Design Exploration of Energy-Efficient Accuracy-Configurable Dadda Multipliers With Improved Lifetime Based on Voltage Overscaling

Hassan Afzali-Kusha, Marzieh Vaeztourshizi, Mehdi Kamal, and Massoud Pedram

***Abstract*— This article investigates an energy-efficient accuracy-configurable Dadda (X-Dadda) multiplier. The structure employs the voltage overscaling and approximate width setting as the approximation knobs for improving the energy consumption as well as the reliability and lifetime of the multiplier. While the former may be set in the design time as well as the runtime, the latter may only be invoked in the design time. For a given accuracy level, the partial product columns and the overscaled voltage for optimizing the energy are determined. Normally, to have the error within a tolerable limit, the voltage overscaled columns are those at lower bit significances which have higher switching activities. The structure makes use of a low number of level shifters for a low-overhead realization. The approximate columns which start from the first column are contiguous. To further improve the efficiency of the multiplier, four-bit truncation of the multiplier output is also suggested. The efficiency of the X-Dadda structure is investigated using a 15-nm FinFET technology. The results indicate that, for example, when the approximate mode with the mean relative error distance (MRED) of 0.11 is considered, up to 43% energy saving is achieved. In addition, for this case, the Bias temperature instability (BTI)-induced delay degradation of the multiplier decreases up to 9.9% compared to 50% in the case of the exact mode. Also, the impact of process variations on the accuracy of the X-Dadda is studied. Finally, the efficacy of the X-Dadda multiplier, when used in neural networks for image classification and image-processing applications, is assessed.**

***Index Terms*— Accuracy configurability, aging reduction, approximate computing (AxC), energy-efficient multiplier.**

I. Introduction

LOWERING the power/energy consumption of embedded processors is of critical concern in the design of embedded processors. It originates from the fact that there are applications with heavy workloads which should be performed without exceeding a given power budget. This implies that the energy consumption for performing a given task should be lowered as much as possible. One of the approaches for improving the energy efficiency of computing systems is to use the approximate computing (AxC) paradigm. This paradigm may be invoked for applications where the impact of computation errors on the output quality degradation is tolerable. AxC may be employed at different levels of the design abstraction. At the hardware level, one of the approximation techniques relies on lowering the circuit supply voltage below the normal errorless one. The technique is called voltage overscaling (VOS) (see [1] and [2]). The use of VOS for realizing the AxC paradigm has the advantages of improving the energy efficiency as well as increasing the lifetime and reliability at the same time. The latter is especially important for highly scaled state-of-the-art technologies where the device dimensions have been considerably reduced. This has made satisfying reliability/lifetime requirements a challenging design task in the design of digital systems. The VOS technique also has another advantage of accuracy configurability, thanks to the ability of changing the VOS level during the runtime. Another approximation technique at the circuit level is based on circuit simplification/pruning where some of the gates are replaced with simpler ones or omitted to improve the energy efficiency of the circuit. Based on these techniques, different types of approximate multipliers have been suggested in the literature [3]. The large number of these approximate multipliers has been the motivation for some published works in the literature comparing the characteristics of these multipliers in terms of error metrics, delay, area, and energy consumption (see [4]).

In this article, an energy-efficient accuracy-configurable Dadda multiplier (X-Dadda) based on VOS is proposed and analyzed. To make the X-Dadda design an area-efficient one, the structure with 4-bit truncation is also considered. The efficiencies of the two designs considering different numbers of approximate columns and VOS voltage levels are thoroughly studied.

The main contributions of this article may be summarized in the following.

1) Design and analysis of the approximate Dadda multiplier structure based on selectively applying the VOS technique.
2) Proposing the hardware for realizing an accuracy-reconfigurable approximate multiplier with a very small overhead for runtime accuracy selection.
3) Investigating the error and energy consumption of the X-Dadda multiplier for different approximate bit widths for the cases of with and without truncation under six

Manuscript received October 26, 2019; revised February 2, 2020; accepted February 20, 2020. Date of publication March 25, 2020; date of current version April 24, 2020. This work was supported in part by a Grant from the Software and Hardware Foundations of the National Science Foundation. *(Corresponding author: Hassan Afzali-Kusha.)*

Hassan Afzali-Kusha, Marzieh Vaeztourshizi, and Massoud Pedram are with the Electrical and Computer Engineering Department, University of Southern California, Los Angeles, CA 90089-2560 USA (e-mail: afzaliku@usc.edu; vaeztour@usc.edu; mpedram@usc.edu).

Mehdi Kamal is with the School of Electrical and Computer Engineering, University of Tehran, Tehran 14399-57131, Iran (e-mail: mehdikamal@ut.ac.ir).

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

Digital Object Identifier 10.1109/TVLSI.2020.2978874

VOS voltage levels for the approximate part (the higher number of VOS voltage levels provide a higher accuracy granularity for the X-Dadda multiplier.)

4) Studying the effect of the VOS technique on the lifetime improvement of the X-Dadda multiplier hardware.
5) Exploring the effect of the process variation (PV) on the output error of the X-Dadda multiplier for different VOS voltage levels.

The rest of this article is organized as follows. A review of the related works is presented in Section II, while background materials related to aging, error metrics, and multiplier structure are presented in Section III. The structures of X-Dadda designs are described in Section IV. The results of investigating the efficiency of the multipliers are discussed in Section V. Finally, Section VI concludes this article.

II. Related Works

In this section, we briefly review a few works which are mainly more relevant to the proposed X-Dadda structure and/or included in the comparative study performed in Section V.

A. Approximate Multipliers Based on Circuit Simplification/Pruning

In [5], to build approximate multipliers, approximate compressors implemented using AND–OR gates were proposed. In addition, a simple algorithm for an efficient use of compressors in the first steps of the partial product reduction (PPR) tree of the multipliers was suggested.

In [6], first, an approximate compressor with a large output error was proposed. Then, by encoding the inputs of the compressor using generate and propagate signals, a more accurate compressor was suggested. By employing these improved approximate compressors, two 4×4 approximate multipliers used as the building blocks for the 16×16 and 32×32 multipliers were designed. In [7], noniterative and iterative approximate logarithmic multipliers (ALMs) were suggested. In the noniterative ALMs, three different approximate adders for the mantissa adder were used. The proposed iterative ALMs used a set-one adder (an n-bit set-one adder consists of two parts of an m-bit inexact adder and an (n-m)-bit exact adder where the sum of the m-bit adder is set to the logic one) for the mantissa adders during an iteration. They also utilized lower-part-OR adders and approximate mirror adders for the final addition. The work described in [8] suggested an approximate Booth multiplier which was based on approximate radix-4 modified Booth encoding (MBE) algorithms and a regular partial product (PP) array using an approximate Wallace tree. The authors also described two approximate Booth encoders for error-tolerant computing applications.

In [9], a simplification of 4:2 and 5:2 approximate compressors using the K-map were proposed. Furthermore, for more simplification, the carry in and carry out (C_{in} and C_{out}) of the compressors were removed. Finally, by realizing the proposed compressors, some approximate 8-bit Dadda multipliers were suggested. Using the approximate computing for simplifying the tree stage of the Dadda multiplier was discussed in [10]. The authors suggested altering the generated partial products to other partial products with some properties useful for the approximation. Based on the properties of the altered partial products, approximate half adder (HA), full adder (FA), and 4:2 compressors were suggested. By utilizing these cells as well as exact OR gates, two approximate multipliers were proposed. In the first one, the approximation was applied to all of the columns of the partial products of the multiplier while for the second one, the approximation circuits were used in the $n-1$ least significant columns of the multiplier.

An approximate multiplier (called BAM) where the carry-save adders (CSAs) were utilized for its PPR stage (an add and shift structure) was proposed in [11]. To improve the power and delay parameters, some of the CSAs were omitted horizontally and vertically. More specifically, some CSAs, in the first rows of the partial product array and some from the least significant columns of the array, were pruned. In [12], by using Cartesian genetic programming (CGP), a library of 8-bit approximate multipliers was generated. The structure of the multiplier was represented by integer chromosomes where the mutation operation was employed to generate new populations in the employed genetic algorithm. The fitness function was defined based on the mean relative error distance (MRED) of the circuits represented by the chromosomes. In addition, to limit the search space, some constraints on the error, power, and delay were defined. Based on this approach, a library containing 471 8-bit approximate multipliers was generated.

In [13], two 4:2 approximate compressors used for realizing approximate Dadda multipliers was suggested. Four approximate multipliers, based on the two suggested approximate compressor structures, were designed. In [14], an inaccurate 4:2 counter for reducing the partial products levels in the Wallace multiplier, without any modification to the other parts of the structure, was proposed. This inaccurate 4:2 counter was used to build a basic power-efficient inaccurate 4×4 Wallace multiplier. Arbitrary large multipliers have been built based on this multiplier. They also proposed an efficient error detection and correction circuit.

In [15], a partial product perforation technique for designing approximate multipliers was introduced. The authors mathematically proved that the error was bounded and predictable based on the input distribution.

B. Approximate Multipliers Based on VOS

Applying the overscaled voltage selectively to some parts of the arithmetic hardware unit has been introduced in prior works (see [16] and [17]) where the term biased VOS (Bi-VOS) was used to differentiate it from the case of applying it to the whole unit. This scheme offers less accuracy loss for a given energy reduction. In [16], the idea of Bi-VOS + SLEEP in a 32-bit floating point multiplier was presented. Since the multiplier hardware was not implemented, the error and energy were not accurately extracted. In [17], the usage of Bi-VOS for minimizing the error with a specified energy budget for an array multiplier was investigated. They derived few theoretical results for assigning the VOS voltage levels when designing an approximate multiplier with a minimum error, given an energy budget constraint [17]. Also, it was stated that the FAs of the same column should have the same voltage [17].

