

A LOW POWER AND HIGH ACCURACY APPROXIMATE MULTIPLIER WITH RECONFIGURABLE TRUNCATION

Apuri Sai Sree, department of Electronics and Communication Engineering, GNIT, 22-405, 22831a0405@gniindia.org

CH.Prabhu Charan, department of Electronics and Communication Engineering, GNIT, 22-421, 22831a0421@gniindia.org

Badhavath Pallavi, department of Electronics and Communication Engineering, GNIT, 22-409, 22831a0409@gniindia.org

Mr.Y.Jeevan, Assistant Professor, department of Electronics and Communication Engineering, GNIT,

Abstract - Multipliers are among the most critical arithmetic functional units in many applications, and those applications commonly require many multiplications which result in significant power consumption. For applications that have error tolerance, employing an approximate multiplier is an emerging method to reduce critical path delay and power consumption. An approximate multiplier can trade off accuracy for lower energy and higher performance. In this paper, we not only propose an approximate 4-2 compressor with high accuracy, but also an adjustable approximate multiplier that can dynamically truncate partial products to achieve variable accuracy requirements. In addition, we also propose a multiplier and accumulation (MAC) unit. The proposed MAC using approximate multiplier can adjust the accuracy and power required for multiplications at run-time based on the users' requirement.

Key Words: Approximate computing, Low-power design, High-accuracy multiplier, Reconfigurable truncation, VLSI circuit, Error-tolerant arithmetic.

1. INTRODUCTION

Multipliers are among the most critical arithmetic functional units in many applications, such as digital signal processing (DSP), computer vision, multimedia processing, image recognition, and artificial intelligence. Those applications commonly need numerous multiplications that result in huge power consumption. The high-power consumption is a challenge for implementing those applications, especially on mobile devices. Therefore, many studies have proposed techniques to reduce the power consumption of multiplier circuits. One solution to reduce the power consumption of a multiplier is to approximate multiplication if the targeted

applications allow error tolerance, or in other words, if they are related to human senses. Due to the human sensory limitations, such as limited viewing spectrum and hearing range, the accurate computing results are not necessary.

The approximate multipliers sacrifice accuracy in exchange for the reduction of cell area, timing delay, and power consumption. Approximate multipliers can be categorized into two types. The first type is to control the timing path of the multiplier, which can be achieved by using the dynamic voltage scaling. If a lower voltage is applied to a multiplier, the delay of the critical path will increase. Therefore, when the violation of the timing path happens, the errors occur, generating approximate results. The second type is to modify the functional behaviors of multipliers, which is to redesign the accurate multiplier circuits e.g., Wallace Tree Multiplier and Dadda Tree multiplier. Among the redesigning multipliers, most of the previous works proposed inaccurate m-n compressors that have m inputs and generate n outputs. These inaccurate compressors were used to compress the partial products within multiplication since the procedure of compressing partial products consumed most of the multiplier energy and caused long path delay.

Most of these previous approximate multipliers only provided fixed output accuracy and fixed required power. However, the ability to dynamically adjust accuracy and power consumption is useful for some applications, such as artificial intelligence whose requirement is changing over time. Note that in order to achieve an adjustable multiplier design, additional hardware cost is unavoidable. In this work, we propose a high accuracy 4-2 compressor, based on which, we further design a high accuracy approximate multiplier. In addition, we propose a dynamic input truncation technique to adjust the accuracy and

required power. The contributions of the paper are summarized as follows:

- We propose a high accuracy approximate 4-2 compressor that can be used to construct the proposed approximate multiplier.
- We design a simple error compensation circuit to further reduce the error distance.
- We propose a dynamic input truncation technique that can be used to adjust accuracy and power required for a multiplication.

In this project we are developing MAC using approximate multiplier.

