Exploring Approximate Computing and Near-Threshold Operation to Design Energy-efficient Multipliers

Vinícius Zanandrea[1], Douglas M. Borges[2], Vagner S. Rosa[2] and Cristina Meinhardt[1,2]
[1]Departamento de Informática e Estatística, Universidade Federal de Santa Catarina - UFSC, Brazil
[2]Centro de Ciências Computacionais, Universidade Federal do Rio Grande - FURG, Brazil
vinicius.zanandrea@posgrad.ufsc.br, douglasborges19@gmail.com, vagner.rosa@furg.br, cristina.meinhardt@ufsc.br

***Abstract*—Multiplier circuits are components of particular relevance for digital systems. Hardware projects often require fast and low-power multipliers. In this regard, this work evaluates a set of multiplier circuits to explore alternative approaches for energy-efficient scenarios. This work explores two techniques for energy efficiency: reducing the operating voltage (near-threshold operation) and through Approximate Computing. Two approximate adders are adopted in the lower bits. Altogether, eight operation scenarios are considered and evaluated at the electrical level, providing an overall discussion of the most indicate approaches for different design requirements. The results show that by applying near-threshold operation, it is possible to achieve a considerable reduction in power consumption, however, with a significant increase in delay times. The replacement of exact Mirror adders by approximate AMA2 provided a reduction of up to 29.6% in energy consumption and up to 4% in delay for the evaluated multiplier circuits.**

***Index Terms*—approximate computing, multipliers, energy efficiency**

I. Introduction

Power consumption has become a critical constraint in digital systems development due to the recent growth in the use of personal devices (portable computers and real-time audio and video-based multimedia applications). In general, higher power dissipation implies high-temperature operation, leading to possible system failures. Multiplier circuits have a significant influence on the performance and power characteristics of the system where they are inserted [1]. In this regard, choosing a specific multiplier topology can bring gain or loss of performance to digital and embedded systems [2], especially in Digital Signal Processor (DSP), image processing, and arithmetic units in microprocessors [1].

The design of multiplier circuits aiming at energy efficiency has traditionally explored two approaches: reducing the operating voltage or exploring new topologies. However, recently, Approximate Computing (AC) technique has been explored in hardware and software development for different contexts in error-tolerant applications. Error tolerant environments include a wide range of applications, such as video and sound processing, Internet of Things (IoT), and computer vision [3]. The central part of these applications has a common point: to be predominantly composed of arithmetic modules composed by adders and multipliers [4] [5]. Approximate Computing exchanges a reduction in the precision of the results generated in a computer system for a reduction in the necessary resources, such as energy or area [6].

There are few recent works in literature exploring approximation on multipliers. These works can be divided into two categories: exploring approximate compressors or exploring approximate adders. Approximate compressors minimize the error probability and the average error [7] [8]. In [7], the approximate compressors are implemented by using AND-OR gates and adopted to build approximate multipliers. The circuits have been synthesized using a 40nm CMOS technology. In [8] a 15-4 compressor is proposed for a 16×16 bit multiplier. Approximate adders are explored in the lower bits of 16×16 multipliers in [9]. Simulation results show that area, delay, and power are significantly improved by the proposed multiplier compared to the conventional binary multiplier. All these studies are proposed and evaluated at the register-transfer-level. To the best of our knowledge, no work investigates multipliers at the electrical level. The evaluation at the electrical level allows the investigation in detail the energy and delay characteristics. However, electrical simulations are very time-consuming, restricting the bit-width of the multipliers on evaluation.

The main goal of this work is to present the design aspects of multiplier circuits by using techniques to increase energy efficiency. In particular, two techniques are explored: reducing the operating voltage (near-threshold operation) and through AC at a logical level by adopting approximate adders in the lower bits. The evaluation is conducted at the electrical level, considering four 4-bits Multipliers topologies, and discussing eight operation scenarios, from the traditional nominal operation to mixed near-threshold and AC alternatives. The overall evaluation indicates design approaches to be explored on different design constraints.

Next section presents an overview of multiplier circuits and approximate adders. Section III describes the methodology followed to define the approximate multipliers and the near-threshold operation. The results are discussed in Section IV and Section V. Section VI presents an overall evaluation. Finally, Section VII concludes this work.