In this article, we focus on investigating an energy-efficient accuracy-configurable Dadda multiplier (X-Dadda) based

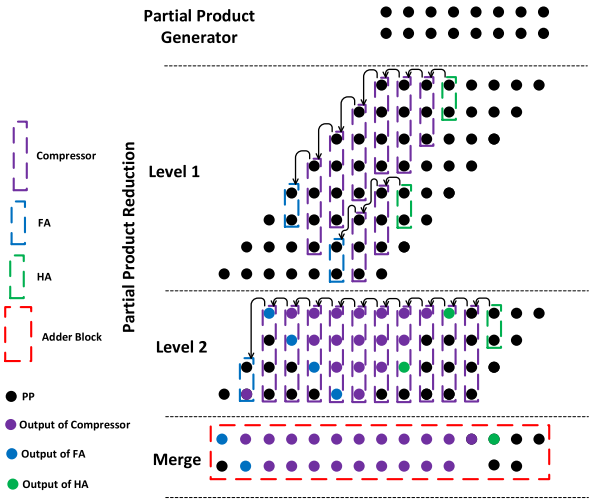

Fig. 1. Structure of an 8-bit Dadda multiplier [13].

on VOS. In our X-Dadda structure, we also apply the use of the VOS technique to some columns of the PPR and merge the stages of the Dadda multiplier (see Fig. 1). There are, however, several differences between our work and the ones suggested in [16] and [17] which focused on the floating point multipliers and array multipliers, respectively. Both these designs did not use voltage-level shifter circuits (for the transition from the low voltage to the high voltage) in their structures. The use of the level shifter reduces the output error as well as its uncertainty. In addition, none of these works suggested runtime accuracy configurability for the proposed multipliers. Both works of [16] and [17] did not implement the full hardware realizations of the multipliers with the Bi-VOS technique. Hence, a complete and accurate analysis of the output error and energy determination through the whole circuit simulation is not possible and hence they will not be included in the comparative study of Section V. More specifically, in [16], for calculating the output error of the 32-bit floating point multiplier, a simulator (based on the C programming language) was developed. The simulator used the error of a 1-bit FA calculated using HSPICE simulation. Also, for calculating the energy, they used HSPICE to calculate the energy consumed when the sum and carry (i.e., outputs) of a 1-bit FA was toggled [16]. Then, the number of toggles that occurred in a multiplier was obtained by simulating the Verilog hardware description language (HDL) of the circuit, and next, based on the amount of the toggles, the total energy consumed in a multiplier was determined [16]. The authors assumed that the energy was consumed only when any output changed [16]. Moreover, the static energy consumption was not considered in this article. Lau *et al.* [17] used a high-level analytical expression for measuring only the output error of the array multiplier. In addition, for extracting the energy, short-circuit, leakage, and dc current drawn from the supply voltage were not considered. Finally, none of these works reported the standard error metrics including mean error distance (ED), MRED, and mean normalized ED (MNED) for evaluating the accuracy of the Bi-VOS based approximate multipliers.

C. Accuracy-Configurable Approximate Multipliers

Supporting output accuracy configurability is a feature which has received renewed attention for designing approximate multipliers in dynamically adjustable quality-of-service (QoS) applications. In [18], four configurable approximate compressors for implementing approximate multipliers were suggested. The multiplier structures had the flexibility of switching between the exact and approximate operation modes. In [19], methods for designing quality-configurable circuits were developed. The method was based on CGP in which the exact and approximate modes of the circuit were realized at the same time. The method was experimentally evaluated by designing configurable approximate multipliers which had both exact and approximate modes. This multiplier only has two modes of exact and approximate computations and hence was not considered in our comparative study. In [20], an accuracy-configurable multiplier is proposed. The final product of the multiplier is generated by a carry-maskable adder (CMA). The accuracy of the proposed multiplier can dynamically be configured by changing the length of the carry propagation. Furthermore, the partial product tree of the multiplier is approximated by the proposed tree compressor. An extension of this design which had smaller area and power consumption was described in [21]. The authors used a simpler approximate tree compressor. In [22], an approximate multiplier based on an approximate adder design which limits its carry propagation to the nearest neighbors was suggested. The adder was used for fast partial production accumulation. Different levels of accuracy were provided through their configurable error recovery circuit realized using OR gates or the proposed approximate adders. Since the proposed structure in this article does not use error recovery unit, this multiplier was not included in the comparative study of this article.

III. Preliminaries

A. Aging Mechanisms

While improving transistor aging and reliability has been always one of the design challenges, the problem becomes worse as the technology scales more [23]. There are several circuit aging mechanisms including, but not limited to, Bias temperature instability (BTI), hot carrier injection (HCI), time-dependent dielectric breakdown (TDDB), and electromigration (EM). Among these mechanisms, the first three are directly related to the circuit lifetime where HCI and BTI cause some increase in the threshold voltage (V_{th}). The latter causes gate oxide deterioration which eventually would lead to the device failure. When the threshold voltage of the transistor increases, the delay of the circuit which has a direct relation to the threshold voltage increases. This eventually can lead to the violation of the circuit timing for a given clock frequency. In the case of device failure (transistor breakdown), the circuit will not operate properly. All of these three mechanisms have strong dependencies on the electric field which is proportional to the supply voltage. When lower supply voltages are used in the case of overscaling technique, the amount of the delay degradation induced by the circuit aging will be lower than that of case of the nominal supply voltage. Also, for TDDB,

TABLE I
BTI MODEL PARAMETERS [25]

Device	A	κ	α	β	γ
NMOS	3.12×10^{-2}	50	0.158	1/6	5.2
PMOS	2.02×10^{-2}	50	0.173	1/6	3

lower electric fields in the device gives rise to less damage to the dielectric of the device. Therefore, the device degradation takes place at a slower pace. In this article, we only consider the impact of BTI on the threshold voltage increase.

The BTI affects both NMOS and PMOS transistors by creating interface traps at the Si/SiO_2 interface [24]. These traps cause a threshold voltage shift modeled by [25]

$$\Delta V_{\text{th,BTI}} \cong A e^{-\frac{\kappa}{\theta}} t^{\alpha} E_{\text{OX}}^{\gamma} f^{\beta} \tag{1}$$

where t is the total stress time (s), θ is the temperature (K), f is the duty factor of the stress signal, and $E_{\text{OX}} = (V_{\text{DD}} - V_{\text{th}})/T_{\text{INV}}$ is the electric field across the gate oxide (where T_{INV} denotes the thickness of the gate inversion layer). This relation shows strong dependence of the aging mechanism to V_{DD}. The technology-dependent parameters of A, κ, α, β, and γ are reported in Table I.

B. AxC Quality Metrics

The output quality of the approximated hardware may be evaluated by some quality/accuracy metrics [18]. One of the parameters is the ED which is defined by the absolute difference between the exact and approximated values of the output. In addition, there are other closely related parameters to ED which include the mean of ED (MED), MRED, and MNED [18]. In this article, we considered these parameters for our evaluation of the approximate designs.

C. Dadda Multiplier

The structure of an 8-bit Dadda multiplier is shown in Fig. 1 [13]. For most of this article, including the comparative study part, the focus is on the 8-bit X-Dadda multiplier because of the high usage of low bit multipliers in image processing and machine learning applications. This multiplier is constructed from three stages. In the partial product generation (PPG) stage, "AND" operations of the bits of the multiplicand with the bits of the multiplier are performed. In the PPR stage, eight partial product terms are compressed to two terms by adding them in some way. Finally, in the merge stage, the final two terms are added to produce the output. Among these stages, PPR has the highest delay, area, and power consumption [26]. To reduce the delay of the PPR stage, 4:2 and 5:2 compressors are vastly used [26]. For the X-Dadda design, the exact 4:2 compressor, which is comprised of two FAs connected serially, is used [18].

IV. X-DADDA STRUCTURES

A. Design Based on Bi-VOS

Applying the VOS technique to the circuit increases its delay. A simple first-order approximation circuit delay model based on the alpha power law is given by

$$\text{Task}_{\text{Delay}} \propto \frac{V_{\text{DD}}}{(V_{\text{DD}} - V_{\text{th}})^{\alpha}} \tag{2}$$

where $\text{Task}_{\text{Delay}}$ is the delay of the circuit and α is a technology-dependent parameter considered to be 1.3 for sub-20-nm technologies [27]. This suggests that the use of the VOS technique for the combinational circuits may cause the violation of the setup time (of the destination flip-flops operating at nominal voltage).

This violation may occur for the outputs corresponding to the longest timing paths of the circuit [2]. The number of timing paths with failed timing requirement increases as the supply voltage is decreased to lower VOS voltage levels where the output error and quality degradation increase significantly [2]. This implies using VOS for all the bits including most significant bits (MSBs) and least significant bits (LSBs) will hinder taking the full advantage of efficient use of the large energy reduction potential of the VOS technique [28]. To overcome this problem, in this article, inspired by the idea of Bi-VOS employed in [16] and [17], we have applied different VOS levels to different bit positions of parallel multipliers. Additionally, we consider two designs for the Dadda multiplier where the first design only uses Bi-VOS while the second one employs Bi-VOS along with some bit truncation.

Using the VOS for LSBs decreases the energy and improves the lifetime without increasing the error of the multiplier output excessively.

Owing to the considerable impact on the output accuracy, in the hardware part of the MSBs, we use the nominal voltage to make its operation errorless. The structure of the considered Bi-VOS design is shown in Fig. 4 (in the case of with and without truncation). In this structure, there are two supply voltages, where one is for the accurate part, denoted here as $V_{\text{DD\_AC}}$, and the other one is for the approximate part, denoted as $V_{\text{DD\_AP}}$, whose value is set through a power switch box (see Fig. 4). The accuracy of the multiplier can be configured in the runtime by changing $V_{\text{DD\_AP}}$ with a negligible delay overhead. It should be noted that the approximate width (width__region2), which is another approximation knob, is set only at the design time.

We utilize level shifters for matching the voltages of the outputs produced by the half adders, FAs, and compressors of the approximate columns to the supply voltage of the half adders, FAs, and compressors of the accurate columns. To lower the latency overhead of the level shifter, a low-latency level shifter circuit proposed in [29] was used. Of course, depending on the design constraints and requirements, other level shifters (see [30]) in the X-Dadda structure may be utilized. To minimize the area overhead, all of the transistors used in the level shifter of [29] were minimum sized. In Fig. 2, we have depicted the circuit diagram of the level shifter whose delay versus $V_{\text{DD\_AP}}$ characteristic is shown in Fig. 3.