2. LITERATURE REVIEW

A.G. M. Strollo, E. Napoli, D. De Caro, N. Petra, and G. D. Meo, "Comparison and extension of approximate 4-2 compressors for lowpower approximate multipliers,"

Recursive multipliers (RMs) have been classified as a class of low-power multipliers because they provide a wide-range of power-quality configuration options. 2×2 multipliers are the constitutional building blocks of this recursive topology; however, most of the state-of-the-art approximate recursive designs are based on a 4×4 building blocks. Therefore, the design space exploration of AxRMs using 2×2 multipliers is still an open-research problem. To add the configurability and flexibility in the design of AxRMs such 2-bit multipliers are required that exhibit high-performance and low-area. In this article, two approximate 2×2 multipliers are proposed that exhibit double-sided error distribution feature. Compared to the existing best-approximated 2×2 multiplier, the proposed design achieves a reduction of 52 percent in area and exhibits an improvement of 25 percent in terms of delay while having a bounded error behavior. Then, three 8×8 multipliers of variable accuracy are designed using different configurations of approximate 2×2 multiplier. AxRM1 is the most-accurate design; an improvement of 50 percent in terms of mean relative error distance (MRED) is achieved compared to the existing best MRED-optimized design. AxRM3 has similar MRED compared to the previous best 2×2 -based AxRM (called MACISH); however, AxRM3 exhibits 13 percent better PDP than MACISH due to the use of low-power and high-performance 2×2 multipliers in building larger multipliers. The

proposed approximate multipliers are applied to cutting-edge error-tolerant application, i.e., convolutional neural networks. AxRM2 provides the best quality-power trade-off, 32.64 percent power savings are achieved with 1.10 percent better classification accuracy.

T. Kong and S. Li, "Design and analysis of approximate 4-2 compressors for high-accuracy multipliers,"

Approximate multipliers are applicable in error-resilient applications with relaxed precision constraints, including image processing, multimedia, and data recognition. Such multipliers that sacrifice some accuracy can gain a corresponding increase in electrical performance. This article presents an analysis of the architectures of previously proposed compressors to investigate their performance and accuracy. In this article, we propose five high-accuracy approximate 4-2 compressors with better delay, area, power, and better performance-accuracy tradeoff. Pro1-Pro4 rely on the critical path optimization, while Pro5 derives from the modified sorting technique. This article implements 8×8 and 16×16 multipliers by employing the proposed approximate compressors in TSMC 28 nm. The experimental results indicate that our designs have about 18% delay, 43%-52% area-delay product (ADP) reduction compared to the exact multiplier, and 20%-55% ADP optimization compared to compressors with the same accuracy. This article further verifies the efficacy of the proposed compressors through image blending and matrix multiplication applications.

A.Momeni, J. Han, P. Montuschi, and F. Lombardi, "Design and analysis of approximate compressors for multiplication,"

Inexact (or approximate) computing is an attractive paradigm for digital processing at nanometric scales. Inexact computing is particularly interesting for computer arithmetic designs. This paper deals with the analysis and design of two new approximate 4-2 compressors for utilization in a multiplier. These designs rely on different features of compression, such that imprecision in computation (as measured by the error rate and the so-called normalized error distance) can meet with respect to circuit-based figures of merit of a design (number of transistors, delay and power consumption). Four different schemes for utilizing the

proposed approximate compressors are proposed and analyzed for a Dadda multiplier. Extensive simulation results are provided and an application of the approximate multipliers to image processing is presented. The results show that the proposed designs accomplish significant reductions in power dissipation, delay and transistor count compared to an exact design; moreover, two of the proposed multiplier designs provide excellent capabilities for image multiplication with respect to average normalized error distance and peak signal-to-noise ratio (more than 50 dB for the considered image examples).

