

Maximum Delay Models for Parallel-Prefix Adders in the Presence of Threshold Voltage Variations

Kleanthis Papachatzopoulos and Vassilis Paliouras

Electrical and Computer Engineering Department

University of Patras, Patras, Greece

Email: papachatz@ece.upatras.gr; paliouras@ece.upatras.gr

Abstract—This paper introduces a delay modeling formulation for several Parallel-Prefix Adders in the presence of threshold voltage variability. A path-based model is derived for the delay variability of Kogge-Stone, Knowles, Sklansky, Brent-Kung, Han-Carlson, Ladner-Fischer, and New Adder architectures. The delay model accuracy is evaluated for the specific adders on the basis of SPICE Monte-Carlo Simulations at 45 nm and 16 nm nodes. The presented analysis reveals that the proposed path-based model estimates the maximum delay Probability Density Function of the particular adder architectures with sufficient accuracy, assuming 3σ intra-die threshold voltage variations as high as 10% of nominal value. Delay yield estimations produced by the proposed model are found to agree with those of Monte-Carlo Simulations for a number of highly probable critical paths, presenting an error less than 2%. For the particular adders and technology nodes, an approximately 10-fold reduction in simulation time is obtained when exploiting the proposed model. The particular observation indicates that the computational time for delay yield estimation of Parallel-Prefix Adders can be exponentially reduced with negligible accuracy loss when the analysis focuses solely on the Nominal-Maximum Delay critical path. Finally, a quantitative comparison of prefix adders to the Borrow-Save Adder is offered, in terms of complexity and susceptibility to variations.

Index Terms—parallel-prefix adders, critical path, threshold voltage variations, variability, delay yield

1. Introduction

Under the persistent scaling of integration in nanometer regime, process parameter variations have emerged as a major challenge in both design and integration flow [1], [2]. Process imperfections affect cell delays, leading to performance and power degradation. Furthermore, the need of satisfying excessively tight specifications combined with the timing degradation leads to greater guard-bands as integration scales, delaying design closure [3]. Hence, consideration of variation mechanisms during timing analysis reduces chip failure rates and time to market.

Robustness of timing analysis cannot be ensured by a corner-based analysis, as it cannot capture the prominent-in-denser-nodes intra-die variations [4]. Hence, a shift in

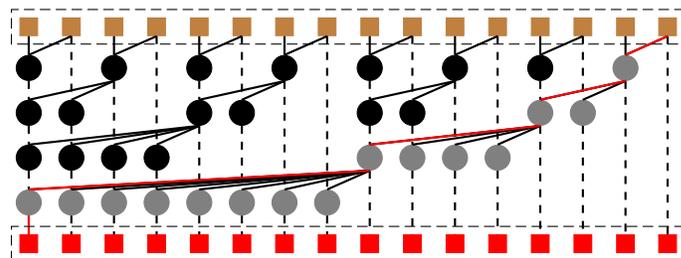


Figure 1. Propagate-Generate Network of Sklansky adder. Brown nodes refer to first-stage propagate-generate cells, black to group propagate-generate cells, gray to group generate cells, and red to XOR cells. Red line connects the nodes forming the NMD critical path.

statistical-based methodologies, such as *Monte-Carlo* and *Statistical Static Timing Analysis*, provides the generalization of traditional techniques in a variation-prone framework. The aforementioned timing analysis techniques handle delays as Random Variables (RVs) and Probability Density Functions (PDFs), respectively. Nowadays, research efforts focus on Statistical Static Timing Analysis (SSTAs) using *non-Gaussian* models [5], [6] and dealing with correlated RVs [7]. Regarding the statistical nature of delay variables, lognormal models are suitable for the approximation of cell delays in sub-threshold regime and present an error of 11% and 16% for the mean and standard deviation of delay, respectively, compared to Monte-Carlo Simulations [8]. Additionally, it is shown that a unified Log-Skew Normal model for the cell delays and for operation stages ranging from sub-threshold to super-threshold can accurately predict the delay yield compared to Monte-Carlo Simulations [9]. However, beyond 40 nm, OCV and AOCV are the mainstream signoff approaches for EDA tools, with the latter providing an accuracy compared to that of SSTA with much less computational effort [10], as well, *Liberty Variation Format*-based timing analysis that is able to model the non-Gaussian timing behavior of path delays [3]. For early stage delay modeling analysis, Alioto *et al.* [11] present a framework for the estimation of path delay variability that relies on fan-out-of-4 metric and on technology library characteristics.