978-1-6654-2170-6/21/$31.00 ©2021 IEEE

II. Multiplier Circuits and Approximate Adders

The operation of multiplier architectures can be divided into three stages: partial products generation stage, partial products addition stage, and the final addition stage. The speed of multiplication operation can be increased by decreasing the number of partial products [10].

This work has considered four multiplier architectures: Array [11], Booth [12], Baugh-Wooley [13], and Vedic [14], which are presented in Fig. 3. The Array multiplier is an architecture based on addition and shifting procedures. Partial products are generated by AND gates, and the addition of the products is done using half adders and full adders. A 4-bit Array multiplier is shown in Fig. 3(a). This topology is composed by 392 transistors.

The Booth multiplier is an algorithm to perform 2's complement numbers multiplication. The circuit architecture is illustrated in Fig. 3(b). It is composed of 808 transistors. Booth algorithm requires examination of the multiplier bits, starting from the least significant bit. This topology consists of two types of cells: CAS (Controlled Add/Subtract) and CNTL (Control), which is responsible for determining which operation (addition, subtraction, or shifting) is done in CAS.

Baugh-Wooley circuit is shown in Fig. 3(c). This algorithm was developed in order to multiply 2's complement numbers. Each of the multiplier cells receives four inputs: the multiplier input (horizontal blue line), multiplicand input (vertical red line), carry from previous cells (vertical black line), and sum from previous cells (diagonal black line). They produce two outputs: sum output (diagonal black line) and carry output (vertical black line). This 4-bit architecture is composed of 644 transistors.

The Vedic multiplier follows the principle of Vedic Mathematics, which is based on sixteen Sutras. The Urdhva Tiryakbhyam (Vertically Crosswise) Sutra is the method applied for multiplications. A 4-bit Vedic multiplier is presented in Fig. 3(d). This topology is composed by 690 transistors for Vedic Carry-lookahead Adder (CLA) and 506 transistors for Vedic Ripple-Carry Adder (RCA).

In multipliers, the Full Adder is responsible for adding partial products in order to obtain the final product of multiplication. This circuit has a fundamental role in the operation of any computer system, being the main cell of the Arithmetic Logic Unit [15]. Approximate adders based on XOR/XNOR gates and pass transistors are proposed in [16]. These circuits have a reduced number of transistors when compared to the exact mirror adder CMOS.

In this work, the approximation was applied by replacing the exact adders inserted in the multiplier circuits by approximate adders. In this regard, the topologies chosen are: Approximate Mirror Adder (AMA2) and Approximate XNOR Adder (AXA2), which are presented in Fig. 1 and Fig. 2, respectively. The red square area in each multiplier circuit shown in Fig. 3 illustrates where the approximate adders were adopted in this study.

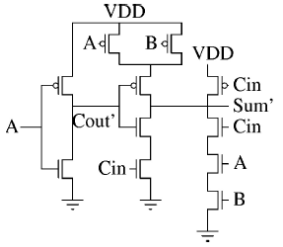


Fig. 1: Approximate Mirror Adder 2 (AMA2) [17]