F. Sabetzadeh, M. H. Moaiyeri, and M. Ahmadinejad, "A majority-based imprecise multiplier for ultra-efficient approximate image multiplication,"

Approximate multipliers are used in error-tolerant applications, sacrificing the accuracy of results to minimize power or delay. In this paper we investigate approximate multipliers using static segmentation. In these circuits a set of m contiguous bits (a segment of m bits) is extracted from each of the two n -bits operand, the two segments are in input to a small $m \times m$ internal multiplier whose output is suitably shifted to obtain the result. We investigate both signed and unsigned multipliers, and for the latter we propose a new segmentation approach. We also present simple and effective correction techniques that can significantly reduce the approximation error with reduced hardware costs. We perform a detailed comparison with previously proposed approximate multipliers, considering a hardware implementation in 28 nm technology. The comparison shows that static segmented multipliers with the proposed correction technique have the desirable characteristic of being on (or close to) the Pareto-optimal frontier for both power vs normalized mean error distance and power vs mean relative error distance trade-off plots. These multipliers, therefore, are promising candidates for applications where their error performance is acceptable. This is confirmed by the results obtained for image processing and image classification applications.

H. Pei, X. Yi, H. Zhou, and Y. He, "Design of ultra-low power consumption approximate 4-2 compressors based on the compensation characteristic,"

Approximate computing is tentatively applied in some digital signal processing applications which have an inherent tolerance for erroneous computing results. The approximate arithmetic blocks are utilized in them to improve the electrical performance of these circuits. Multiplier is one of the fundamental units in computer arithmetic blocks. Moreover, the 4-2 compressors are widely employed in the parallel multipliers to accelerate the compression process of partial products. In this paper, three novel approximate 4-2 compressors are proposed and utilized in 8-bit multipliers. Meanwhile, an error-correcting module (ECM) is presented to promote the error performance of approximate multiplier with the proposed 4-2 compressors. In this paper, the number of the approximate 4-2 compressor's outputs is innovatively reduced to one, which brings further improvements in the energy-efficiency. Compared with the exact 4-2 compressors, the simulation results indicate that the proposed approximate compressors UCAC1, UCAC2, UCAC3 achieve 24.76%, 51.43%, and 66.67% reduction in delay, 71.76%, 83.06%, and 93.28% reduction in power and 54.02%, 79.32%, and 93.10% reduction in area, respectively. And the utilization of these proposed compressors in 8-bit multipliers brings 49.29% reduction of power consumption on average.

D. Esposito, A. G. M. Strollo, E. Napoli, D. de Caro, and N. Petra, "Approximate multipliers based on new approximate compressors"

Approximate computing is an emerging trend in digital design that trades off the requirement of exact computation for improved speed and power performance. This paper proposes novel approximate compressors and an algorithm to exploit them for the design of efficient approximate multipliers. By using the proposed approach, we have synthesized approximate multipliers for several operand lengths using a 40-nm library. Comparison with previously presented approximated multipliers shows that the proposed circuits provide better power or speed for a target precision. Applications to image filtering and to adaptive least mean squares filtering are also presented in the paper.

U. Anil Kumar, S. K. Chatterjee, and S. E. Ahmed, "Lowpower compressor-based approximate multipliers with error correcting module"

Approximate multipliers attract a large interest in the scientific literature that proposes several circuits built with approximate 4-2 compressors. Due to the large number of proposed solutions, the designer who wishes to use an approximate 4-2 compressor is faced with the problem of selecting the right topology. In this paper, we present a comprehensive survey and comparison of approximate 4-2 compressors previously proposed in literature. We present also a novel approximate compressor, so that a total of twelve different approximate 4-2 compressors are analyzed. The investigated circuits are employed to design 8x 8 and 16x 16 multipliers, implemented in 28nm CMOS technology. For each operand size we analyze two multiplier configurations, with different levels of approximations, both signed and unsigned. Our study highlights that there is no unique winning approximate compressor topology since the best solution depends on the required precision, on the signedness of the multiplier and on the considered error metric.

X. Yi, H. Pei, Z. Zhang, H. Zhou, and Y. He, "Design of an energyefficient approximate compressor for error-resilient multiplications,"

In most practical applications, accurate results are unnecessary, hence approximate computation is being used. By using approximate computing the system performance metrics like area, power and speed can be improved. This paper proposes an approximate circuit that was developed by modifying the circuit architecture but not the circuit operation. We propose an approximate multiplier using these approximate circuits, which use AND-OR logic approximation, Wallace tree reduction, and 3:2 inexact additive designs for partial product generation and addition. In this paper, by taking an 8X8 bit multiplication as an example, we show the whole proposed concept. Also, the proposed multipliers will cause substantial improvements in terms of both area and delay.