B. Design Based on Combination of Bi-VOS and Truncation

To reduce the area of the previous X-Dadda circuit, one may use the truncation technique as well, which causes some

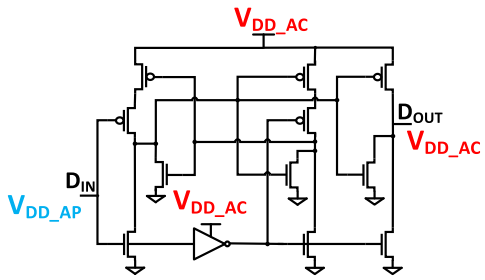

Fig. 2. Circuit diagram of the level shifter employed in this article [29].

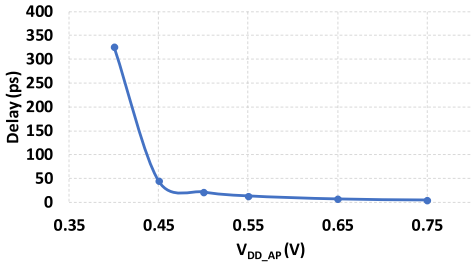

Fig. 3. Delay versus V_{DD_AP} of the level shifter.

output error increase (see [31]). Normally, as the truncation width increases, the output error becomes larger. In this article, without loss of generality, as an example, we only considered truncation of the first four LSB bits of the circuit. In Fig. 4, the boundary of the truncation is indicated by a yellow dashed line. Based on the energy, area, and error budget constraints, other truncation widths may be chosen by the designer. Also, it should be mentioned that the same level shifter was used for this design too.

V. Results and Discussion

A. Simulation Setup

All the studies were performed by employing the 15-nm FinFET-based Open Cell Library (OCL) technology [32] with a nominal operating voltage level of 0.80 V (V_{DD_AC}). We considered six VOS voltage levels for the approximate part (V_{DD_AP}) including 0.40, 0.45, 0.50, 0.55, 0.65, and 0.75 V. The level shifter cell was added to the technology file which was characterized for all the considered VOS voltage levels. For the evaluation, first, we described each 8-bit X-Dadda structure with Verilog HDL separately based on the considered approximate part width and also the length of the truncation. For this article, in the case of without truncation (with 4-bit truncation), the length of the approximate part was varied from 2-bit to 14-bit (1-bit–10-bit).

More specifically, 13 designs for the X-Dadda structure without truncation and ten designs for the X-Dadda structure with truncation were studied. The designs were synthesized by considering V_{DD_AC} and V_{DD_AC} of 0.8 and 0.45 V, respectively. In this article, Synopsys Design Compiler (DC) tool was utilized for synthesizing the designs. For each width_region2, the synthesized output gate-level netlist of its corresponding design was employed for studying the impact of the V_{DD_AP} value on the design parameters of the X-Dadda.

Thus, for each chosen operating VOS voltage level of the approximate part, the DDC (Synopsys internal database format) of the structure with the considered approximate bit

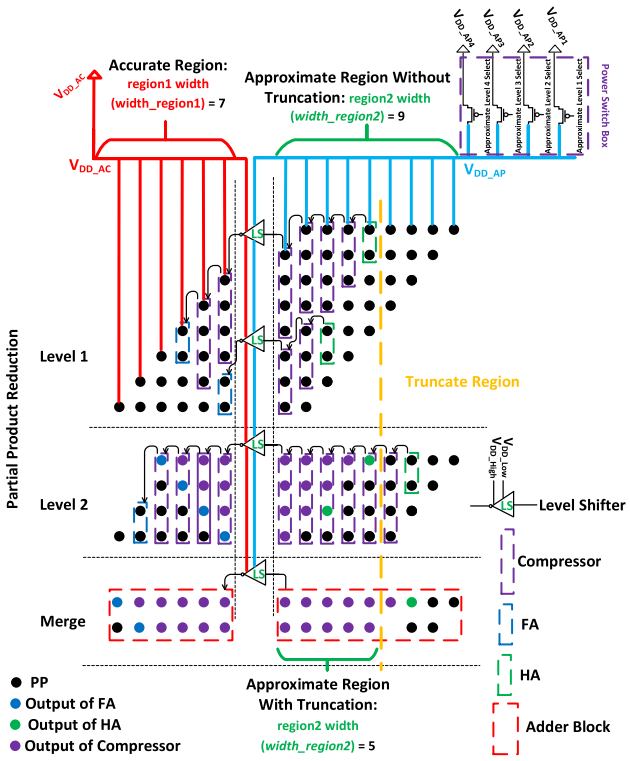

Fig. 4. Structure of the 8-bit X-Dadda design with width_region2 of 9 and 5, illustrating the cases of without and with truncation.

width and the proper technology file were fed to the synthesis tool. Due to the existence of two voltage domains in the X-Dadda structure, the united power format (UPF), a standard for specifying the low-power design intent for the chip design, was utilized with the netlist during synthesize.

It should be noted that each of the gate-level netlists had different number of level shifters whose positions in the circuit netlist tended to be different. These two factors caused the design parameters of the studied X-Dadda to deviate slightly from the expected trend of the design parameters when the approximation width (width_region2) is varied. Note that the width of the exact part is denoted as width_region1.

As mentioned before, in the VOS-based design, the target delay is not changed by changing the operating voltage level. In our investigation, the target delay for each of the 23 (13 + 10) designs was determined by the postsynthesis simulation of the gate-level netlist of the design by considering $V_{DD\_AP} = 0.8$ V (exact operation). For the post-synthesis simulations, based on the chosen operating VOS voltage level of the approximate part, the proper standard delay format (SDF) file was generated. The file was generated using the synthesis tool when employing the characterized technology file at that operating VOS voltage level. We used the Synopsys VCS tool for the simulations during which one million random inputs were injected as the inputs to the circuit.

The target delays of the considered designs are provided in Table II for the cases of with and without truncation, respectively. These delays were used for setting the clock frequency of the flip-flops (operating at 0.80 V) which sampled the X-Dadda outputs. When lowering V_{DD_AP}, this timing requirement was not satisfied, potentially causing erroneous

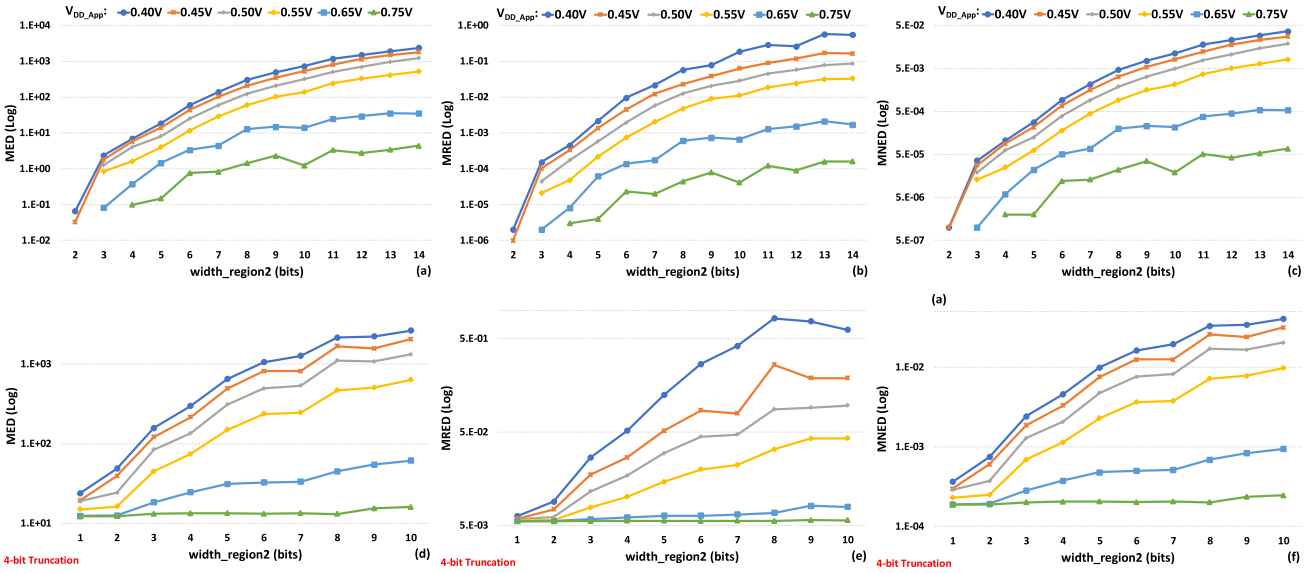

Fig. 5. MED, MRED, and NED of the 8-bit X-Dadda multiplier (a)–(c) without and (d)-(f) with 4-bit truncation under different V_{DD_AP} and width__region2.

TABLE II
TARGET PERFORMANCES FOR DIFFERENT DESIGNS WITHOUT AND WITH TRUNCATION AT THE SUPPLY VOLTAGE OF 0.80 V

Designs without Truncation		Designs with Truncation	
width_region2	Target Performance @ 0.80V (ps)	*width_region2*	Target Performance @ 0.80V (ps)
1-4	750	6	680
5, 6, 8, 13	760	5	690
9, 11	770	1	710
10, 14	780	8	720
12	800	9	730
NA	NA	3	740
NA	NA	4, 10	750
NA	NA	2, 7	760

outputs (therefore, there was no need for considering special timing requirements for the flip-flops (register) used with the proposed multiplier). Finally, we ought to emphasize that we forced the DC synthesis tool to generate the same structure for all the designs with the same approximate part widths. The energy consumptions of the designs were also extracted based on the design parameters reported by DC.

B. Results

1) Nominal X-Dadda Structures: The error metric parameters of MRED, MED, and MNED of the 8-bit X-Dadda multiplier in the two cases of without truncation (*w/o_t*) and with truncation (*w_t*) under different width__region2 and VOS voltage levels (V_{DD_AP}) are plotted in Fig. 5. As was expected, by increasing width__region2 and also decreasing V_{DD_AP}, the increases in output error increase MED, MRED, and MNED parameters. The results for width__region2 equal to 1 has not been shown since the effect of the technique will be limited. Among the studied structures, the maximum MED ($\sim$2.65E + 03), MRED ($\sim$8.22E–01), MNED ($\sim$4.05E–02) belonged to the *w_t* X-Dadda multiplier with width__region2 of 10, 8, and 10 with $V_{DD_AP} = 0.40$V, respectively. It should be noted that as discussed before, there are few points in the results that do not follow quite the expected trend. This may be attributed to different employed gates and corresponding sizes resulted from synthesizing different X-Dadda structures independently.

