VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

SCHOOL OF ELECTRICAL ENGINEERING

APRIL 2023

APPROXIMATE MULTIPLIER USING 4:2 COMPRESSOR

PROJECT REPORT

For

ADVANCED DIGITAL SYSTEM DESIGN WITH FPGAs – EEE4019

Under the Guidance of,

Prof. MARIMUTHU R.

Submitted in partial fulfilment for the award of the degree of

B. TECH

In

ELECTRICAL AND ELECTRONICS ENGINEERING

By

ANUJ MATHUR	**--**	**20BEE0278**
SIDRAMESHWER	**--**	**20BEE0289**
HARISH R	**--**	**20BEE0299**

CERTIFICATE

This certificate is to affirm that the project entitled "**Approximate Multiplier using 4:2 Compressor**" was carried out by **Anuj Mathur (20BEE0278), Sidrameshwer (20BEE0289), and Harish R (20BEE0299)** under the supervision of **Marimuthu R. in the Department of Electrical and Electronics Engineering** at **Vellore Institute of Technology, Vellore**, during the academic year 22-23. We highly recommend the approval of this project work to be considered for the partial fulfillment of the requirements for the award of the 6th semester in the Bachelor of Electrical and Electronics Engineering degree.

ACKNOWLEDGMENT

The successful completion of a project relies on the collaboration, coordination, and combined efforts of multiple sources of knowledge. This report expresses gratitude towards several individuals who provided guidance, supervision, motivation, and inspiration. Firstly, we express our gratitude to the divine power for the blessings bestowed upon us to accomplish this project work triumphantly. It is our honor to convey our utmost respect to our project mentors, **Prof. Marimuthu.R,** for his valuable insights, competent guidance, motivation, wholehearted support, and constructive feedback during the entire duration of our project.

ABSTRACT:

This paper presents an approximate multiplier using a 4:2 compressor that provides an approximate result instead of an exact one. The circuit is designed to reduce power consumption and improve performance by utilizing simpler circuits and fewer operations. Several techniques are employed to implement approximate multiplication, such as truncation, quantization, and stochastic rounding. The use of an approximate multiplier offers several benefits, including reduced power consumption, increased speed, and cost-effectiveness. However, the accuracy of the approximate multiplier depends on the level of error introduced in the multiplication result, which is specific to the implementation. We will be implementing the desired circuit using **VERILOG HDL in XILINX VIVADO 2018.3.**

INTRODUCTION:

Multiplication is a critical arithmetic operation in digital circuits, but its hardware implementation can be expensive in terms of power consumption and performance. An alternative solution to traditional multipliers is the use of an approximate multiplier, which can provide an approximate result instead of an exact one. This approach aims to reduce power consumption and improve performance by using simpler circuits and fewer operations. Techniques used to implement approximate multiplication include truncation, quantization, and stochastic rounding. In this paper, we propose an 8-bit unsigned multiplier using a 4:2 compressor that approximates the least significant region partial products using four different methods and reduces the approximate region partial products using proposed approximate compressors. An error-correcting module compensates for the error introduced by the proposed compressor. The partial products in the accurate region are reduced using exact full adders and 4:2 compressors.

This approach exploits the redundancy in the partial products generated during multiplication and uses a compressor circuit to reduce the number of adders required. The 4:2 compressor is a simple and efficient circuit that reduces two 4-bit partial products to a 3-bit sum and a 1-bit carry. The use of a 4:2 compressor in a multiplier results in a reduction in power consumption and hardware resources while maintaining the accuracy of the multiplication operation. This approach has applications in digital signal processing, computer arithmetic, and other areas where multiplication is a critical operation.

METHODOLOGY

Stage division is an essential technique used in the design of multipliers using a 4:2 compressor. The multiplier is divided into multiple stages, each of which performs a specific task. In the proposed 8-bit unsigned multiplier, the multiplication operation is divided into three stages:

1. **The Least significant region (LSR),**
2. **The Approximate region (AR), and**
3. **The Accurate region (AC).**

The LSR is the least significant part of the multiplier, and the partial products in this region contribute the least to the final product. Therefore, these partial products are approximated using one of four different methods. The AR is the middle part of the multiplier, and the partial products in this region are reduced using proposed approximate compressors. An error-correcting module is used to compensate for the error introduced by the proposed compressor. Finally, the AC is the most significant part of the multiplier, and the partial products in this region are reduced using exact full adders and 4:2 compressors. The use of stage division allows for better control and optimization of the hardware resources and power consumption of the multiplier while maintaining its accuracy.