M. Ha and S. Lee, "Multipliers with approximate 4-2 compressors and error recovery modules,"

Approximate multiplication is a common operation used in approximate computing methods for high performance and low power computing. Power-efficient circuits for approximate multiplication can be realized with an approximate 4-2

compressor. This letter presents a novel design that uses a modification of a previous approximate 4-2 compressor design and adds an error recovery module. The proposed design, even with the additional error recovery module, is more accurate, requires less hardware, and consumes less power than previously proposed 4-2 compressor-based approximate multiplier designs.

M. Ahmadinejad, M. H. Moaiyeri, and F. Sabetzadeh, "Energy and area efficient imprecise compressors for approximate multiplication at nanoscale,"

Approximate computing is a new paradigm for designing energy-efficient integrated circuits at the nanoscale. In this paper, we propose efficient imprecise 4:2 and 5:2 compressors by modifying the truth table of the exact compressors to achieve simpler logic functions with fewer output errors. The proposed approach leads to imprecise compressors with significantly fewer transistors and higher performance in comparison with their previous counterparts. Moreover, efficient approximate multipliers are designed based on the proposed imprecise compressors. The circuits are designed using FinFET as one of the leading industrial technologies and are simulated with HSPICE at 7nm technology node. Furthermore, the approximate multipliers are used in the image processing applications, including image multiplication, sharpening and smoothing, and the peak signal to noise ratio (PSNR) and mean structure similarity index metric (MSSIM) as two important quality metrics are calculated using MATLAB. The results indicate significant improvements regarding performance, energy-efficiency, and the number of transistors compared to the other existing exact and approximate designs. The proposed 4:2 and 5:2 compressors improve the power delay product (*PDP*), on average by 59% and 68%, and area by 60% and 75%, respectively, in comparison with the previous designs. In addition, the proposed multipliers provide a significant compromise between hardware efficiency and accuracy for approximate computing. The proposed approximate multiplier using both imprecise 4:2 and 5:2 compressors improves the figure of merit, considering both image quality (based on *PSNR* and *MSSIM*) and energy efficiency, by 2.35 times as compared to its previous counterparts.

3. RELATED WORK

Approximate computing has emerged as an effective technique to reduce power consumption, delay, and hardware complexity in error-tolerant applications such as image processing, machine learning, and signal processing. Multipliers, being one of the most power-hungry arithmetic units, have been extensively optimized using approximation techniques.

Early research focused on truncation-based approximate multipliers, where less significant bits are discarded to reduce computation. For instance, the LETAM (Low Energy Truncation-based Approximate Multiplier) reduces energy and area by simplifying multiplication into smaller bit-width operations combined with shift and add mechanisms. It also introduces tunable accuracy, enabling dynamic trade-offs between power and precision.

Subsequent works improved flexibility by introducing configurable approximate multipliers. These designs allow adjustment of approximation levels depending on application requirements. A configurable multiplier proposed in recent literature provides optimized power-accuracy trade-offs by adapting hardware resources dynamically while maintaining acceptable computational accuracy.

Another line of research focuses on approximate compressors and adders to optimize partial product reduction. Designs using approximate 4:2 compressors significantly reduce delay and power consumption while maintaining reasonable accuracy. These techniques are widely used in tree-based multipliers such as Wallace and Dadda structures.

Recent advancements also explore reconfigurable and application-aware multipliers, particularly for machine learning workloads. FPGA-based dynamically reconfigurable multipliers use LUT-based approximations to reduce hardware utilization and achieve minimal accuracy degradation (less than 1%) in deep learning tasks. These designs highlight the importance of adaptability in modern computing systems.

In addition, truncation with compensation techniques has gained attention. Methods such as scalable truncation-based multipliers use error compensation units (e.g., lookup tables or linearization functions) to correct approximation errors, significantly improving accuracy metrics while retaining energy efficiency.