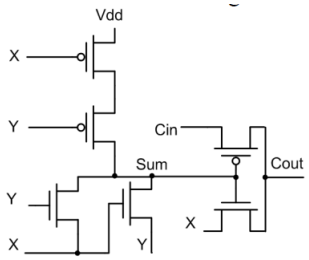


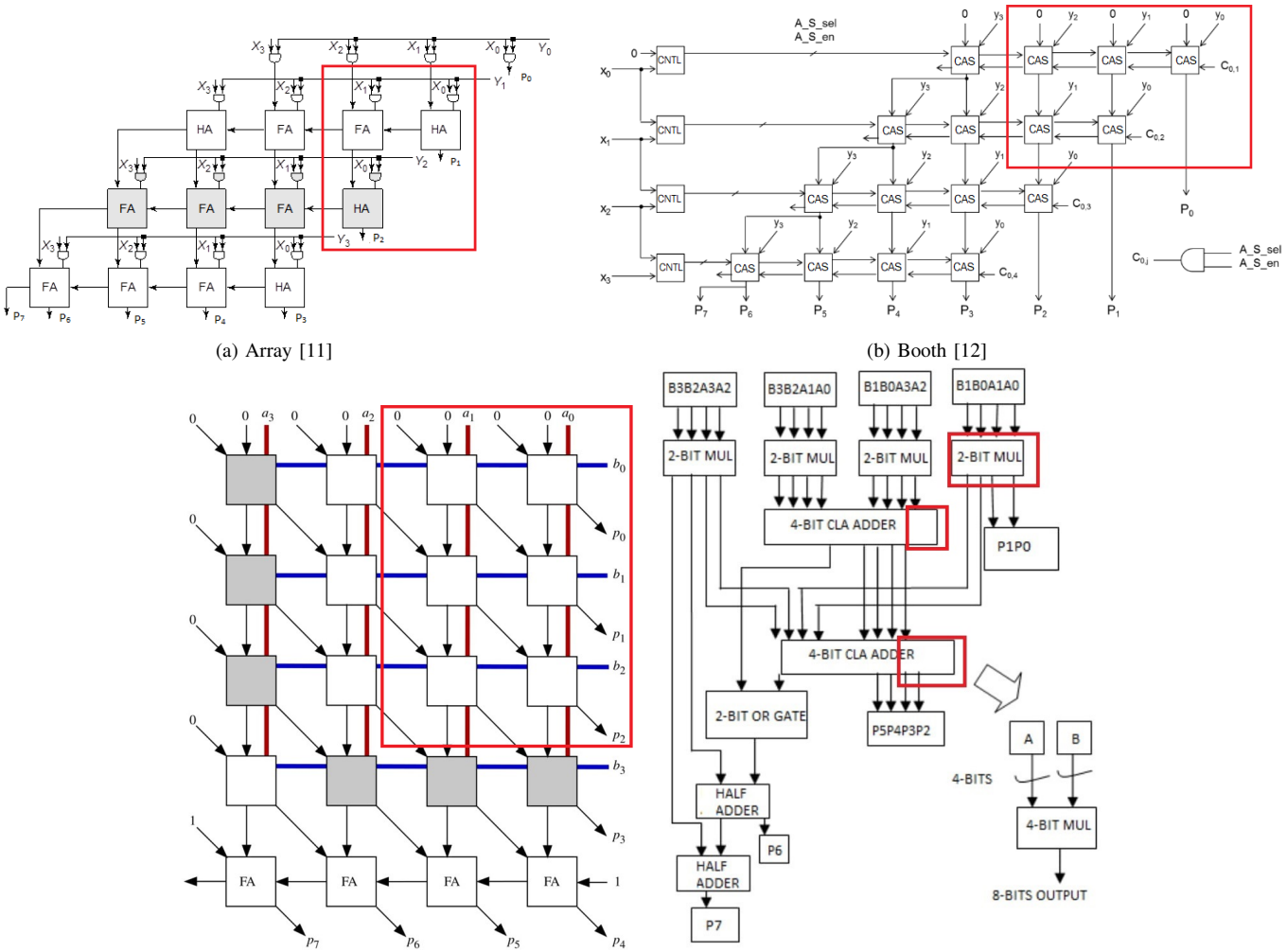
Fig. 2: Approximate XNOR-based Adder 2 (AXA2) [17]

III. Methodology

This work focuses on the power reduction provided by the use of Approximate Computing and near-threshold operation. Therefore, delay, power consumption, Power-Delay Product (PDP) and Error Distance (ED) are evaluated for four multiplier circuits and energy-efficient scenarios. The bit length of these multipliers is defined to 4-bits due to the high time demanded by the electrical simulations. In order to analyze data, the multipliers were described at the electric level and simulated using NGSPICE circuit simulator. The topologies are designed using 16nm High Performance technology provided by PTM, which is based on Bulk CMOS [18]. The high-performance model is adopted instead of the low-power model with the main objective to found a good trade-off among power, delay, and accuracy. To emulate a more realistic scenario, two inverters were used as load in each input, and a capacitor of 1 fF was connected in each output.