Optimizations related to arithmetic circuits, and, especially, to adder structures, have always been of great

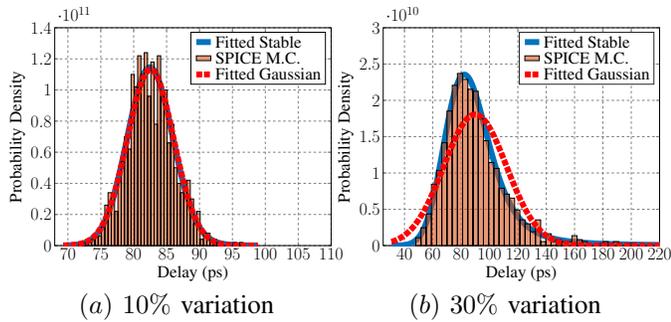


Figure 2. Probability Density Functions of the sum delay of a Mirror FA cell in the presence of (a) 10% and (b) 30% intra-die threshold voltage variability.

importance, as they constitute fundamental blocks of digital processing units. Furthermore, they imply timing and power insights for more complex structures as well. Given the process integration variability and atomic-scale uncertainty in nanometer integration nodes, evaluation of delay variability guides necessary design countermeasures. In this context, Alioto and Palumbo propose a model that evaluates the sensitivity of Full-Adder cells to supply voltage variations [12]. Bernstein *et al.* quantify the delay variability of selected 16-bit adders under certain variation mechanisms [13]. Furthermore, Papachatzopoulos and Paliouras evaluate the variability of Ripple-Carry Adders and Borrow-Save Adders under threshold voltage (V_{th}) variations, and propose two interconnection transformations for the Borrow-Save Adder that reduce delay variability [14]. Also, they introduce a delay model, suitable for SSTA, for the estimation of maximum delay PDF for Ripple-Carry and Borrow-Save Adders under delay variations [15]. Furthermore, the advantages of Residue Number System (RNS)-based Multiplier-Accumulators (MACs) for variation-tolerant architectures have been revealed compared to their binary counterparts [16]. Clearly, the selection of an adder architecture can significantly affect the delay variability characteristics of the overall processing system.

Parallel-prefix adders are of great interest as they are a common choice for fast addition. This paper proposes a delay model for the radix-2 Kogge-Stone, Knowles [2, 1, 1, 1], Sklansky, Brent-Kung, Han-Carlson, Ladner-Fischer and New (1, 1, 1) adders under V_{th} variations at 45-nm and 16-nm technology nodes. The proposed model assumes a Gaussian distribution, as Stable and Log-Normal distributions are not found to substantially improve fitting accuracy for values of intra-die V_{th} variability as much as 10% in the particular context. Fitting parameters of the proposed model are derived using SPICE Monte-Carlo (M.C.) Simulations of the Nominal Maximum-Delay critical path of the Parallel-Prefix Adders under study.

The main contribution of this paper is that it demonstrates that a relatively simple model can accurately describe the maximum delay behavior under a certain variability. Furthermore and in the process of model derivation, several Parallel-Prefix Adders are evaluated and compared. In addition,

TABLE 1. TRANSISTOR COUNT COMPLEXITY OF TARGETED ADDERS

| Architecture | #Transistors | Architecture | #Transistors |
|--------------|--------------|----------------|--------------|
| Knowles | 1076 | Kogge-Stone | 1076 |
| Sklansky | 838 | Brent-Kung | 754 |
| Han-Carlson | 838 | Ladner-Fischer | 768 |
| New | 838 | Borrow Save | 896 |

tion, the Borrow-Save Adder (BSA) is also considered, as it has been recently reported to perform well under variation. It is found that under the given assumptions, several trade-offs can be exploited in terms of complexity and susceptibility to variations.

The remainder of the paper is structured as follows. Section 2 introduces the proposed path-based modeling formulation for the Parallel-Prefix Adders, while Section 3 assesses the fitting accuracy of the proposed path-based PDFs compared to SPICE-level Monte-Carlo Simulations. Finally, Section 4 discusses conclusions.

2. Proposed Path-Based Delay Model

The proposed formulation provides estimations for the maximum delay distribution of a set of radix-2 Parallel-Prefix Adders using statistical estimations of the Nominal Maximum-Delay (NMD) critical path. The model targets delay variations that origin from atomic-level deviations of dopant atoms, commonly known as Random Dopant Fluctuations, whose impact can be abstracted in circuit level as threshold voltage deviations [13]. Interest is focused on threshold voltage variations as this is the leading contributor of variability regarding timing in several adder architectures [13]. Furthermore, variations manifested in channel length can be mapped to threshold voltage deviations as explained by the DIBL effect [17]. However, channel length-related variations take place systematically in a technology process and their effect is predictable [1]. Less major mechanisms, such as mobility and gate oxide thickness variations, impact on threshold voltage as well.