To study the impact of approximate knob adjustment on the fluctuations of the error metrics of the proposed multiplier, we make use of an error sensitivity (ES) parameter. The sensitivity of the error metric X (MED, MRED, or NED) to the change of V_{DD_AP} (from $V_{DD\_AP,i}$ to $V_{DD\_AP,i+1}$) for a given width__region2$_k$ is defined as (3), shown at the bottom of the next page.

where the function E_X is the error metric X for the specified $V_{DD\_AP,i}$ and width__region2$_k$. The index $i (0 \leq i < 5)$ is used to point to the voltages in the set of {0.75, 0.65, 0.55, 0.50, 0.45, and 0.40 V}. The results of this investigation for MRED revealed that, on average (for all considered width__region2), ES_{MRED} increased 319$\times$ (484$\times$) by decreasing V_{DD_AP} from 0.75 to 0.45 V for the *w/o_t* (*w_t*) design. This implies a much higher sensitivity for the structure at lower V_{DD_AP} values. In other words, at lower VOS voltage levels, going from one level to the next higher level, improves MRED considerably. Similarly, the ES to the change of width__region2 was examined. The obtained results indicated that, on average (for all considered V_{DD_AP} values), ES_{MRED} increased 1204$\times$ (123$\times$) by enlarging width__region2 from 2 to 13 (1–8) for the case of *w/o_t* (*w_t*) design. Again, this indicates that when the number of approximation bits is large, reducing the number of approximation bits by even one bit improves MRED significantly.

Next, we study the energy consumptions of the X-Dadda structure versus width__region2 with V_{DD_AP} as the running parameter. As may be observed from the results, which are shown in Fig. 6, the energy consumption reduces

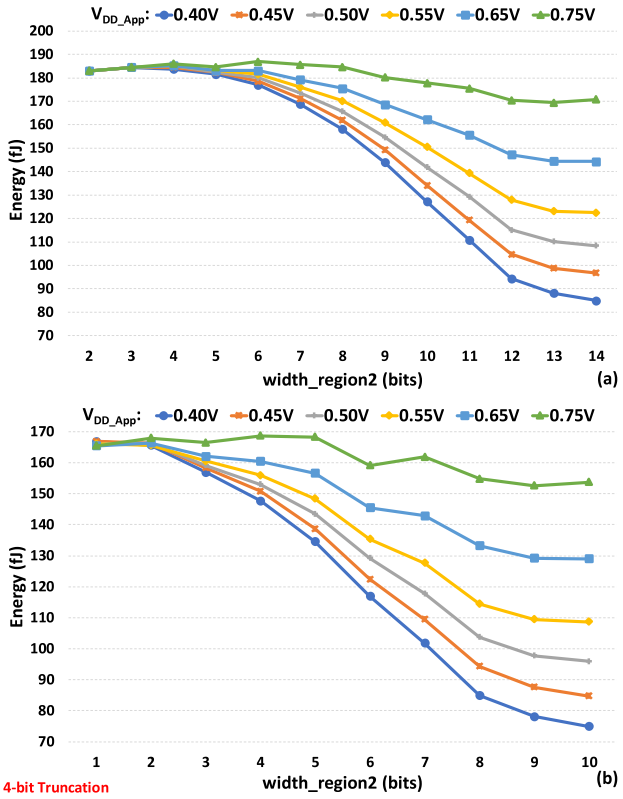

Fig. 6. Energy consumption of the 8-bit X-Dadda structures versus width__region2 for (a) without and (b) with truncation designs for different VOS voltage levels.

when V_{DD_AP} decreases and/or width__region2 increases. More specifically, increasing width__region2 from 2 to 14 [see Fig. 6(a)], the energy consumption reduction enlarges from ∼0.4% to 42% for both *w/o_t* and *w_t* designs. For instance, in the case of $V_{DD\_AP} = 0.4$ V, increasing width__region2 from 2 to 14 (1–10) causes energy reduction changes from 0% to 60% compared to the nominal energy consumption the case of *w/o_t* (*w_t*) design, respectively. For $V_{DD\_AP} = 0.75$ V, the corresponding reductions are from ∼0% to 18% (1%–21%) which is lower than the other VOS voltage level. To further investigate the characteristic of the structure, let us define energy reduction sensitivity (ERS) to width__region2. The sensitivity, which may also be considered as the energy gain efficiency, is calculated from the change in the energy reduction when increasing width__region2 by one. As a specific example, by increasing width__region2 from 4 to 12 (2–7) for the case of *w/o_t* (*w_t*) design, we obtain a ∼2× (∼2.2×) increase in the ERS. The values were obtained by averaging over all the VOS voltage levels.

It should be noted that the delay and area of the switches for connecting lower V_{DD}'s are important parameters in the power switching approaches. Since in this article, we have used FinFET devices, the number of the fins determines the delay and area overheads of the switches. We used seven devices for connecting the seven voltage levels to region 2. The highest number of fins corresponded to the case of the largest width__region2 (largest capacitance for region 2) which was 14 in the case of the 8-bit X-Dadda structure. It should be stated that the number of fins of the switches were determined based on keeping the VOS voltage IR drop below 10%. The number of the fins for the voltages of 400, 450, 500, 550, 650, and 750 mV were chosen to be 37, 27, 22, 19, 14, and 12. The area overheads of the employed switch boxes were less than 0.1% of the area of the proposed X-Dadda multiplier.

Finally, we provide some error results for the 16-bit structure in Fig. 7 which show MRED, MED, and MNED versus the width__region2. As is observed from Fig. 7 that similar characteristics to those of the 8-bit structures exist for this structure. The trend is the increase in the error metric as width__region2 increases. The energy reduction of the structure versus width__region2 is plotted in Fig. 8. Again, the energy consumption improves as the width increases.

While the energy reduction rate is higher for lower supply voltages, the error increase rate is also higher for these voltages. Due to the space limitation, the characteristics for the case of the bit truncation are not provided here.

2) Impact of PV on X-Dadda Structure: Operating a circuit at lower voltages than the nominal one in the presence of the PV may exacerbate the impact of variation on the circuit characteristics. This is owing to the fact that the operating region of devices may become close to the near threshold region [33]. In this part, we study the impact of the PV on the accuracy of the proposed multiplier structure in the case of without truncation. The study was performed for two low VOS voltage levels of 0.40 and 0.50 V applied to region 2 when considering different width__region2 values.

The study included both local and global variations. For the global PV, the Gaussian distributions for the channel length (L_g), the fin thickness (t_{si}), and the height of fin (H_{fin}) with $3\sigma = 10\%$ of their nominal values and for the gate oxide thickness with $3\sigma = 5\%$ of its nominal value were considered [34]. Also, the local variability was only assumed for t_{si} and L_g due to line edge roughness (LER) phenomenon [34].

Based on the above variation models and by employing the Monte Carlo simulation with the Synopsys HSPICE tool, delay variations of the gates were extracted. Then, for each design, 5000 SDF files were generated based on the delay variations of the gates. To do this, we developed an in-house tool that applied the delay variation of the gates to the SDF files of the design which were extracted by the synthesis tool (DC) without considering the PV. This way, we used the transistor-level specs for considering the impact of the PV on the gates which, in turn, was utilized to investigate the impact of the PV on the proposed multipliers.

Based on the generated SDF files, each design was simulated by Synopsys VCS tool 5000 times (based on the

$$ES_X(V_{DD\_AP,i}, \text{width\_\_region2}_k) = \frac{E_X(V_{DD\_AP,i+1}, \text{width\_\_region2}_k) - E_X(V_{DD\_AP,i}, \text{width\_\_region2}_k)}{V_{DD\_AP,i} - V_{DD\_AP,i+1}} \quad (3)$$

TABLE III

ACCURACY VARIATION OF THE ERROR PARAMETERS OF THE PROPOSED APPROXIMATE MULTIPLIER (WITHOUT TRUNCATION) UNDER THE PV FOR VOLTAGE LEVELS OF 0.40 AND 0.50 V FOR THE APPROXIMATE PART

width_region2		(V_{DD_AC}, V_{DD_AP})= (0.80V, 0.40V)			(V_{DD_AC}, V_{DD_AP})= (0.80V, 0.50V)		
		MED	MRED	MNED	MED	MRED	MNED
5	Nominal	1.84E+01	2.16E-03	2.82E-04	8.13E+00	5.81E-04	1.25E-04
	Mean	1.86E+01	2.18E-03	2.86E-04	8.02E+00	5.78E-04	1.23E-04
	Std	3.19E-01	3.55E-05	4.92E-06	2.45E-01	1.61E-05	3.78E-06
	Mean/Std	58	61	58	33	36	33
8	Nominal	3.13E+02	5.89E-02	4.81E-03	1.31E+02	1.30E-02	2.02E-03
	Mean	3.14E+02	5.97E-02	4.83E-03	1.32E+02	1.31E-02	2.03E-03
	Std	2.87E+00	7.87E-04	4.41E-05	2.65E+00	1.64E-04	4.08E-05
	Mean/Std	110	76	110	50	80	50
10	Nominal	8.14E+02	2.03E-01	1.25E-02	3.44E+02	3.23E-02	5.28E-03
	Mean	8.14E+02	2.05E-01	1.25E-02	3.42E+02	3.20E-02	5.26E-03
	Std	6.23E+00	2.64E-03	9.58E-05	4.58E+00	4.21E-04	7.04E-05
	Mean/Std	131	78	131	75	76	75
14	Nominal	2.37E+03	5.46E-01	3.65E-02	1.23E+03	8.55E-02	1.89E-02
	Mean	2.40E+03	5.52E-01	3.69E-02	1.21E+03	8.43E-02	1.86E-02
	Std	2.37E+01	9.53E-03	3.65E-04	1.65E+01	8.47E-04	2.53E-04
	Mean/Std	101	58	101	73	100	73