1. **Least significant region (LSR)**

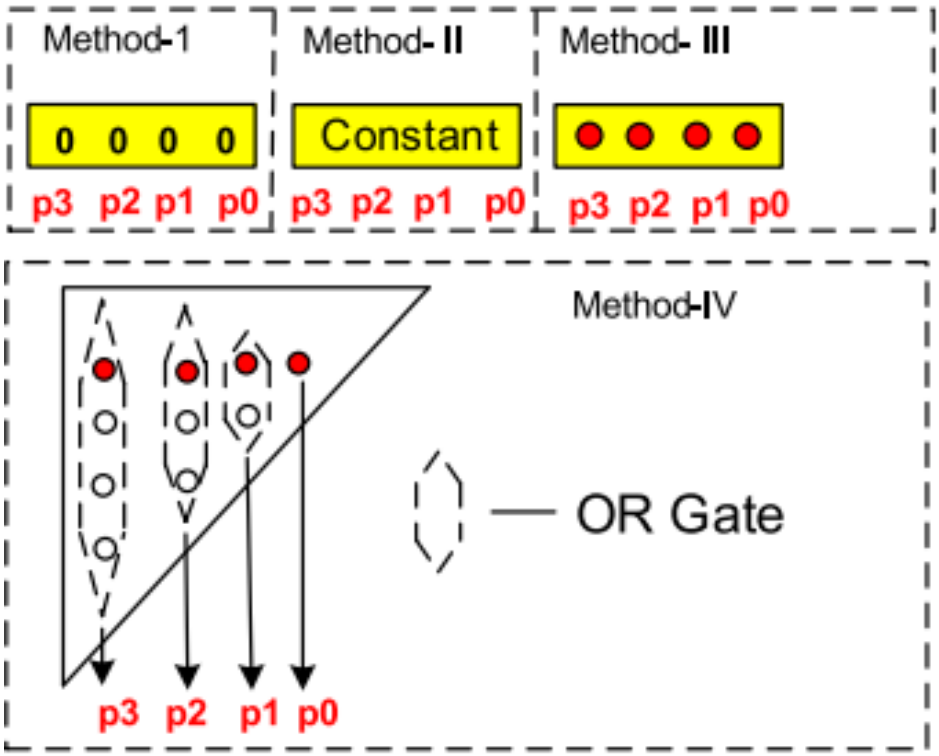
   As this region contributes the least to the final product, Different methods are proposed to approximate the partial product:

   i. **M1-Truncation**: The process of dropping the least significant bits of a binary number and replacing them with zeroes. This method saves area but introduces errors.

   ii. **M2- Constant Correction Term**: The LSP value is assigned to a constant value calculated by averaging the exact value obtained for all the possible input combinations.

   iii. **M3-Bypass the first row**: The LSR value is directly assigned to one partial product term. Method 2 and Method 3 reduce error without incurring additional hardware cost.

   iv. **M4-OR operation**: The LSR value is calculated by using OR gates, which reduces the error significantly however increases the hardware utilization.



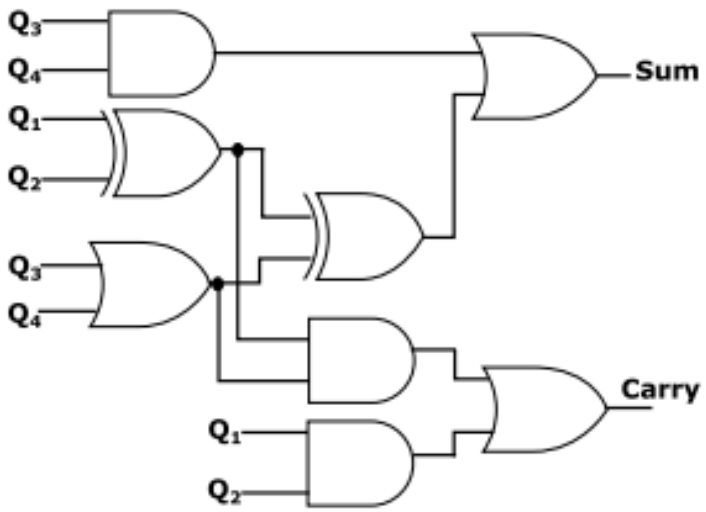
4-Bit Least Significant Region

2. The Approximate region (AR)

The 4-bit approximate region includes the most significant bits of the multiplier, which contribute more significantly to the final product. The partial products in this region are reduced using the proposed approximate compressors, which are more efficient than traditional adders. However, the use of these compressors introduces some error in the result, which must be compensated for.