Parallel-Prefix Adders consist of a network of group propagate and group generate cells that produce the related propagate and generate signals for a set of bit positions. Indicatively, Fig. 1 depicts the structure of Propagate-Generate (PG) network for a 16-bit Sklansky adder. The performance of Parallel-Prefix Adders is of particular interest as it scales proportionally to $\log_2 N + c$, where N is the adder bit length and c is a constant dependent on the particular architecture [18]. Particularly, adders investigated in this paper are Knowles [2, 1, 1, 1] [19], Kogge-Stone [20], Sklansky [21], Brent-Kung [22], Han-Carlson [23], Ladner-Fischer [24], and New (1, 1, 1) [25] architectures. Adder architectures are sized for minimum area and using low-power transistors, *i.e.*, with relatively high threshold voltage in order to substantially reduce leakage power dissipation. Indicative complexity of the adders under study, quantified in terms of transistor count, are reported in Table 1 assuming 16-bit implementations. The reported figures refer to static CMOS implementations for the constituent cells.

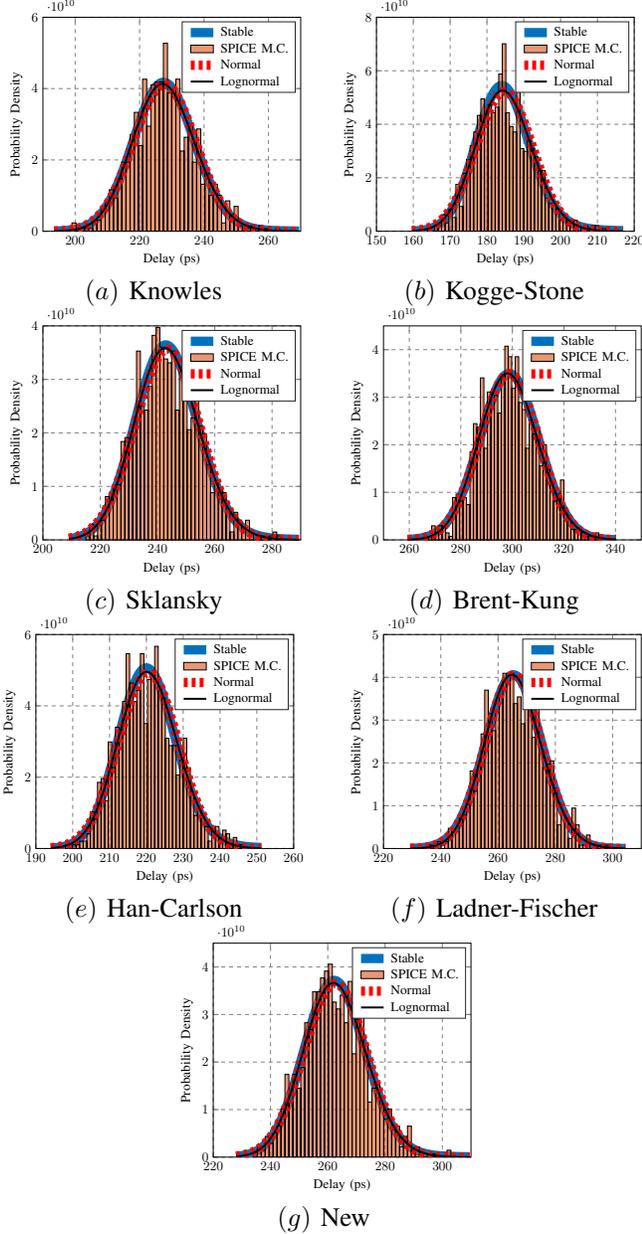


Figure 3. Monte-Carlo histograms and fitted Probability Density Functions for 16-bit (a) Knowles, (b) Kogge-Stone, (c) Sklansky, (d) Brent-Kung, (e) Han-Carlson, (f) Ladner-Fischer, and (g) New adders.