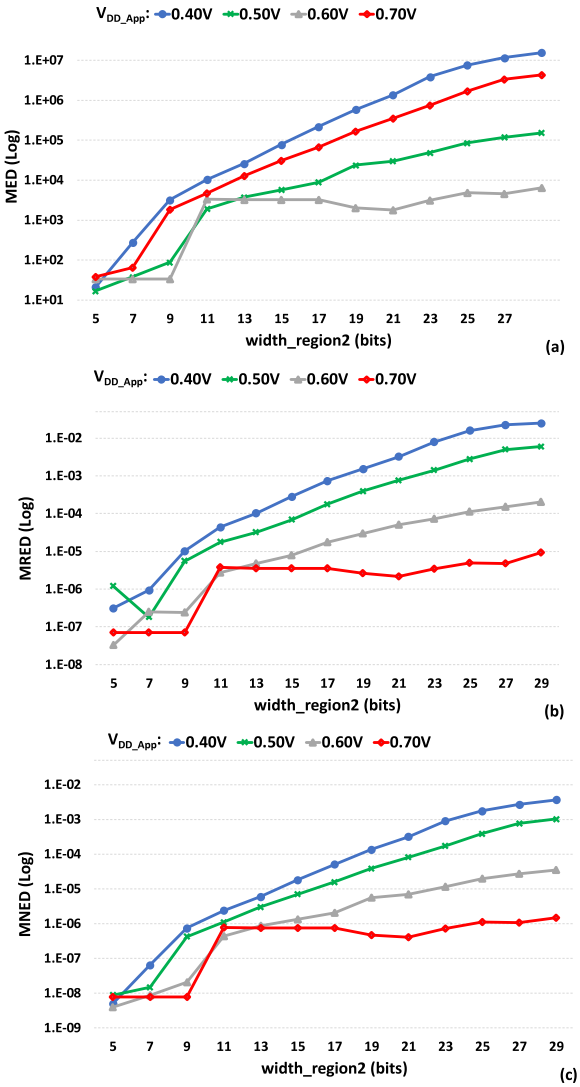

Fig. 7. (a) MED, (b) MRED, and (c) NED of the 16-bit X-Dadda multiplier under different V_{DD_AP} and width__region2.

SDF files), and the mean and standard deviation of the MED, MRED, and MNED parameters were obtained. The extracted accuracy variations are reported in Table III which shows that

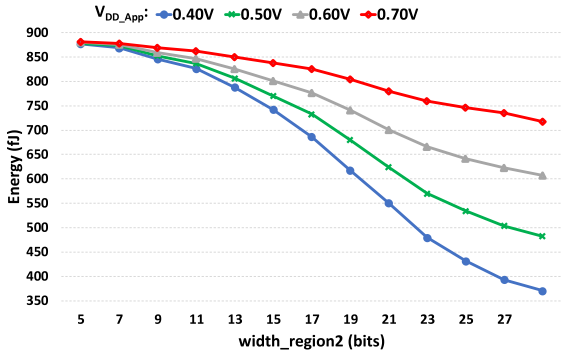

Fig. 8. Energy consumption of the 16-bit X-Dadda structures versus width__region2 for different VOS voltage levels.

the PV does not have a significant impact on the error metrics. The results indicate, the mean to standard deviation ratio of all the parameters for the designs are greater than 6, revealing that the designs are robust with respect to the PV. In three out of the eight designs, MRED decreases in the presence of the PV in comparison to the corresponding error parameters of the nominal design. The maximum increase in the MRED is 1% and the largest reduction of the MRED error parameter is 2% compared to the nominal value of the MRED.

3) Aging Effects: The threshold voltages (V_{th}'s) of NMOS and PMOS of the considered technology at the nominal voltage (i.e., 0.80 V) are 0.175 and −0.190 V, respectively. Also, the corresponding voltages of NMOS and PMOS at the overscaled voltage of 0.40 V are 0.205 and −0.181 V, respectively. The increases in the threshold voltage magnitudes of NMOS and PMOS versus the supply voltage over ten years stress are plotted in Fig. 9. The results, which were obtained using (1), indicate that the magnitude of NMOS and PMOS threshold voltages would increase 0.151 and 0.190 V (∼0.000 and ∼0.001 V), respectively, for the supply voltage of 0.80 V (0.40 V). For the 0.40-V operating voltage level, over ten years, the delay (time of switching) will increase about 0.3% (0.2% for NMOS and 0.4% for PMOS). For determining the delay increase, (2) was utilized. As mentioned before, one of the advantages of the X-Dadda structure

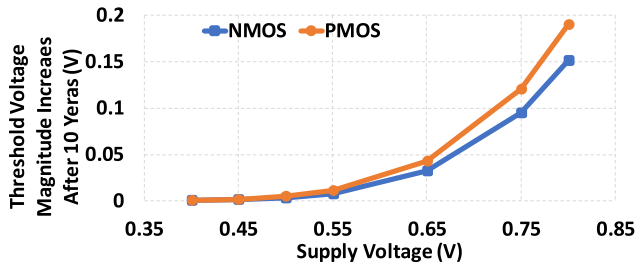

Fig. 9. Increase in the threshold voltage magnitude for NMOS and PMOS versus the supply voltage after ten years.

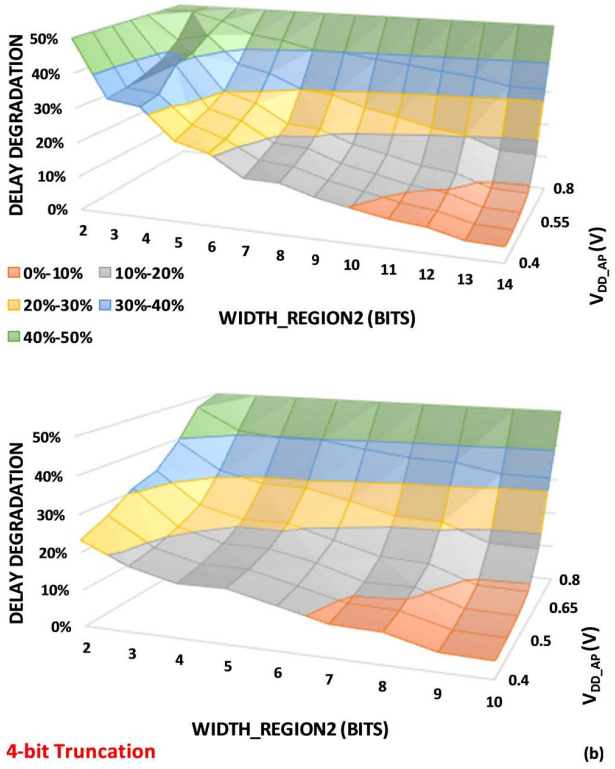

Fig. 10. BTI-induced delay degradation of the explored structures for different V_{DD_AP} and width_region2 values for (a) without and (b) with truncation designs for different VOS levels.

compared to other accuracy-configurable approximate multipliers is its higher lifetime/reliability, thanks to operating at lower operating voltage levels.

To demonstrate this, in Fig. 10, the BTI-induced delay degradation of the explored structures versus V_{DD_AP} has been plotted. Let us consider the notation of TXWYVZ for different X-Dadda structures. In this notation, X indicates the number of truncation bits, Y shows the width of region2, and Z is the VOS voltage level of the region2. Based on the results presented in Fig. 10, in the best (worst) case, after ten years, the delay of the T0W14V0.4 (T0W1V0.8) increases about 4% (∼50%). The delay degradation of the X-Dadda structure over time depends on V_{DD_AP}. By assuming that the X-Dadda structure experiences all the considered operating voltage levels of this article with similar probability ($P(V) = 1/7$), in the case of width_region2 = 14 (8), the delay degradation of the X-Dadda will be 18% (27%). It should be noted that these results are for the case of without truncation design. Also, it is clear that in the case of *w_t*, due to the fewer number of gates in the critical path, the delay degradation is smaller than the case of *w/o_t*.

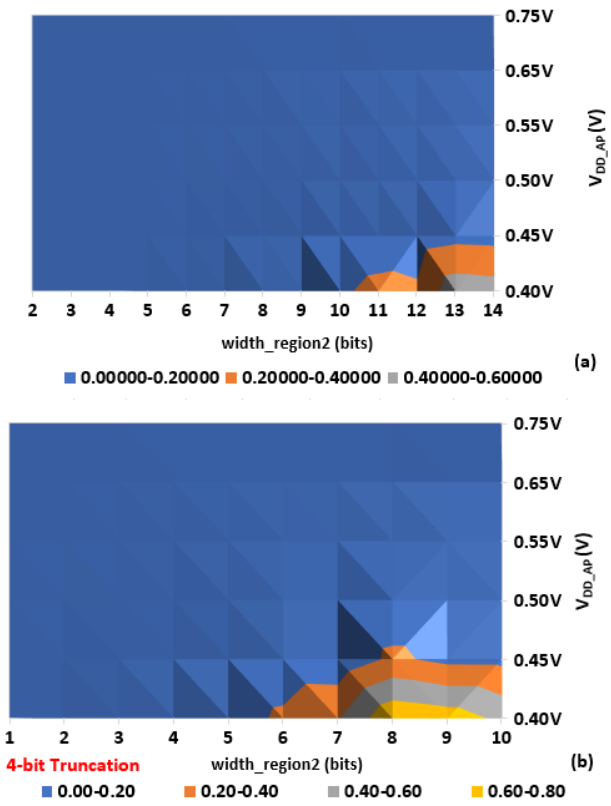

Fig. 11. MRED reduction of the X-Dadda multiplier when working in the approximate mode and the clocks of the designs were set as the delay of the X-Dadda multiplier in exact mode after ten years of operation for (a) without and (b) with truncation.

By increasing the delay of X-Dadda during its lifetime, to guarantee that the X-Dadda multiplier in its exact operating mode does not have any error, the clock period for sampling its output should be increased over time accordingly.

On the other hand, as we have shown, due to lower V_{DD_AP}, the delay degradation rate of the X-Dadda, while operating in the approximate mode, is lower. Hence, it is expected that even when not adjusting the clock period over time, the X-Dadda outputs should have less error compared to the other approximate structures which are based on the circuit simplification/pruning approach (see [5]), in a lifespan of ten years. To illustrate this, in Fig. 11, we have plotted the MRED reduction of the X-Dadda when the clock of the designs was set as the delay of the X-Dadda in the exact mode after ten years of operation. Obviously, the exact X-Dadda works precisely after ten years under the considered clock period.