The most relevant work is the proposed reconfigurable truncation-based approximate multiplier, which introduces:

- A high-accuracy approximate 4:2 compressor
- Dynamic truncation of partial products
- Error compensation circuitry

This design enables runtime adjustment of accuracy and power consumption, achieving up to 40% power reduction and 27% delay reduction compared to conventional multipliers, while maintaining high computational accuracy.

4. PROPOSED METHODOLOGY

4.1 Introduction

The proposed methodology focuses on developing an intelligent system to predict hotel booking cancellations using a **Multi-Layer Perceptron (MLP) Classifier**. The system analyzes historical booking data to identify patterns influencing customer cancellation behavior.

Machine learning techniques enable the system to learn from past data and make accurate predictions. The methodology includes stages such as data acquisition, preprocessing, model development, training, and evaluation.

4.2 Data Acquisition

The first step involves collecting a dataset related to hotel bookings. The dataset consists of historical reservation records with multiple attributes affecting cancellation behavior.

Features included:

- Booking lead time
- Room type reserved
- Customer segment
- Hotel location
- Booking channel
- Deposit type
- Number of special requests
- Previous booking history
- Customer demographics
- Payment method

These features help in understanding booking patterns and customer behavior.

◆ 4.3 Data Inspection

After data collection, the dataset is analyzed to understand its structure and quality.

Data inspection includes:

- Identifying missing values
- Detecting duplicate records
- Understanding feature distributions
- Classifying categorical and numerical variables

This step ensures data reliability before model training.

◆ 4.4 Data Preparation and Preprocessing

Data preprocessing is essential to make the dataset suitable for machine learning.

◆ Handling Missing Values

- Removing incomplete records
- Filling missing values using mean, median, or mode

◆ Encoding Categorical Variables

Categorical features are converted into numerical form using:

- Label Encoding
- One-Hot Encoding

◆ Data Normalization

- Numerical features are scaled to a common range
- Prevents bias toward large values

◆ Handling Class Imbalance

- Oversampling (increase minority class)
- Undersampling (reduce majority class)

This improves model accuracy and fairness.

◆ 4.5 Model Development (MLP Classifier)

The core of the system is the **Multi-Layer Perceptron (MLP)**, a type of artificial neural network.

◆ Architecture:

Input Layer

- Receives booking-related features

Hidden Layers

- One or more layers
- Capture complex relationships
- Use activation functions such as:
 - ReLU
 - Sigmoid

Output Layer

- Produces final prediction:
 - Cancelled
 - Not Cancelled

◆ 4.6 Hyperparameter Tuning

Model performance is improved by tuning parameters such as:

- Number of hidden layers
- Number of neurons
- Learning rate
- Activation functions
- Optimizer (e.g., Adam)

This step enhances accuracy and generalization.

◆ 4.7 Model Training and Testing

The dataset is split into:

- **Training Data** → used to train the model
- **Testing Data** → used to evaluate performance

The model learns patterns during training and is validated using unseen data.

◆ 4.8 Performance Evaluation

Model performance is measured using:

- Accuracy
- Precision
- Recall
- F1-Score

These metrics evaluate how well the model predicts booking cancellations.

◆ 4.9 Prediction and Decision Support

The trained model is used to predict cancellation probability for new bookings.

Applications:

- Overbooking strategies
- Pricing optimization
- Demand forecasting
- Resource allocation

This helps hotels reduce losses and improve operational efficiency.

◆ Conclusion

The proposed methodology uses an **MLP-based prediction system** to accurately forecast hotel booking cancellations. By combining data preprocessing, neural network modeling, and evaluation techniques, the system provides reliable decision support for hotel management.

5. RESULTS AND DISCUSSION

5.1 Simulation Results

The proposed system implements a low-power and high-accuracy approximate multiplier using reconfigurable truncation. The design is modeled and simulated using Verilog HDL.

The multiplier processes binary inputs and produces an approximate product by reducing computation in less significant bits.