2.1. Modeling Approaches

Commonly, statistical methods assume that the underlying variations follow a Gaussian distribution; however, a Stable distribution may more accurately model delay variation in certain cases. As an illustrative example, Fig. 2 shows the delay PDF observed at the sum output of a Mirror Full-Adder (FA) cell under intra-die V_{th} variations and for a specific input transition pattern. The reported (a) and (b) cases in Fig. 2 correspond to Monte-Carlo Simulations with 3σ threshold voltage variations equal to 10% and 30% of the nominal threshold voltage value, respectively. It is shown

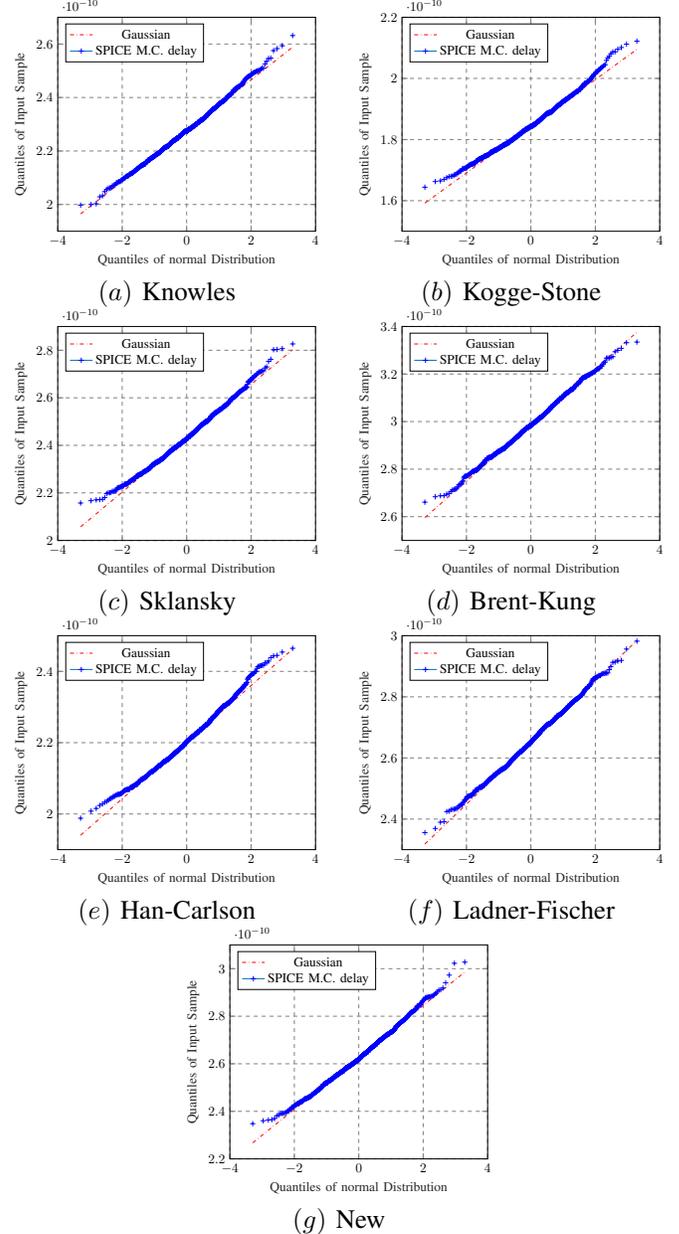


Figure 4. QQ plots for Gaussian distribution and M.C. SPICE delays for (a) Knowles, (b) Kogge-Stone, (c) Sklansky, (d) Brent-Kung, (e) Han-Carlson, (f) Ladner-Fischer, and (g) New adders.

that, as the magnitude of variations increases, FA delay is restricted by a minimum delay threshold, and, hence, creates a positive skewed PDF, *i.e.*, with a heavier tail in the right side of the distribution. Fig. 2 reveals that a Stable distribution describes more accurately the delay of the FA sum output than the fitted Gaussian distribution. In the context of threshold voltage variations not exceeding 10% of their nominal value and, furthermore, for sufficient long logic depth, as that of the examined architectures, Gaussian modeling of delay distribution provides sufficient accuracy as shown in Section 3.

TABLE 2. SPICE DELAY ESTIMATORS UNDER V_{th} VARIATIONS @ 45 nm

| Architecture | M.C. SPICE | | | |
|----------------|------------|---------------|----------------------|--------------|
| | μ (ns) | σ (ps) | $\mu + 3\sigma$ (ns) | σ/μ |
| Knowles | 0.4189 | 9.6303 | 0.4478 | 0.0229 |
| Kogge-Stone | 0.3335 | 7.7172 | 0.3566 | 0.0231 |
| Sklansky | 0.4481 | 10.3807 | 0.4792 | 0.0231 |
| Brent-Kung | 0.5450 | 10.6058 | 0.5768 | 0.0194 |
| Han-Carlson | 0.3950 | 8.4109 | 0.4202 | 0.0212 |
| Ladner-Fischer | 0.4802 | 9.7081 | 0.5094 | 0.0202 |
| New | 0.4795 | 10.3447 | 0.5105 | 0.0215 |
| Borrow-Save | 0.1614 | 4.6073 | 0.1752 | 0.0285 |