In the cases of *w/o_t* and *w_t* designs, as V_{DD_AP} decreases, the MRED reduction increases for different width_region2 values. The highest reduction for the designs *w/o_t* (*w_t*) corresponds to $V_{DD\_AP} = 0.40$ V and width_region2 = 13 ($V_{DD\_AP} = 0.40$ V and width_region2 = 10).

4) Comparing X-Dadda Structures With Some Prior Approximate Multipliers: To evaluate the efficacy of the proposed X-Dadda multiplier, we have compared the characteristics of the X-Dadda multiplier with those of some

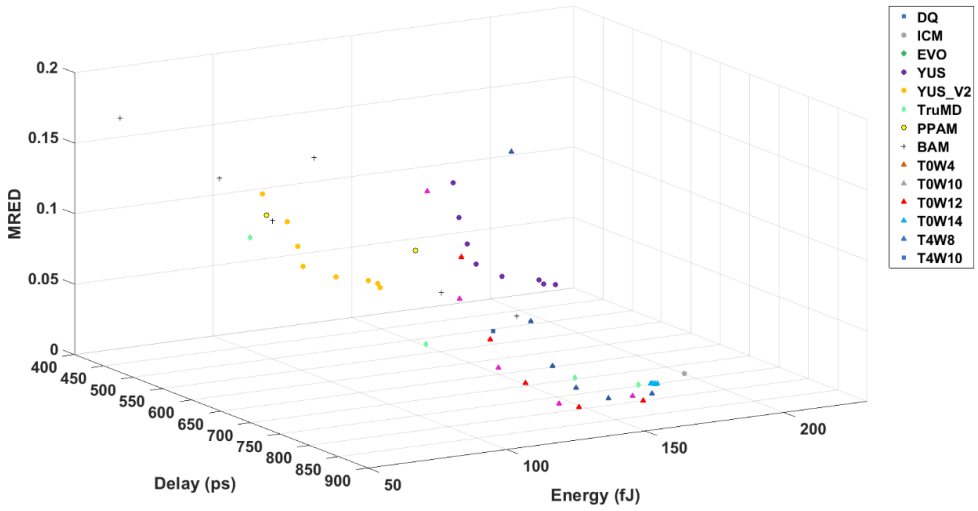

Fig. 12. MRED versus energy and delay for different approximate multipliers.

TABLE IV
FEATURES OF THE STUDIED MULTIPLIERS

Structure	Runtime Accuracy Configurable	Voltage Overscaling	Circuit Simplification/Pruning
YUS [20]	✓	×	✓
YUS-V2 [21]	✓	×	✓
EVO [12]	×	×	✓
BAM [11]	×	×	✓
PPAM[15]	×	×	✓
TruMD	×	×	✓
DQ [18]	✓	×	✓
ICM [14]	×	×	✓
X-Dadda	✓	✓	×

prior art multipliers given in Table IV. Different approximate multipliers may be compared in terms of the design parameters such as error, energy, and delay. The critical design parameter is the energy which is the main purpose of invoking the AxC paradigm in the first place. In the second tier of importance, the delay should also be considered. For the X-Dadda multiplier, different MRED values were obtained by varying the overscaled voltage which also resulted in lower energy consumption.

Other approximate multipliers included in the comparison are BAM [11], two multipliers from EVO [12], ICM [14], PPAM [15], DQ [18], YUS [20], YUS-V2 [21], and TruMD (which has truncated 2, 4, 6, and 8 most LSBs of the Dadda multiplier). Among these multipliers, YUS, YUS_V2, and DQ have accuracy configurability during the runtime which is the key feature (advantage) of the proposed X-Dadda structure. As discussed previously, in [18], four dual-quality reconfigurable approximate 4:2 compressors with two modes of exact and approximate were considered. These compressors had two parts of approximate and supplementary where the supplementary part was power gated during the approximate mode. By providing the exact mode signal, the supplementary part is added to the approximate part of the multiplier to provide the exact mode for the multiplier. Also, the works in [20] and [21] used a CMA in the merge stage of the multiplication. This CMA can be configured to function as carry propagation adder (CPA), a set of bit-parallel OR gates, or a combination of the two. This configurability is realized by using mask signals to mask the carry propagation.

In Fig. 12, we have used a 3-D plot for showing the MRED value versus the energy and delay of the multipliers considered in the comparative study. In the cases of fixed accuracy designs, each point indicates the delay and energy of each structure for its given error. Different points for this type of structures correspond to different instances of the design. For the accuracy-configurable multipliers, different points correspond to the change in the approximation level during the runtime by their proper means. For a better illustration, Figs. 13 and 14 show the energy-MRED and delay-MRED features of the approximate multipliers using 2-D plots. For accuracy-configurable designs, the energy versus the output error characteristic normally shows a pareto-optimal behavior, meaning that improving the accuracy is equivalent to the energy increase. For fixed accuracy structures, different designs are located at different points in the energy-error plane. Of course, changes might be applied to the designs themselves for obtaining multipliers with different points of energy error.

The energy comparison, which is shown in Fig. 13, for most of the characteristics, reveals a pareto-optimal behavior, as was expected. The proposed X-Dadda structure can provide lower energy consumptions for lower MRED values when compared to other structures. When truncation is used for the X-Dadda structure, the energy efficiency improves. More specifically, for small MRED values, T4W10 shows the lowest energy consumption. For MREDs between 0.015 and 0.10, TruMD (b-bit truncation) has the lowest energy consumption. The TruMD structure, however, is not a configurable multiplier.

For the energy consumption below 84 fJ, T4W10, PPAM, BAM, EVO, and DQ are the designs which can provide lower energy consumptions. Between these designs, PPAM is the most accurate design. Our T4W10 design can also provide this amount of energy consumption with a lower accuracy.

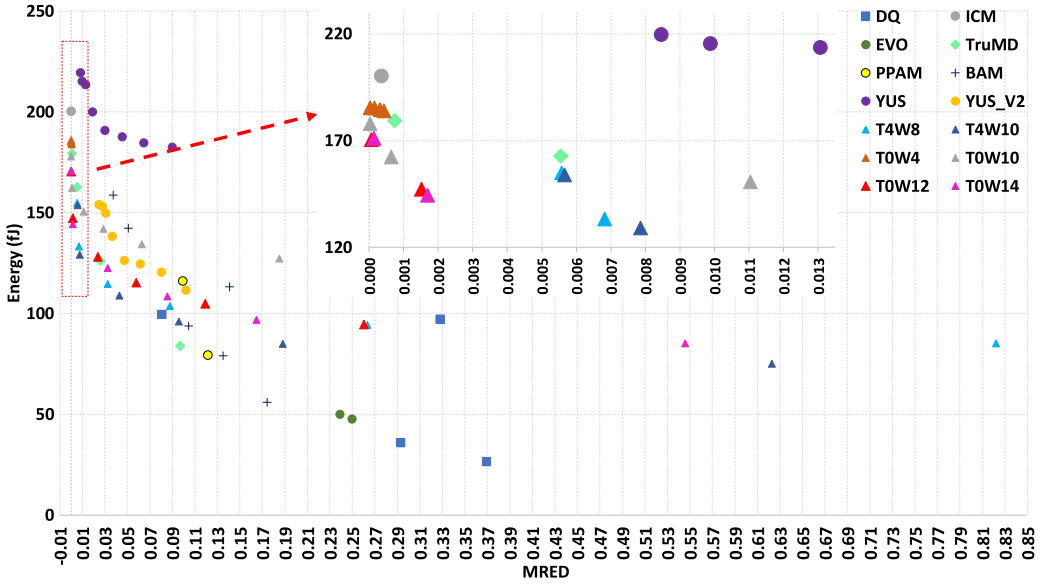

Fig. 13. Energy versus MRED for different approximate multipliers.

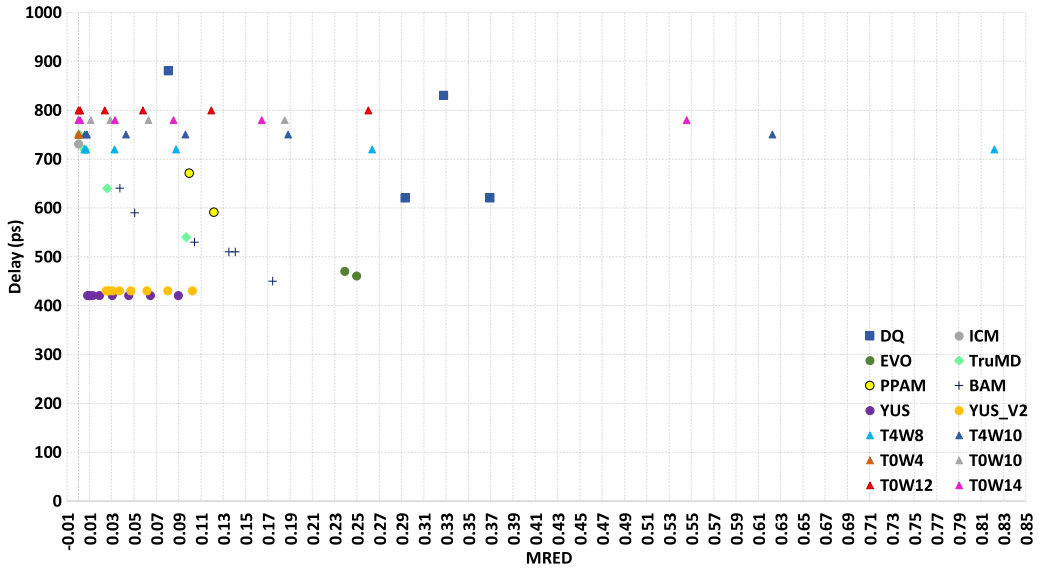

Fig. 14. Delay versus MRED for different approximate multipliers.