The delay is obtained by considering all the timing arcs exhaustively. We identified 4840 timing arcs to unsigned multipliers (Array and Vedic) and 5088 for 2's complement multipliers (Booth and Baugh-Wooley). The power consumption is obtained by the division of the energy consumed and the total period of the simulation. PDP is calculated by the product of critical delay and power consumption [19]. All the process to obtain these measures is automated by a C routine. In order to verify the quality of the outcomes, an implementation that considers all possible multiplications between 4 bits is processed, and we observe the Error Distance. ED is defined as the arithmetic distance between an erroneous output and the expected correct output for a given input [20].

In order to analyze the performance of the multiplier circuits, these eight test cases were created:



(a) Array [11]

(b) Booth [12]

(c) Baugh-Wooley [13]

(d) Vedic [14]

Fig. 3: Multiplier circuits and the approximated regions.

1) Nominal: the nominal voltage (0.7 V) was used.
2) Near-Threshold: the multipliers were simulated using a near-threshold supply voltage of 0.4 V.
3) 100% AXA2: all adders were replaced by AXA2, and nominal voltage was used.
4) 30% AXA2: 30% of the exact adders were replaced by AXA2, and nominal voltage was used.
5) 30% AXA2 in near-threshold: 30% of the AXA2 was used, and the multipliers were simulated at 0.4 V.
6) 100% AMA2: all adders were replaced by AMA2, and nominal voltage was used.
7) 30% AMA2: 30% of the exact adders were replaced by AMA2, and nominal voltage was used.
8) 30% AMA2 in near-threshold: 30% of the AMA2 was used, and the multipliers were simulated at 0.4 V.

IV. Results

This section presents the results according to the simulated scenarios. First, is presented the delay, power, and PDP results for the nominal scenario. Then, the delay and power values of the remaining scenarios are normalized by the values obtained in the nominal approach.

A. Nominal

Table I shows the delay, power, and PDP results of the multiplier circuits at nominal voltage (0.7 V). The Vedic RCA presented the lowest critical delay, being 14.46% and 13.96% lesser than Array and the Vedic CLA, respectively. The Baugh-Wooley presented the lowest critical delay among the 2's complement multipliers, being 3.2% lesser than Booth. The Array multiplier presented the best value for maximum power, being 84.11% lesser than Vedic CLA. Among the 2's complement multipliers, the Baugh-Wooley shows up to 33% of power reduction compared to Booth. As can be observed, the best PDP value was found in Vedic RCA, being 81.86% lesser than the Vedic CLA.

B. Near-threshold

The delay of the multipliers in near-threshold normalized by the nominal scenario is shown in Fig. 4. By analyzing

TABLE I: Delay, power and PDP in nominal scenario

Multiplier Circuits	Delay (ps)	Power (μW)	PDP (fJ)
Array	324.54	**5.77**	1.65
Booth	388.00	10.51	3.37
Baugh-Wooley	375.57	7.00	2.17
Vedic CLA	322.64	36.33	8.49
Vedic RCA	**277.60**	6.28	**1.54**

the results, we note that the supply voltage reduction had a considerable impact on the critical delay of the circuits, bringing increases over eight times. The worst-case was found in the Vedic CLA, being approximately 11 times slower when compared to the nominal scenario.

Fig. 5 shows the power results of the multipliers in near-threshold normalized by the nominal approach. At the voltage of 0.4 V, a reduction of 90% on average is observed. The Vedic CLA multiplier presented the best value for power, achieving reductions of up to 94%. Among the 2's complement multipliers, the Booth presented a power improvement of approximately 91% in relation to the nominal scenario.