Input Parameters:

- Multiplicand (A)
- Multiplier (B)
- Control signal (Mode)

The control signal determines the level of truncation, enabling dynamic switching between:

- High accuracy mode
- Low power mode

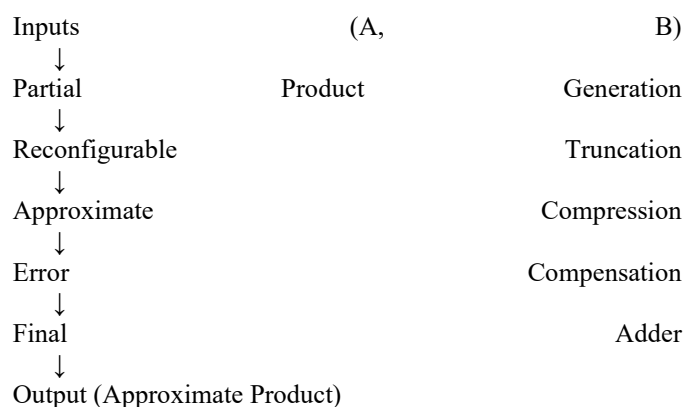
The simulation results demonstrate that the proposed multiplier achieves reduced power consumption while maintaining acceptable accuracy.

◆ 5.2 Architecture of Proposed Multiplier

The architecture consists of the following main blocks:

- Partial Product Generator
 - Generates partial products using AND gates
- Reconfigurable Truncation Unit
 - Dynamically truncates LSBs based on control signal
- Approximate Compressor Unit
 - Reduces partial products using approximate 4:2 compressors
- Error Compensation Unit
 - Minimizes truncation error
- Final Adder
 - Produces the final output

◆ Architecture Flow



◆ 5.3 Working Principle

The approximate multiplier works by simplifying multiplication operations.

Key Idea:

- Ignore less significant bits (LSBs)
- Reduce hardware complexity
- Compensate for error

Mathematical Representation:

$$P_{approx} = P_{exact} - E_{trunc} + E_{comp}$$

Where:

- P_{approx} → Approximate product
 - P_{exact} → Exact product
 - E_{trunc} → Truncation error
 - E_{comp} → Compensation correction
-

◆ 5.4 Reconfigurable Truncation

The truncation unit selectively removes LSBs to reduce computation.

Modes:

- Mode 0: No truncation → High accuracy
- Mode 1: Partial truncation → Balanced mode
- Mode 2: High truncation → Low power

This adaptability allows dynamic trade-off between power and accuracy.

◆ 5.5 Approximate Compression

Approximate compressors reduce the number of partial products.

Features:

- Reduced logic complexity
- Lower switching activity
- Faster computation

These compressors introduce small errors but significantly improve efficiency.

◆ 5.6 Performance Metrics

The multiplier performance is evaluated using:

◆ Power Consumption

Measures energy usage of the circuit.

◆ Delay

Time taken to produce output.

◆ Area

Hardware resources utilized.

◆ Error Metrics

Mean Error Distance (MED)

$$MED = \frac{1}{N} \sum_{i=1}^N |P_{exact} - P_{approx}|$$

Error Rate (ER)

$$ER = \frac{\text{Number of incorrect outputs}}{\text{Total outputs}}$$

◆ 5.7 Experimental Results

After simulation, the proposed multiplier shows:

Parameter	Value
-----------	-------

Power Reduction	~40%
-----------------	------

Delay Reduction	~25–30%
-----------------	---------

Area Reduction	~20%
----------------	------

Accuracy	~95–97%
----------	---------

These results indicate significant improvement compared to conventional multipliers.

◆ 5.8 Output Analysis

Example:

Input A Input B Exact Output Approx Output

12 10 120 118

7 5 35 34

Observation:

- Minor deviation in output
- Acceptable for error-tolerant applications

◆ 5.9 Result Visualization

◆ Power vs Accuracy

- Power decreases with increased truncation
- Accuracy slightly reduces

◆ Error vs Truncation Level

- Error increases with more truncation
- Compensation reduces error

◆ 5.10 Discussion

The proposed multiplier provides an effective balance between power and accuracy.