Approximate Compressor 1: Approximate compressor 1 (AC1) has four inputs and two outputs. It has been observed that AC1 it produces an error for two cases only.

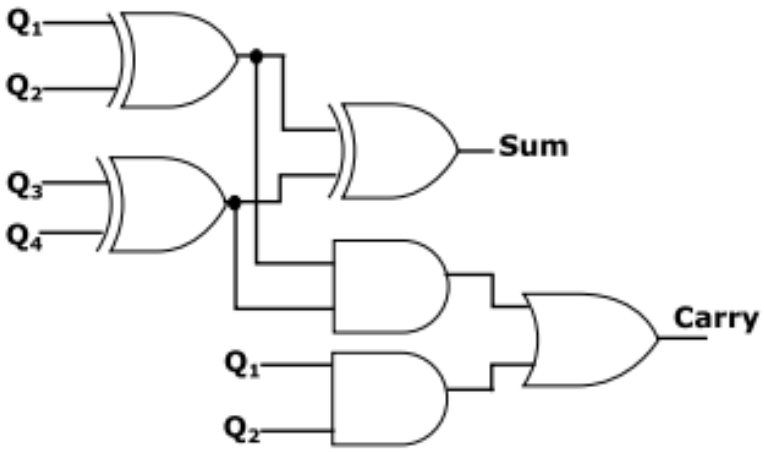
The PPs in the accurate region are reduced using exact full adders and 4:2 compressors, which maintain the accuracy of the multiplication operation. Overall, the division of the multiplier into these three regions and the use of different techniques to approximate and reduce partial products in each region is an effective approach to improve the performance and reduce the hardware resources required for multiplication.



Approximate compressor 1

- **Boolean Expression:** $Sum = p1 \oplus p2 + Q3Q4$
  $Carry = p1\ p2 + Q1Q2$
- **Where** $p1 = Q1 \oplus Q2$
  $p2 = Q3 + Q4$

Approximate compressor 2 (AC2) is used to correct errors caused by approximate compressor 1 in the most significant column.



Approximate compressor 2

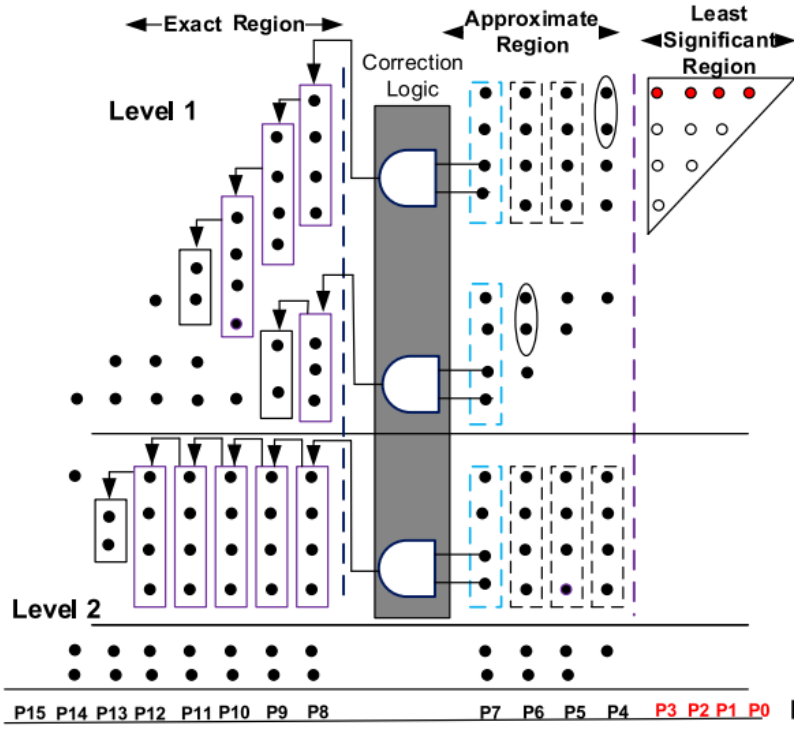
An additional approximate compressor is designed to replace the MSC and can be easily converted to an exact compressor using an error-correcting module.