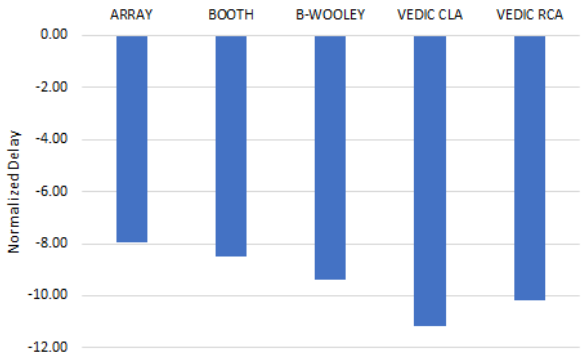


Fig. 4: Normalized delay in NT compared to the nominal scenario

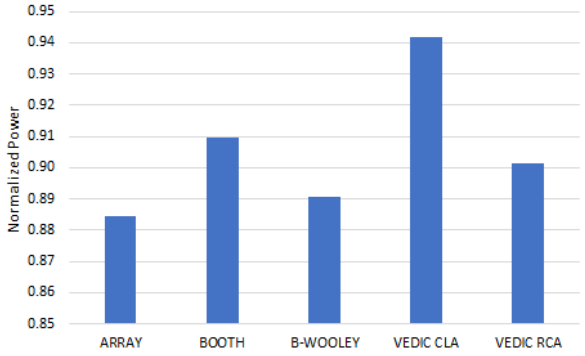


Fig. 5: Normalized power in NT compared to the nominal scenario

C. AXA2 Scenarios

The Fig. 6 illustrates the delay of the multipliers in AXA2 scenarios normalized by the nominal approach. As can be observed, the critical delay was significant increased in the test cases. In particular, the 30% AXA2 in near-threshold operation presented the worst case, being 13 times slower than the nominal scenario for the Vedic RCA multiplier.

The power results in AXA2 scenarios normalized by the nominal approach are shown in Fig. 7. The 100% AXA2 test case shows that only the 2's complement multipliers have presented power reduction, with improvements of up to 63% and 51% for Booth and Baugh-Wooley, respectively. Observing the 30% AXA2, none of the analyzed circuits showed gains. The 30% AXA2 in near-threshold operation presented on average 91% of power reduction.

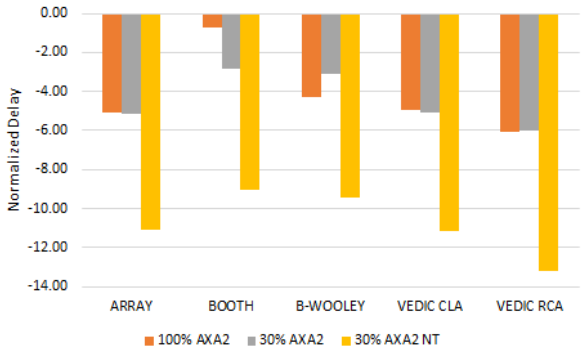


Fig. 6: Normalized delay in AXA2 scenarios compared to the nominal

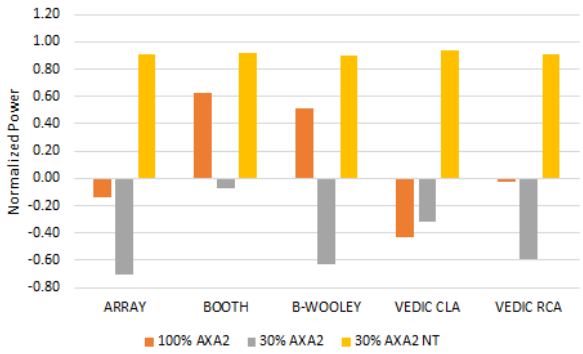


Fig. 7: Normalized power in AXA2 scenarios compared to the nominal

D. AMA2 Scenarios

The Fig. 8 shows the delay for the AMA2 scenarios normalized by the nominal approach. The 100% AMA2 scenario shows up to 64.52% and 77.67% of improvement in delay times for Array and Booth, respectively. The 30% AMA2 scenario, in which 30% of the exact adders were replaced by AMA2 and nominal voltage was used, shows up to approximately 4% of improvements for the evaluated multipliers. The critical delay of the multipliers in 30% AMA2 in NT test case was significantly impacted, with Vedic CLA being up to 11 times slower when compared to its nominal version.