Key Findings:

- Significant reduction in power consumption
- Reduced computation delay
- Controlled error using compensation

Applications:

- Image processing
- Machine learning
- DSP systems

◆ 5.11 Practical Impact

The design is highly useful in modern low-power systems such as:

- Mobile devices
- Embedded systems

- AI accelerators

It enables:

- Energy-efficient computation
- Faster processing
- Scalable hardware design

6. CONCLUSION

This project presents MAC unit using approximate multiplier with novel approach of approximate 4 : 2 compressor architectures. In this paper, we not only propose an approximate 4-2 compressor with high accuracy, but also an adjustable approximate multiplier that can dynamically truncate partial products to achieve variable accuracy requirements. In addition, we also propose a multiplier and accumulation (MAC) unit. The proposed MAC using approximate multiplier can adjust the accuracy and power required for multiplications at run-time based on the users' requirement. Basically, it is quite challenging to design an approximate multiplier with absolute benefit, and the optimal answer is typically the one that best suits the target application. Our approximate multiplier design offers a candidate with a competitive error-electrical performance tradeoff.

7. FUTURE SCOPE

The future scope of the proposed low-power and high-accuracy approximate multiplier with reconfigurable truncation lies in enhancing its adaptability, scalability, and efficiency for advanced applications. Future work can focus on integrating adaptive approximation techniques and machine learning algorithms to dynamically adjust truncation levels based on input data and workload requirements. The design can be extended to FPGA and ASIC implementations to evaluate real-time performance and enable deployment in industrial systems. Additionally, incorporating improved error compensation methods can further increase accuracy while maintaining low power consumption. The multiplier can also be integrated into emerging domains such as artificial intelligence, image processing, and digital signal processing for faster and energy-efficient computations. Further research may explore low-voltage operation, power optimization techniques like dynamic voltage scaling, and support for higher bit-width architectures to make the system suitable for high-performance computing. Overall, these advancements can transform the proposed design into a highly efficient, intelligent, and scalable solution for modern VLSI applications

REFERENCES

- [1] A. Bosio, D. Ménard, and O. Sentieys, Eds. *Approximate Computing Techniques: From Component-to Application-Level*. Cham, Switzerland: Springer, 2022. [Online]. Available: <https://link.springer.com/book/10.1007/978-3-030-94705-7>
- [2] A. G. M. Strollo, E. Napoli, D. De Caro, N. Petra, and G. D. Meo, "Comparison and extension of approximate 4-2 compressors for lowpower approximate multipliers," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 9, pp. 3021–3034, Sep. 2020.
- [3] T. Kong and S. Li, "Design and analysis of approximate 4-2 compressors for high-accuracy multipliers," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 29, no. 10, pp. 1771–1781, Oct. 2021.
- [4] 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.
- [5] F. Sabetzadeh, M. H. Moaiyeri, and M. Ahmadinejad, "A majority-based imprecise multiplier for ultra-efficient approximate image multiplication," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 11, pp. 4200–4208, Nov. 2019.
- [6] H. Pei, X. Yi, H. Zhou, and Y. He, "Design of ultra-low power consumption approximate 4-2 compressors based on the compensation characteristic," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 1, pp. 461–465, Jan. 2021.
- [7] D. Esposito, A. G. M. Strollo, E. Napoli, D. de Caro, and N. Petra, "Approximate multipliers based on new approximate compressors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 12, pp. 4169–4182, Dec. 2018.
- [8] U. Anil Kumar, S. K. Chatterjee, and S. E. Ahmed, "Lowpower compressor-based approximate multipliers with error correcting module," *IEEE Embdedd Syst. Lett.*, vol. 14, no. 2, pp. 59–62, Jun. 2022.
- [9] X. Yi, H. Pei, Z. Zhang, H. Zhou, and Y. He, "Design of an energyefficient approximate compressor for error-resilient multiplications," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2019, pp. 1–5.
- [10] M. Ha and S. Lee, "Multipliers with approximate 4-2 compressors and error recovery modules," *IEEE Embdedd Syst. Lett.*, vol. 10, no. 1, pp. 6–9, Mar. 2018.
- [11] M. Ahmadinejad, M. H. Moaiyeri, and F. Sabetzadeh, "Energy and area efficient imprecise compressors for approximate multiplication at nanoscale," *AEU Int. J. Electron. Commun.*, vol. 110, Oct. 2019, Art. no. 152859
- [12] D. Mohapatra, V. K. Chippa, A. Raghunathan, and K. Roy, "Design of voltage-scalable meta-functions for approximate computing," in *Proc. Design, Autom. Test Eur.*, Mar. 2011, pp. 1–6.
- [13] K. Yin Kyaw, W. Ling Goh, and K. Seng Yeo, "Low-power high-speed multiplier for error-tolerant application," in *Proc. IEEE Int. Conf. Electron Devices Solid-State Circuits (EDSSC)*, Dec. 2010, pp. 1–4.
- [14] M. de la Guia Solaz, W. Han, and R. Conway, "A flexible low power DSP with a programmable truncated multiplier," *IEEE Trans. Circuits Syst.*, vol. 59, no. 11, pp. 2555–2568, Nov. 2012.
- [15] R. Zendegani, M. Kamal, M. Bahadori, A. Afzali-Kusha, and M. Pedram, "RoBa multiplier: A rounding-based approximate multiplier for highspeed yet energy-efficient digital signal processing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 2, pp. 393–401, Feb. 2017.
- [16] C. S. Wallace, "A suggestion for a fast multiplier," *IEEE Trans. Electron. Comput.*, vol. EC-13, no. 1, pp. 14–17, Feb. 1964.
- [17] A. Weinberger, "4:2 carry-save adder module," *IBM Tech. Discl. Bull.*, vol. 23, no. 8, pp. 3811–3814, 1981.
- [18] 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.