- **Boolean Equations:** $Sum = p2 \oplus p1$

$$Carry = Q1Q2 + p2p1$$

- **Where** $p1 = Q3 \oplus Q4$ **and** $p2 = Q1 \oplus Q2$

An error-correcting module is included in the multiplier design to address the '-2' error generated by AC2 in 4 out of 16 cases in the MSC of the approximate region. This module identifies whether there is an error that needs to be compensated and generates a '1' to signal the error or '0' otherwise. The '-2' error in the cases '0011', '0111', '1011', and '1111' is compensated by adding a carry into the next position, resulting in an effective addition of +2 to negate the error completely. After the error correction, AC2 functions like an exact compressor with lower hardware complexity.



4 Bit Approximate Region

Level 1 of the error-correcting module includes two AND gates that analyze the PPs at the MSC of the approximate region. The outputs of these gates serve as carry-ins to the exact compressors in the least significant column of the exact region. This reduces the error distance and enhances the overall accuracy of the multiplier. Similarly, in level 2, an AND gate is used at the MSC of the approximate region to further improve the accuracy of the multiplier.

RESULTS

Variance according to different LSB designs of multiplier:

By varying the techniques used in the LSR (approximate and accurate region remains same), different variants of 8-bit and 16-bit multiplier architectures are proposed, namely D1, D2, D3, and D4, and are tabulated in Table 3. In D1, when the LSR is truncated it saves area but introduces error. In contrast, in D2, the LSP value is assigned to a constant value calculated by averaging the exact value obtained for all the possible input combinations. In D3, the LSR value is directly assigned one partial product term. Now we can see that in the case of D2 and D3, replacing the LSR portion with a constant term or partial product term reduces the error without incurring additional hardware cost. Finally, in D4, the LSR value is calculated by using OR gates, which reduces the error significantly however increases the hardware utilization. This is further evaluated using a numerical example.

Consider A= (10100111)2 and B=(11101101)2, the exact output is shown in Fig. 3(a). The impact of the four approximation methods on the output can be observed in Fig. 3(b–e). In design D1, lower four bits are truncated, and the remaining approximate and exact portions are reduced using proposed and exact compressors, respectively. The resulting output is (39552)10, which is close to the exact output. In design D2, lower four bits are replaced with a constant term, and the resultant output is (39558)10. Replacing the lower four bits directly with the partial products of the first row in D3, the output is (39559)10, which is closer to the exact output than D1 and D2 outputs. In design D4, the output obtained by reducing the PP bits using OR gates is $(39567)_{10}$, which is the most accurate of all the four outputs. Fig. 3(f) shows the design D1 without error correcting module, and the output obtained here is (39296)10. Comparing the two outputs from Fig. 3(b) and 3(f), it is clear that the error correcting module is indispensable for our proposed designs.

Summary:

1. *Four distinct techniques (D1, D2, D3 and D4) are proposed to lower the area and power with a compromise on accuracy in the least significant part of the PPR structure.*
2. *To simply the hardware complexity, two new approximate 4:2 compressors with high accuracy and low power consumption are deployed in the approximate portion of the PPR structure.*
3. *The approximate compressor is designed to perform like an exact compressor using a simple but robust error-correcting module.*