One of the DQ design has the lowest energy consumption at an MRED value of 0.37. The second highest accuracy design is ICM which has one of the highest energy consumptions. The YUS design has the highest energy consumption, while YUS_V2 has lower energy consumption compared to that of YUS. The YUS and YUS_V2 designs also have the advantage of accuracy configurability during the runtime. One of the other observations is that the MRED values of the X-Dadda structure spans over a wide range of MRED values. For instance, in T4W10, MRED spans from 0.005 to 0.6 (energy consumption of 75–154 fJ), which provides the user with different accuracy levels based on the energy budget and application accuracy requirement.

Comparison of the multipliers in terms of delay shows the lowest delays for YUS and YUS_V2 structures. The X-Dadda structure without truncation which keeps the critical paths of the exact multiplier has a higher delay. The delay for the truncated version of this structure is obviously lower than the case of without truncation. It should be noted that the delays for the X-Dadda structures were determined based on the exact mode and hence are not changed versus MRED. For fixed accuracy structures, lower MRED values are achieved based on less simplification/pruning and hence higher delays are expected. This would normally result in a pareto-optimal behavior.

Another design parameter of considerably less importance is the area. In the case of approximate multipliers based on circuit simplification/pruning, the area would be lower than that of the exact circuit, while in the case of approximate structures relying on VOS, the area will be slightly larger than

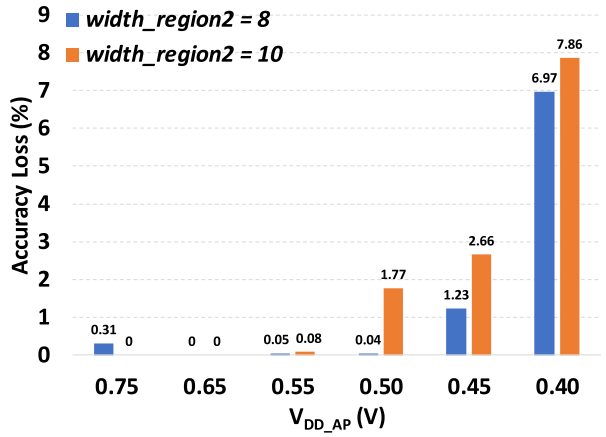

Fig. 15. Accuracy loss of the considered NN by using different X-Dadda structures as well as V_{DD_AP} values compared to case of the exact implementation.

that of the exact circuit due to the overhead associated with the power switches and level shifters.

Finally, it should be stated that since low-bit multipliers have a wide usage in image processing and machine learning applications, we focused on presenting accuracy-configurable 8-bit approximate multipliers. For the applications requiring multipliers with higher numbers of bits, similar comparative studies may be performed.

5) Using the X-Dadda Multiplier in Error Resilient Applications: Now, we evaluate the efficacy of the proposed X-Dadda multiplier for realizing neural networks (NNs) in the applications of classifying the MNIST data set, image sharpening, image smoothing, and discrete cosine transform (DCT).

a) Classifying the MNIST data set: The NN structure consisted of two hidden layers (with 100 neurons with the ReLU activation function in each layer) and trained offline by 60K images. Then, the NN was described by Verilog HDL language, while the inputs and the weights as well as the width of the multipliers were 8 bits.

While for the training of the NN, we used floating-point operations, in the hardware implementation used for the evaluation study, the inputs, weights, and biases were converted to 8-bit, 8-bit, and 16-bit integer numbers, respectively. The conversion caused some accuracy degradation for the network. Obviously, training the NN considering integer inputs, weights, biases as well as the proposed approximate multiplier will increase the accuracy of the NN.

Other operations (e.g., 24-bit addition in each neuron of the layers and the ReLU function of each neuron of the layers) in the hardware implementation of the NN were described in behavioral level (exact mode). The clock period of the design was determined based on the width__region2 parameter and the VOS voltage level considered for V_{DD_AP} (see Table II). The results obtained in this article showed a classification accuracy of 92.67% for the nominal voltage (exact multipliers).

The accuracy loss (compared to the exact implementation) versus V_{DD_AP}, when the X-Dadda structures (without truncation) with width__region2*s* of 8 and 10 bits were employed, are drawn in Fig. 15. For extracting the accuracy, 10K test

TABLE V

PSNR OF THE SHARPENING APPLICATION FOR DIFFERENT IMAGES AND APPROXIMATE MULTIPLIERS WITH TWO VOS VOLTAGE LEVELS

VOS Voltage Level	400mV					500mV				
width_region2	**4**	**6**	**8**	**10**	**12**	**4**	**6**	**8**	**10**	**12**
Lena	inf	41.53	33.75	33.34	29.55	inf	47.24	36.37	32.09	31.04
Baboon	inf	39.94	33.94	31.11	27.49	inf	45.22	36.80	33.18	30.78
Splash	inf	42.72	37.16	37.33	31.36	inf	47.80	41.60	36.86	35.35
Airplane	inf	43.56	31.34	32.60	30.95	inf	50.05	39.09	29.73	31.15
Pepper	inf	40.57	34.06	33.97	28.92	inf	45.77	36.62	33.25	32.07
Energy Reduction (%)	3%	10%	23%	36%	44%	2%	9%	20%	32%	36%

TABLE VI

PSNR OF THE SMOOTHING APPLICATION FOR DIFFERENT IMAGES AND APPROXIMATE MULTIPLIERS WITH TWO VOS VOLTAGE LEVELS

VOS Voltage Level	400mV					500mV				
width_region2	4	6	8	10	12	4	6	8	10	12
Lena	Inf	42.17	34.24	24.81	23.81	Inf	44.98	36.83	34.17	33.83
Baboon	Inf	39.39	31.35	21.05	20.09	Inf	41.01	35.69	32.36	31.00
Splash	Inf	44.84	36.70	29.45	28.17	Inf	50.98	40.83	36.33	38.43
Airplane	Inf	44.34	35.97	26.56	25.52	Inf	46.46	41.19	37.46	36.43
Pepper	Inf	41.33	33.92	25.64	24.40	Inf	42.41	36.53	33.54	33.15
Energy Reduction (%)	3%	10%	23%	36%	44%	2%	9%	20%	32%	36%

TABLE VII

PSNR OF THE DCT-IDCT APPLICATION FOR DIFFERENT IMAGES AND APPROXIMATE MULTIPLIERS WITH TWO VOS VOLTAGE LEVELS

VOS Voltage Level	400mV							
width_region2 (8-bit, 16-bit)	(2, 5)	(2, 9)	(4, 5)	(4, 9)	(6, 5)	(6, 9)	(6, 11)	(6, 13)	(6, 15)
Lena	49.84	49.83	49.84	47.13	49.79	42.04	38.29	34.75	30.18
Baboon	50.66	50.66	50.66	49.31	50.58	41.61	37.37	34.44	31.90
Splash	48.69	48.69	48.69	47.55	48.49	41.90	37.52	34.18	29.91
Airplane	49.34	49.34	49.34	44.33	49.31	40.82	37.88	33.39	27.83
Pepper	49.22	49.22	49.22	49.22	49.16	40.45	37.00	33.80	29.74
Energy Reduction (%)	1%	6%	1%	6%	2%	7%	12%	17%	23%
VOS Voltage Level	500mV							
Lena	49.83	49.83	49.83	49.83	49.83	49.83	49.09	48.01	45.45
Baboon	50.66	50.66	50.66	50.66	50.66	50.66	48.78	47.95	46.29
Splash	48.69	48.69	48.69	48.69	48.69	48.69	48.32	47.16	45.08
Airplane	49.34	49.34	49.34	49.34	49.34	49.34	48.98	47.84	45.25
Pepper	49.22	49.22	49.22	49.22	49.22	49.22	48.32	46.85	43.37
Energy Reduction (%)	1%	5%	1%	5%	2%	6%	9%	14%	19%

images were injected to the NN. The highest accuracy loss for both selected designs belonged to $V_{DD\_AP} = 0.40$ V (about 7.86%), while for most of the VOS voltage levels, the accuracy degradation was negligible or even zero. As an example, decreasing V_{DD_AP} from 0.80 to 0.50 V led to, on average, only 1.49% accuracy loss, while consuming lower energy for performing the classification.

b) Image sharpening: For the Sharpening application, each pixel can be extracted by [35]

$$Y(i,j) = 2X(i,j) - \frac{1}{273}\sum_{m=-2}^{2}\sum_{n=-2}^{2} X(i+m, j+n) \cdot \text{Mask}_{\text{Sharpening}}(m+3, n+3) \quad (4)$$

$$D = \begin{bmatrix} 91 & 91 & 91 & 91 & 91 & 91 & 91 & 91 \\ 126 & 106 & 71 & 25 & -25 & -71 & -106 & -126 \\ 118 & 49 & -49 & -118 & -118 & -49 & 49 & 118 \\ 106 & -25 & -126 & -71 & 71 & 126 & 25 & -106 \\ 91 & -91 & -91 & 91 & 91 & -91 & -91 & 91 \\ 71 & -126 & 25 & 106 & -106 & -25 & 126 & -71 \\ 49 & -118 & 118 & -49 & -49 & 118 & -118 & 49 \\ 25 & -71 & 106 & -126 & 126 & -106 & 71 & -25 \end{bmatrix} \quad (9)$$

where X and Y are the input and output images, respectively, and

$$\text{Mask}_{\text{Sharpening}} = \begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}. \quad (5)$$

The PSNR values and the energy saving for this application for five input images are reported in Table V (where *inf* corresponds to no changes in the image quality). The results show that by decreasing the VOS voltage level from 500 to 400 mV causes, on average, the PSNR (energy saving) decrease (increase) by 7% (15%) compared to that of the exact computation. Also, in the studied VOS voltage level, decreasing the width__region2 from 12 to 6 bits, on average, leads to an increase (decrease) of 30% (76%) in the PSNR (energy saving) compared to the case of the exact computation. In the case of width__region2 = 4, while the output quality does not decrease, an energy reduction of 2.5% is achieved.

c) Smoothing application: For the smoothing application, the relation between the output and input images are given by [35]

$$Y(i,j) = \frac{1}{60} = \sum_{m=-2}^{2}\sum_{n=-2}^{2} X(i+m, j+n) \cdot \text{Mask}_{\text{Smoothing}}(m+3, n+3) \quad (6)$$

where X and Y are the input and output images, respectively, and

$$\text{Mask}_{\text{Smoothing}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 4 & 4 & 1 \\ 1 & 4 & 12 & 4 & 1 \\ 1 & 4 & 4 & 4 & 1 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}. \quad (7)$$

The PSNR values and the energy reduction for the approximate smoothing application for five input images have been reported in Table VI. Similar to the Sharpening application, in the case of the width__region2 = 4, there was not any output quality loss, while about 2.5% energy reduction has been achieved. Also, based on the reported results, by decreasing the VOS voltage level from 500 to 400 mV, the PSNR (energy saving) is, on average, reduced (increased) by 18% (14%) compared to that of the exact computation. Also, for the studied VOS voltage levels, by decreasing the width__region2 from 12 to 6 bits, the PSNR (energy saving), on average, is increased (reduced) by 33% (76%).

d) DCT application: In this application, we approximate the two multiplication operations of the 8-point DCT expressed by [36]

$$\text{DCT} = D \times A \times D' \quad (8)$$

where A is a block of an input image and D is the DCT matrix (D' is the inverse of D). In our study, the values in the D matrix were transformed to integer values by multiplying them by 2^8. The transformed D matrix is (9), as shown at the top of this page.