Fig. 9 illustrates the power results of the multipliers in AMA2 test cases normalized by the nominal scenario. In 100% AMA2 approach, the Booth multiplier showed 67% of power-saving. By analyzing the 30% AMA2 test case, the multipliers

have presented on average 12% of power reduction, with the best value found in Booth, having 29.6% of improvements. The 30% AMA2 NT scenario shows gains for each of the evaluated circuits, with an average reduction of 92%.

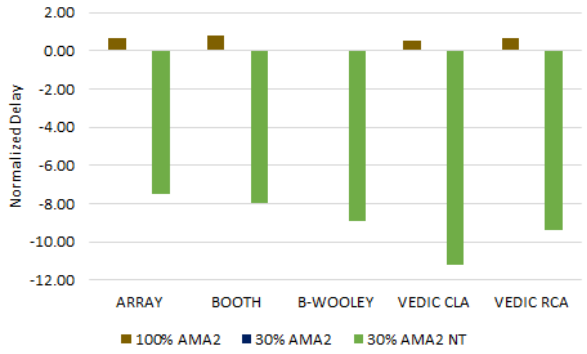


Fig. 8: Normalized delay in AMA2 scenarios compared to the nominal

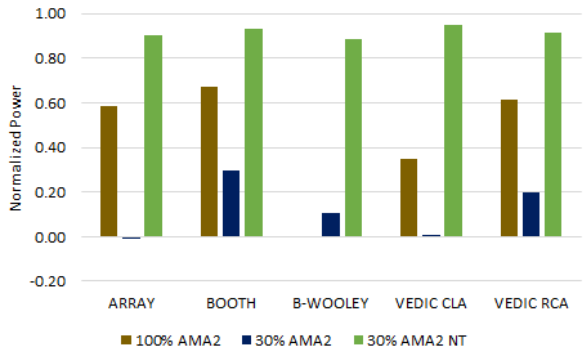


Fig. 9: Normalized power in AMA2 scenarios compared to the nominal

V. Approximation Impact on Quality

Despite the reduction in power and delay, the adoption of the approximate adders impacts the outputs inserting some errors. These errors may influence the quality (accuracy) of the solution depending on the application. This work observes the total Error Distance metric and the accumulated error by output bit to estimate this impact. Tables II and III present the Error Distance of the multipliers in AXA2 and AMA2 scenarios, respectively. We emphasise that the accuracy is not affected by the voltage operation of near-threshold.

In 100% AXA2, the Booth circuit presented the lowest total and mean values of ED. The dispersion of the results was similar for all the evaluated topologies, with standard deviation (SD) ranging from 26 to 29. The 30% AXA2 scenario showed a reduction of up to 94% for the Array and up to 82% for the Baugh-Wooley in total ED compared to the 100% AXA2.

In 100% AMA2, the 2's complement multipliers presented the lowest total, mean and SD values of ED. In 30% AMA2 scenario, it was possible to achieved, on average, a reduction of 82.4% in total ED when compared to the 100% AMA2 approach.

The total Error Distance for each output bit of the multipliers in 30% AXA2 and 30% AMA2 scenarios is shown in Fig. 10 and Fig. 11, respectively. In general, we observe that AC in the least significant bits of the multipliers produces a considerable reduction on error for P4 to P7 output bits compared to the full approximate multiplier. As can be observed, the total ED for P3, P2, and P1 is lower in AMA2 test case when compared to the AXA2 approach, especially for Vedic CLA and Vedic RCA, except for the Baugh-Wooley multiplier.

TABLE II: Error Distance of the multipliers in AXA2

Multipliers	100% AXA2			30% AXA2		
	Total	Mean	SD	Total	Mean	SD
Array	15744	61.50	26.19	912	3.56	2.88
Booth	6094	23.80	28.67	2400	9.38	20.97
Baugh-Wooley	10848	42.38	26.39	1988	7.77	4.207
Vedic CLA	20032	78.25	29.29	2688	10.50	4.213
Vedic RCA	20212	78.95	28.21	2688	10.50	4.213

TABLE III: Error Distance of the multipliers in AMA2

Multipliers	100% AMA2			30% AMA2		
	Total	Mean	SD	Total	Mean	SD
Array	6160	24.06	23.31	544	2.13	2.00
Booth	4152	16.22	14.60	1088	4.25	9.88
Baugh-Wooley	4096	16.00	14.36	1380	5.39	4.09
Vedic CLA	13980	54.61	52.32	1488	5.81	4.66
Vedic RCA	13436	52.48	51.77	1104	4.31	3.45