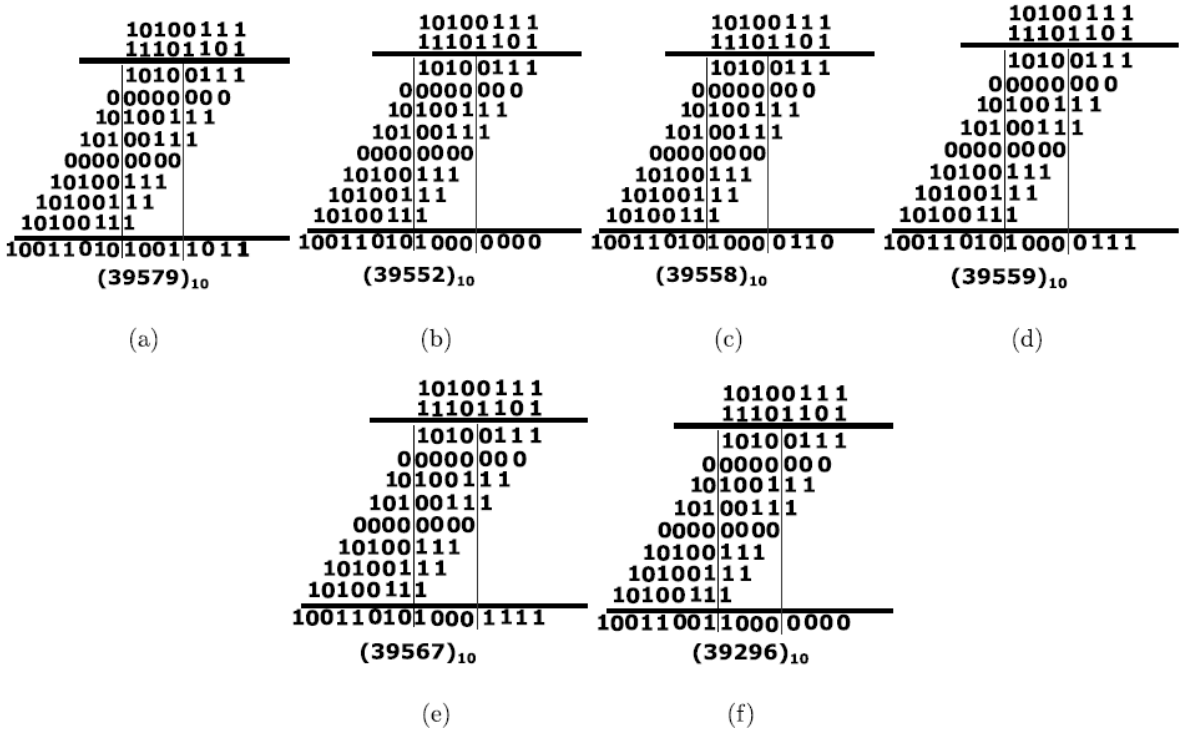
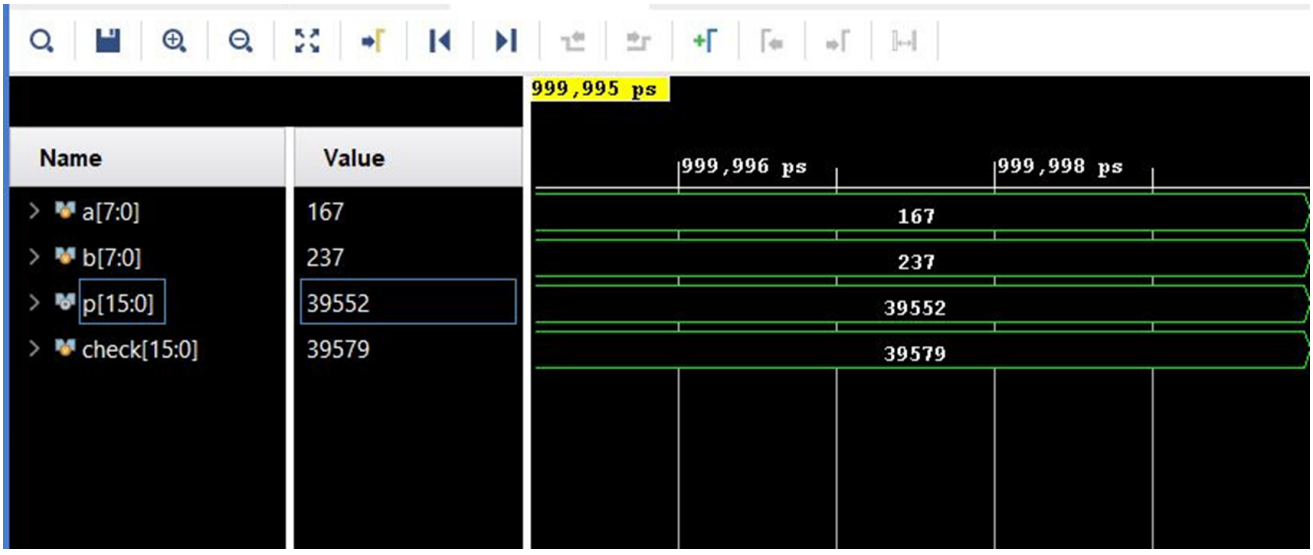