For implementing this part, for the first (second) multiplication, an 8-bit (16-bit) approximate Dadda multiplier was used. To evaluate the output quality of the approximated DCT block, the results of DCT were applied to an inverse DCT (IDCT) block to retrieve the image. The PSNR values of the output (approximate) images as well as the energy reduction of the DCT part are reported in Table VII. The considered width__region2s of these two multipliers are indicated by the tuple of (x, y), where $x(y)$ denotes the width of the 8-bit (16-bit) proposed approximate Dadda multiplier. From the results, one can observe that decreasing the widths from (6, 15) to (2, 5), the output quality (energy saving) is increased (reduced), on average, about 24.3% (94%).

VI. Conclusion

In this article, we have explored the design of the accuracy-configurable approximate Dadda multiplier (X-Dadda) based on the VOS technique. The use of VOS provided lower energy consumption and higher lifetime. The X-Dadda multiplier had the flexibility of controlling the output accuracy during the runtime of the system by adjusting the operating voltage level of the approximate part and could change its operation mode to the exact mode. The error, energy consumption, and lifetime of the multiplier versus the approximation width and applied VOS voltage level were determined. The exploration results indicated that, for example, when the width of the approximate part was half of the width of the multiplier and its operating voltage was the half of the nominal voltage, the multiplier consumed 21% lower energy consumption at the cost of only 0.057 MRED. In addition, in this case, the delay increase due to the BTI effect was only 28% of the delay increase in the multiplier in the exact operating mode. The study also included the investigation of the impact of the PV on the error metric of the multiplier where the results did not indicate a considerable deterioration of the error. Finally, we determined the efficacy of the proposed X-Dadda structure in realizing NNs for the

applications of MNIST classification, image smoothing and sharpening, and DCT.

REFERENCES

[1] F. Nakhaee, M. Kamal, A. Afzali-Kusha, M. Pedram, S. M. Fakhraie, and H. Dorosti, "Lifetime improvement by exploiting aggressive voltage scaling during runtime of error-resilient applications," *Integration*, vol. 61, pp. 29–38, Mar. 2018.

[2] H. Afzali-Kusha, O. Akbari, M. Kamal, and M. Pedram, "Energy consumption and lifetime improvement of coarse-grained reconfigurable architectures targeting low-power error-tolerant applications," in *Proc. Great Lakes Symp. VLSI (GLSVLSI)*, 2018, pp. 431–434.

[3] S. Reda and M. Shafique, *Approximate Circuits—Methodologies and CAD*. Cham, Switzerland: Springer, 2019.

[4] H. Jiang *et al.*, "Characterizing approximate adders and multipliers optimized under different design constraints," in *Proc. Great Lakes Symp. VLSI (GLSVLSI)*, 2019, pp. 393–398.

[5] 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.

[6] M. S. Ansari, H. Jiang, B. F. Cockburn, and J. Han, "Low-power approximate multipliers using encoded partial products and approximate compressors," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 3, pp. 404–416, Sep. 2018.

[7] W. Liu, J. Xu, D. Wang, C. Wang, P. Montuschi, and F. Lombardi, "Design and evaluation of approximate logarithmic multipliers for low power error-tolerant applications," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 9, pp. 2856–2868, Sep. 2018.

[8] W. Liu, L. Qian, C. Wang, H. Jiang, J. Han, and F. Lombardi, "Design of approximate Radix-4 booth multipliers for error-tolerant computing," *IEEE Trans. Comput.*, vol. 66, no. 8, pp. 1435–1441, Aug. 2017.

[9] A. Gorantla and P. Deepa, "Design of approximate compressors for multiplication," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 13, no. 3, pp. 1–17, Apr. 2017.

[10] S. Venkatachalam and S.-B. Ko, "Design of power and area efficient approximate multipliers," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 5, pp. 1782–1786, May 2017.

[11] H. R. Mahdiani, A. Ahmadi, S. M. Fakhraie, and C. Lucas, "Bio-inspired imprecise computational blocks for efficient VLSI implementation of soft-computing applications," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 4, pp. 850–862, Apr. 2010.

[12] V. Mrazek, R. Hrbacek, Z. Vasicek, and L. Sekanina, "EvoApproxSb: Library of approximate adders and multipliers for circuit design and benchmarking of approximation methods," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2017, pp. 258–261.

[13] 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.

[14] 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.

[15] G. Zervakis, K. Tsoumanis, S. Xydis, D. Soudris, and K. Pekmestzi, "Design-efficient approximate multiplication circuits through partial product perforation," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 10, pp. 3105–3117, Oct. 2016.

[16] A. Gupta *et al.*, "Low power probabilistic floating point multiplier design," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, Jul. 2011, pp. 182–187.

[17] M. Lau, K. V. Ling, and Y. C. Chu, "Energy-aware probabilistic multipliers: Design and analysis," in *Proc. CASES*, Oct. 2009, pp. 281–290.

[18] 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.

[19] V. Mrazek, Z. Vasicek, and L. Sekanina, "Design of quality-configurable approximate multipliers suitable for dynamic environment," in *Proc. NASA/ESA Conf. Adapt. Hardw. Syst. (AHS)*, Aug. 2018, pp. 264–271.

[20] T. Yang, T. Ukezono, and T. Sato, "A low-power high-speed accuracy-controllable approximate multiplier design," in *Proc. 23rd Asia South Pacific Design Autom. Conf. (ASP-DAC)*, Jan. 2018, pp. 605–610.

[21] H. Baba, T. Yang, M. Inoue, K. Tajima, T. Ukezono, and T. Sato, "A low-power and small-area multiplier for accuracy-scalable approximate computing," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*, Jul. 2018, pp. 569–574.

[22] H. Jiang, C. Liu, F. Lombardi, and J. Han, "Low-power approximate unsigned multipliers with configurable error recovery," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 1, pp. 189–202, Jan. 2019.

[23] B. Bailey. (2018). *Chip Aging Becomes Design Problem*. Accessed: Aug. 14, 2019. [Online]. Available: https://semiengineering.com/chip-aging-becomes-design-problem/

[24] J. W. McPherson, *Reliability Physics and Engineering Time-to-Failure Modeling*, 2nd ed. Boston, MA, USA: Springer, 2013.

[25] E. Cai, D. Stamoulis, and D. Marculescu, "Exploring aging deceleration in FinFET-based multi-core systems," in *Proc. 35th Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2016, pp. 1–8.

[26] C.-H. Chang, J. Gu, and M. Zhang, "Ultra low-voltage low-power CMOS 4-2 and 5-2 compressors for fast arithmetic circuits," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 10, pp. 1985–1997, Oct. 2004.

[27] A. Shafaei, Y. Wang, L. Chen, S. Chen, and M. Pedram, "Maximizing the performance of NoC-based MPSoCs under total power and power density constraints," in *Proc. 17th Int. Symp. Qual. Electron. Design (ISQED)*, Mar. 2016, pp. 49–56.

[28] H. Afzali-Kusha, O. Akbari, M. Kamal, and M. Pedram, "Energy and reliability improvement of voltage-based, clustered, coarse-grain reconfigurable architectures by employing quality-aware mapping," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 3, pp. 480–493, Sep. 2018.

[29] H. Kaul, M. Anders, S. Hsu, A. Agarwal, R. Krishnamurthy, and S. Borkar, "Near-threshold voltage (NTV) design: Opportunities and challenges," in *Proc. 49th Annu. Design Autom. Conf. (DAC)*, 2012, pp. 1149–1154.

[30] A. Shapiro and E. G. Friedman, "Power efficient level shifter for 16 nm FinFET near threshold circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 2, pp. 774–778, Feb. 2016.

[31] S. Vahdat, M. Kamal, A. Afzali-Kusha, and M. Pedram, "TOSAM: An energy-efficient truncation- and rounding-based scalable approximate multiplier," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 5, pp. 1161–1173, May 2019.

[32] M. Martins *et al.*, "Open cell library in 15nm FreePDK technology," in *Proc. Symp. Int. Symp. Phys. Design (ISPD)*, 2015, pp. 171–178.

[33] P. Pandey, P. Basu, K. Chakraborty, and S. Roy, "GreenTPU: Improving timing error resilience of a near-threshold tensor processing unit," in *Proc. 56th Annu. Design Autom. Conf. (DAC)*, 2019, pp. 1–6.

[34] M. Ansari, H. Afzali-Kusha, B. Ebrahimi, Z. Navabi, A. Afzali-Kusha, and M. Pedram, "A near-threshold 7T SRAM cell with high write and read margins and low write time for sub-20 nm FinFET technologies," *Integration*, vol. 50, pp. 91–106, Jun. 2015.

[35] H. R. Myler and A. R. Weeks, *The Pocket Handbook of Image Processing Algorithms*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2009.

[36] R. J. Cintra, F. M. Bayer, and C. J. Tablada, "Low-complexity 8-point DCT approximations based on integer functions," *Signal Process.*, vol. 99, pp. 201–214, Jun. 2014.