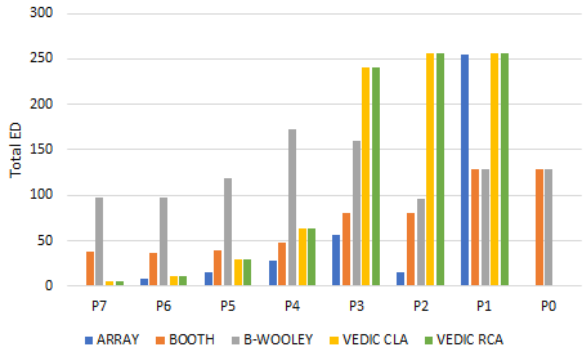


Fig. 10: Total ED per bit in 30% AXA2

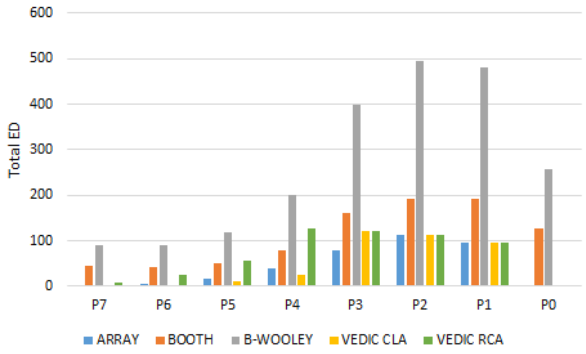


Fig. 11: Total ED per bit in 30% AMA2

VI. Overall Evaluation

Finally, to provide an overall evaluation of the scenarios and the main design constraints, the Fig. 12 illustrates the relation between power, delay, and accuracy for each of the test cases created. The near-threshold operation provides accuracy and power improvements of up to 90%. However, the critical delay had a severe increase. In AXA2 scenarios, it was possible to achieve a significant reduction in power consumption. Nevertheless, delay times were up to 13 times slower than the nominal scenario.

Furthermore, we observed the inaccuracy in the outcomes, especially in 100% AXA2 test case. The 30% AMA2 scenario showed up to approximately 4% of improvements in delay times and 29.6% of reduction in power consumption. On average, we observed an improvement of up to 82.4% in accuracy compared to the 100% AMA2 scenario.

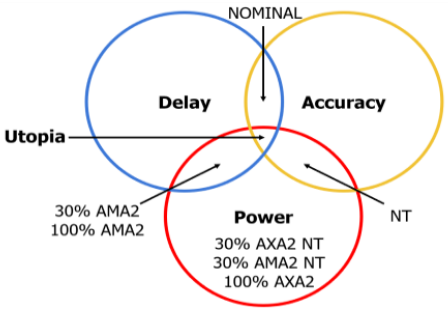


Fig. 12: Overall Scenario Evaluation

VII. Conclusion

The present work provides alternatives for the design of energy-efficient multipliers by exploring two approaches: reducing the operating voltage (near-threshold operation) and approximation by adopting Approximate Computing techniques. Simulations were performed in eight test scenarios. In the first two, the exact multipliers were simulated at nominal voltage and at near-threshold. The other scenarios addressed replacing the exact mirror adder by AXA2 and AMA2 at two levels (30% and 100%), with simulations in nominal voltage and near-threshold.

The results show that by applying near-threshold operation (0.4 V), it is possible to achieve a considerable reduction in power consumption, however, with a significant increase in delay times. In this work, we observed that the use of 30% AMA2 provided a reduction of up to 29.6% in power consumption and up to 4% in critical delay. Aiming at error-tolerant applications, it is noticeable that the adoption of AC techniques can result in scenarios with a better relation between energy, delay, and accuracy.

As future works, the following possibilities are listed: explore the use of different approximate adders in multiplier circuits; investigate the design methodology with different bit-width; evaluate the approximation impact on different applications, and explore an error correction logic for approximate multipliers.