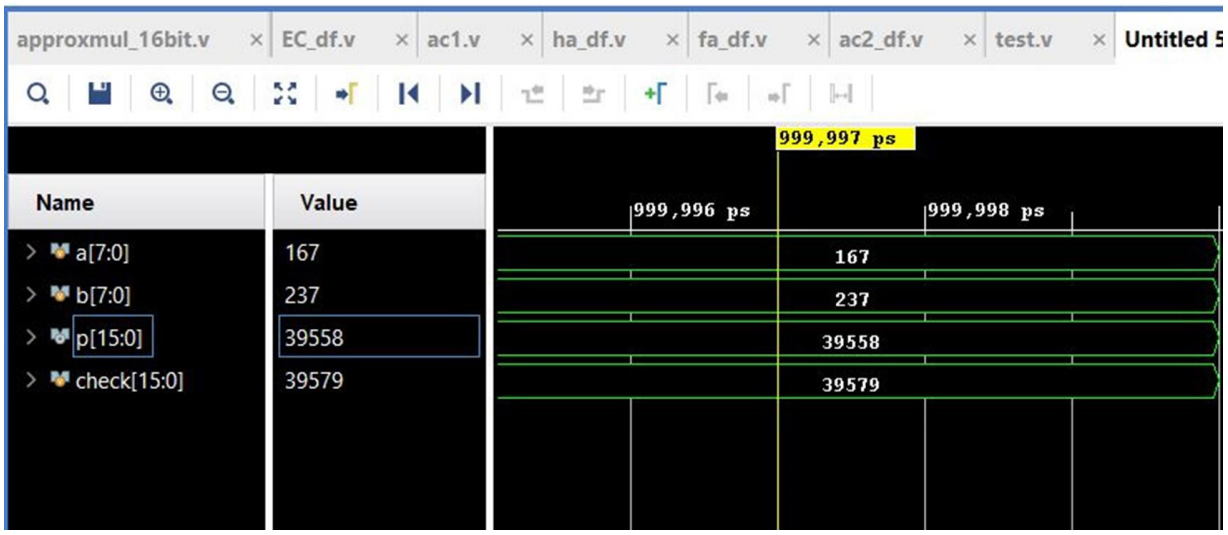
fig. 3

Synthesis and Simulation (Verilog):

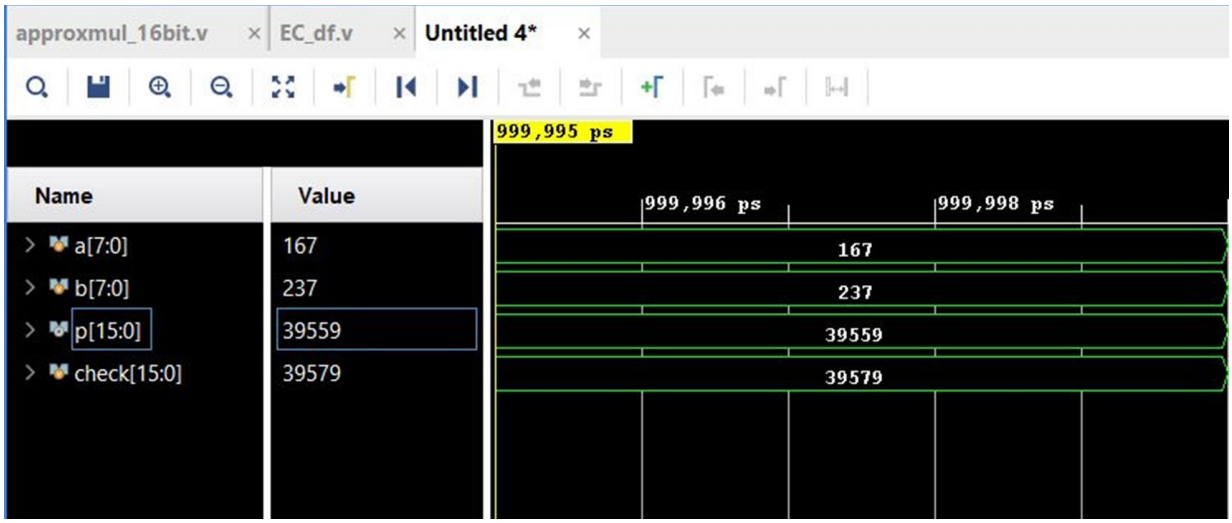
1. Method 1:



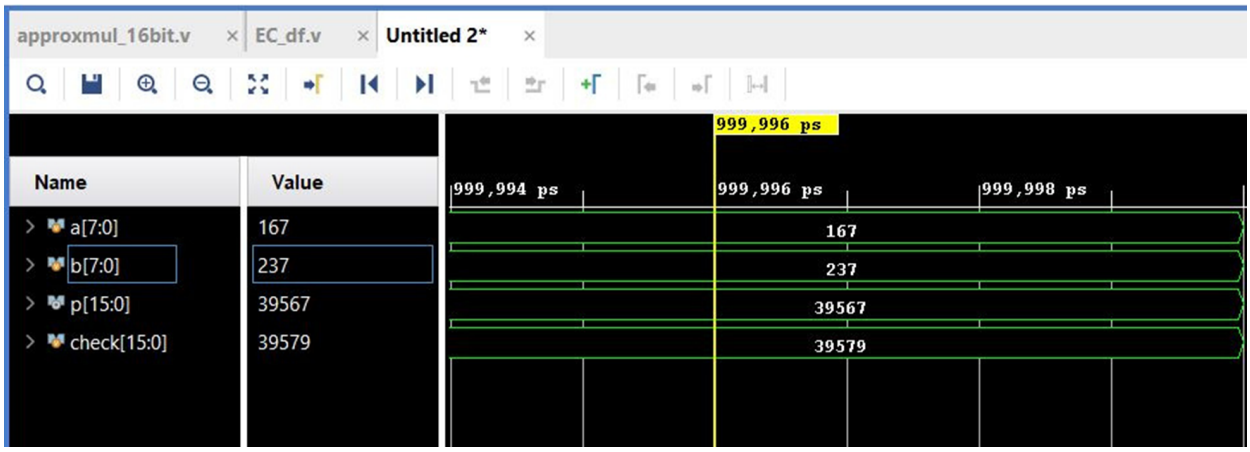
2. Method 2:



3. Method 3:



4. Method 4:



CONCLUSION

Error Analysis & Discussion

Exhaustive error analysis is performed on the existing and proposed 8-bit and 16-bit multiplier architectures using 65,536 and 1 million (random) input samples, respectively, in MATLAB. The results are tabulated in Table 4. Metrics such as max error distance

1. **(EDmax),** normalized mean error distance (NMED), worst case relative error distance (WCRE), and mean relative error distance.
2. **(MRED)** are used to quantify the efficacy of multiplier designs. The error distance represents the difference between accurate and approximate results. MRED indicates the mean of all REDs where RED is relative error distance.

It can be observed from Table 4 that D1, D2, D3 and D4 have lower $EDmax$, NMED, and MRED compared to all existing multiplier designs except Ax8-1 and AxRM1 due to efficient proposed approximate compressors, error-correcting module, and techniques used in the least significant portion of the PPR structure. Similarly, the NMED, MRED, and WCRE values are lower for 16-bit designs than existing ones.

Table 4
Error Analysis of various 8-bit and 16-bit approximate multipliers.

Multiplier	8-bit			16-bit		
	MRED (%)	NMED (10–3)	EDmax	MRED (%)	NMED (10–3)	WCRE
D1	0.68	0.28	213	0.01	0.0032	0.42
D2	0.49	0.23	207	0.01	0.0032	0.40
D3	0.48	0.229	202	0.01	0.0029	0.31
D4	0.21	0.14	198	0.009	0.0031	0.21
Ax8-1 [9]	0.1	0.46	96	0.007	0.001	0.61
Ax8-2 [9]	1.26	1.64	2034	0.09	0.04	0.97
Ax8-3 [9]	2.83	6.12	3954	0.18	0.08	0.98
Multiplier2 [15]	1.36	1.56	1020	0.12	0.44	0.99
Yang [12]	0.77	0.49	641	0.02	0.057	0.65
Minho [13]	0.78	0.43	385	0.01	0.05	0.31
Xilin [14]	1.40	1.33	560	0.05	0.18	0.42
MUL2 [16]	2.43	1.06	264	0.09	0.021	0.99
M3 [8]	1.7	2.1	1188	1.6	1.2	0.43
M4 [8]	2.2	3.2	2244	2.2	1.9	0.44
AxRM1 [10]	0.765	0.25	50	0.01	0.001	34.1
AxRM2 [10]	5.6	4.28	850	0.27	0.025	191.5
AxRM3 [10]	7.566	5.22	1650	0.49	0.03	885.1
C-N [17]	0.10	0.13	520	0.004	0.002	0.19