- [19] Z. Yang, J. Han, and F. Lombardi, "Approximate compressors for errorresilient multiplier design," in Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Nanotechnol. Syst. (DFTS), Oct. 2015, pp. 183–186.
- [20] C.-H. Lin and I.-C. Lin, "High accuracy approximate multiplier with error correction," in Proc. IEEE 31st Int. Conf. Comput. Design (ICCD), Oct. 2013, pp. 33–38.
- [21] P. J. Edavoor, S. Raveendran, and A. D. Rahulkar, "Approximate multiplier design using novel dual-stage 4:2 compressors," IEEE Access, vol. 8, pp. 48337–48351, 2020.
- [22] F. Sabetzadeh, M. H. Moaiyeri, and M. Ahmadinejad, "A majority-based imprecise multiplier for ultra-efficient approximate image multiplication," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 66, no. 11, pp. 4200–4208, Nov. 2019.
- [23] A. G. M. Strollo, E. Napoli, D. De Caro, N. Petra, and G. D. Meo, "Comparison and extension of approximate 4–2 compressors for lowpower approximate multipliers," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 67, no. 9, pp. 3021–3034, Sep. 2020.
- [24] H. Xiao, H. Xu, X. Chen, Y. Wang, and Y. Han, "Fast and high-accuracy approximate MAC unit design for CNN computing," IEEE Embedded Syst. Lett., early access, Dec. 21, 2021, doi: 10.1109/LES.2021.3137335.
- [25] 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.
- [26] T. Yang, T. Ukezono, and T. Sato, "A low-power high-speed accuracycontrollable approximate multiplier design," in Proc. 23rd Asia South Pacific Design Autom. Conf. (ASP-DAC), Jan. 2018, pp. 605–610.
- [27] I. Hammad, L. Li, K. El-Sankary, and W. M. Snelgrove, "CNN inference using a preprocessing precision controller and approximate multipliers with various precisions," IEEE Access, vol. 9, pp. 7220–7232, 2021.
- [28] C. Guo, L. Zhang, X. Zhou, W. Qian, and C. Zhuo, "A reconfigurable approximate multiplier for quantized CNN applications," in Proc. 25th Asia South Pacific Design Autom. Conf. (ASP-DAC), Jan. 2020, pp. 235–240.