Syst (JETC).* 2017;13(3):1–17.
28. Venkatachalam S, Ko S-B. Design of power and area efficient approximate multipliers. *IEEE Trans Very Large Scale Integ (VLSI) Syst.* 2017;25(5):1782–1786.
29. Yang Z, Han J, Lombardi F. Approximate compressors for error-resilient multiplier design. In: 2015 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFTS) IEEE; 2015; Amherst, MA, USA:183–186.
30. Ha M, Lee S. Multipliers with approximate 4–2 compressors and error recovery modules. *IEEE Embedded Syst Lett.* 2017;10(1):6–9.
31. Ansari MS, Jiang H, Cockburn BF, Han J. Low-power approximate multipliers using encoded partial products and approximate compressors. *IEEE J Emerg Sel Top Circ Syst.* 2018;8(3):404–416.
32. Esposito D, Strollo AGM, Napoli E, De Caro D, Petra N. Approximate multipliers based on new approximate compressors. *IEEE Trans Circ Syst I: Reg Papers.* 2018;65(12):4169–4182.
33. Reddy KM, Vasantha MH, Kumar YBN, Dwivedi D. Design and analysis of multiplier using approximate 4-2 compressor. *AEU-Int J Electron Commun.* 2019;107:89–97.
34. Edavoor PJ, Raveendran S, Rahulkar AD. Approximate multiplier design using novel dual-stage 4:2 compressors. *IEEE Access.* 2020;8:48337–48351.
35. Zimmermann R, Fichtner W. Low-power logic styles: CMOS versus pass-transistor logic. *IEEE J Solid-State Circ.* 1997;32(7):1079–1090.
36. Arasteh A, Moaiyeri MH, Taheri M, Navi K, Bagherzadeh N. An energy and area efficient 4:2 compressor based on FinFETs. *Integration.* 2018;60:224–231.
37. Liang J, Han J, Lombardi F. New metrics for the reliability of approximate and probabilistic adders. *IEEE Trans Comput.* 2012;62(9):1760–1771.
38. Pitas I. *Digital Image Processing Algorithms and Applications.* New York: John Wiley & Sons; 2000.
39. Berkeley segmentation dataset, reports, 2015. <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>
40. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* 2004;13(4):600–612.

How to cite this article: Khaleqi Qaleh Jooq M, Ahmadinejad M, Moaiyeri MH. Ultraefficient imprecise multipliers based on innovative 4:2 approximate compressors. *Int J Circ Theor Appl.* 2021;49:169–184. <https://doi.org/10.1002/cta.2876>