This work presented an unsigned multiplier using approximate computing techniques to reduce computing complexity by approximating the middle and least significant portions of the PPR structure. Two approximate compressors and error-correcting modules were proposed to reduce partial products and improve accuracy. The proposed designs were implemented and demonstrated power improvements of 26.5% compared to exact designs and 18.5% compared to existing designs. Furthermore, the proposed designs showed superior computation quality tradeoff when validated using image processing and CNN applications. Future research can investigate utilizing several approximate compressors for each column in Level 2 and reconfigurable platform operations.

REFERENCES

[1] Liu Weiqiang, Lombardi Fabrizio, Shulte Michael. A retrospective and prospective view of approximate computing [point of view]. Proc IEEE 2020;108(3):394–9.

[2] Afzali-Kusha Hassan, Vaeztourshizi Marzieh, Kamal Mehdi, Pedram Massoud. Design exploration of energy-efficient accuracy-configurable dadda multipliers with improved lifetime based on voltage overscaling. IEEE Trans Very Large Scale Integr (VLSI) Syst 2020;28(5):1207–20.

[3] Ansari Mohammad Saeed, Mrazek Vojtech, Cockburn Bruce F, Sekanina Lukas, Vasicek Zdenek, Han Jie. Improving the accuracy and hardware efficiency of neural networks using approximate multipliers. IEEE Trans Very Large Scale Integr (VLSI) Syst 2019;28(2):317–28.

[4] Mrazek Vojtech, Vasicek Zdenek, Sekanina Lukas, Jiang Honglan, Han Jie. Scalable construction of approximate multipliers with formally guaranteed worst case error. IEEE Trans Very Large Scale Integr (VLSI) Syst 2018;26(11):2572–6.

[5] Leon Vasileios, Zervakis Georgios, Soudris Dimitrios, Pekmestzi Kiamal. Approximate hybrid high radix encoding for energy-efficient inexact multipliers. IEEE Trans Very Large Scale Integr (VLSI) Syst 2017;26(3):421–30.

[6] Kulkarni Parag, Gupta Puneet, Ercegovac Milos. Trading accuracy for power with an underdesigned multiplier architecture. In: 2011 24th Internatioal Conference on VLSI Design. IEEE; 2011, p. 346–51.

[7] Jiang Honglan, Liu Cong, Liu Leibo, Lombardi Fabrizio, Han Jie. A review, classification, and comparative evaluation of approximate arithmetic circuits. ACM J Emerg Technol Comput Syst (JETC) 2017;13(4):1–34.

[8] Ansari Mohammad Saeed, Jiang Honglan, Cockburn Bruce F, Han Jie. Low-power approximate multipliers using encoded partial products and approximate compressors. IEEE J Emerg Sel Top Circuits Syst 2018;8(3):404–16.

[9] Waris Haroon, Wang Chenghua, Liu Weiqiang, Han Jie, Lombardi Fabrizio. Hybrid partial product-based high-performance approximate recursive multipliers. IEEE Trans Emerg Top Comput 2020.

[10] Waris Haroon, Wang Chenghua, Xu Chenyu, Liu Weiqiang. AxRMs: Approximate recursive multipliers using high-performance building blocks. IEEE Trans Emerg Top Comput 2021.

[11] Momeni Amir, Han Jie, Montuschi Paolo, Lombardi Fabrizio. Design and analysis of approximate compressors for multiplication. IEEE Trans Comput 2014;64(4):984–94.

[12] Yang Zhixi, Han Jie, Lombardi Fabrizio. Approximate compressors for error-resilient multiplier design. In: 2015 IEEE international symposium on defect and fault tolerance in VLSI and nanotechnology systems. IEEE; 2015, p. 183–6.