VIII. Acknowledgment

The authors would like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and the Brazilian Council for Scientific and Technological Development (CNPq).

References

[1] Navdeep Goel and Lalit Garg. Comparative analysis of 4-bit cmos multipliers. *IJCA Proceedings on International Conference on VLSI, Communications and Instrumentation (ICVCI)*, (4):33–36, 2011.

[2] A. Khatibzadeh, K. Raahemifar, and M. Ahamdi. A novel multiplier for high-speed applications. In *Proceedings 2005 IEEE International SOC Conference*, pages 305–308, 2005.

[3] T. Moreau, A. Sampson, and L. Ceze. Approximate computing: Making mobile systems more efficient. *IEEE Pervasive Computing*, 14(2):9–13, 2015.

[4] Sparsh Mittal. A survey of techniques for approximate computing. 48(4), March 2016.

[5] M Khadra. An introduction to approximate computing. *arXiv preprint arXiv:1711.06115*, 2017.

[6] Krishna Palem and Avinash Lingamneni. Ten years of building broken chips: The physics and engineering of inexact computing. *ACM Trans. Embed. Comput. Syst.*, 12(2s), May 2013.

[7] Darjn Esposito, Antonio Giuseppe Maria Strollo, Ettore Napoli, Davide De Caro, and Nicola Petra. Approximate multipliers based on new approximate compressors. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(12):4169–4182, 2018.

[8] S. Pabithra and S. Nageswari. Analysis of approximate multiplier using 15–4 compressor for error tolerant application. In *2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCCT)*, pages 410–415, 2018.

[9] Sunghyun Kim and Youngmin Kim. High-performance and energy-efficient approximate multiplier for error-tolerant applications. In *2017 International SoC Design Conference (ISOCC)*, pages 278–279, 2017.

[10] Manjunath, V. Harikiran, K. Manikanta, S. Sivanantham, and K. Sivasankaran. Design and implementation of 16×16 modified booth multiplier. In *2015 Online International Conference on Green Engineering and Technologies (IC-GET)*, pages 1–5, 2015.

[11] Jan M. Rabaey, Anantha Chandrakasan, and Borivoje Nikolic. *Digital Integrated Circuits*. Prentice Hall Press, USA, 3rd edition, 2008.

[12] D. M. Borges, A. Borba, V. Rosa, and C. Meinhardt. Performance evaluation of arithmetic blocks at 16nm technology. In *WCAS 2019 Proceedings*, pages 1–4, 2019.

[13] Rajmohan Vijayan and uma maheswari Oorkavalan. Design of compact baugh-wooley multiplier using reversible logic. *Circuits and Systems*, 07:1522–1529, 01 2016.

[14] R. Bathija, R. Meena, S. Sarkar, and Rajesh Sahu. Low power high speed 16x16 bit multiplier using vedic mathematics. *International Journal of Computer Applications*, 59:41–44, 12 2012.

[15] A. Islam, M. W. Akram, A. Imran, and M. Hasan. Energy efficient and process tolerant full adder design in near threshold region using finfet. In *2010 International Symposium on Electronic System Design*, pages 56–60, 2010.

[16] Z. Yang, A. Jain, J. Liang, J. Han, and F. Lombardi. Approximate xor/xnor-based adders for inexact computing. In *2013 13th IEEE International Conference on Nanotechnology (IEEE-NANO 2013)*, pages 690–693, 2013.

[17] P. A. Silva and C. Meinhardt. Energy-efficient design of approximated full adders. In *2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 1–4, 2020.

[18] Y. Cao, T. Sato, M. Orshansky, D. Sylvester, and C. Hu. New paradigm of predictive mosfet and interconnect modeling for early circuit simulation. In *Proceedings of the IEEE 2000 Custom Integrated Circuits Conference (Cat. No.00CH37044)*, pages 201–204, 2000.

[19] Neil Weste and David Harris. *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley Publishing Company, USA, 4th edition, 2010.

[20] J. Liang, J. Han, and F. Lombardi. New metrics for the reliability of approximate and probabilistic adders. *IEEE Transactions on Computers*, 62(9):1760–1771, 2013.