TABLE 3. SPICE DELAY ESTIMATORS UNDER V_{th} VARIATIONS @ 16 nm

| Architecture | M.C. SPICE | | | |
|----------------|------------|---------------|----------------------|--------------|
| | μ (ns) | σ (ps) | $\mu + 3\sigma$ (ns) | σ/μ |
| Knowles | 0.2774 | 9.6893 | 0.2568 | 0.04254 |
| Kogge-Stone | 0.1847 | 7.6339 | 0.2076 | 0.04132 |
| Sklansky | 0.2434 | 11.1929 | 0.2770 | 0.04598 |
| Brent-Kung | 0.2987 | 11.4040 | 0.3329 | 0.03817 |
| Han-Carlson | 0.2205 | 8.0932 | 0.2448 | 0.03669 |
| Ladner-Fischer | 0.2653 | 9.8294 | 0.2948 | 0.03703 |
| New | 0.2626 | 10.9370 | 0.2954 | 0.04163 |
| Borrow-Save | 0.1221 | 5.3839 | 0.1383 | 0.04407 |

Given the need for sufficiently accurate modeling of maximum-delay PDFs, SPICE-level Monte-Carlo Simulations are performed for the examined adder architectures under intra-die V_{th} variations at 45-nm (1.1 V nominal voltage) and 16-nm (0.9 V nominal voltage) nodes. In the following analysis, values of threshold voltage are assumed to be independent and identically distributed, following a Gaussian distribution with $3\sigma_{V_{th}} = 0.1V_{th,nom}$ unless otherwise stated. Fig. 3 depicts the histograms of the adder delays for a sufficiently large number of paths, obtained by SPICE simulations under V_{th} variation, along with the fitted Gaussian, Stable and Log-Normal PDFs at a 16-nm node. Delays reported correspond to 16-bit architectures and 1000 simulation iterations for the remainder of this paper. The methodology of path selection for the Monte-Carlo Simulations is detailed in Section 3. As Fig. 3 indicates, Stable and Log-Normal distributions provide almost the same fitting accuracy as the Gaussian; hence, adoption of Gaussian distribution for the modeling of maximum delay behavior provides sufficient accuracy and ease modeling formulations for the particular circuits and in the particular variability context. The choice of Gaussian distribution is further justified by the Quantile-Quantile (QQ) plots of Fig. 4 that evaluate the achieved Gaussian distribution fitting on the SPICE maximum-delay simulations under variations.

2.2. Proposed Delay Model

In the following we propose a simple delay model for the adders under study that describes the PDF of the maximum delay. As intra-die variation mechanisms are expected to dominate over inter-die variations in technologies of a few nanometers, the proposed model focuses on them.

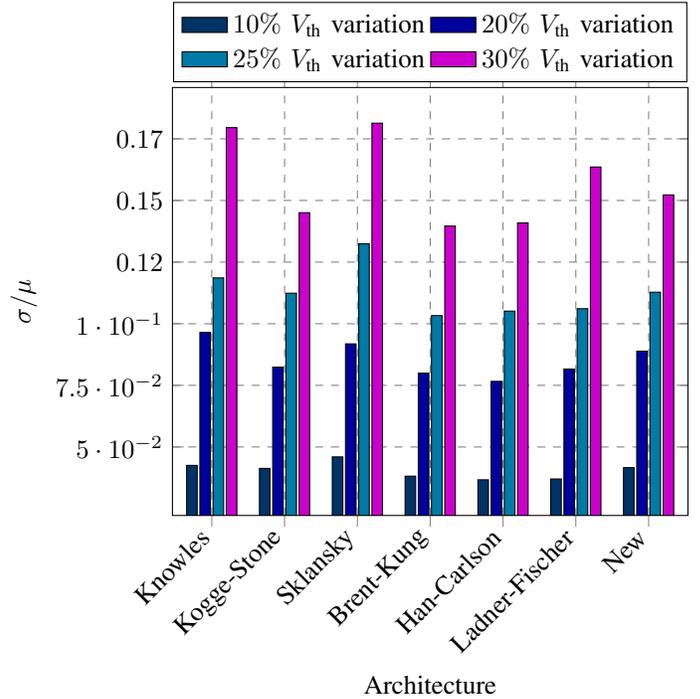


Figure 5. Variability (σ/μ) for certain percentages of threshold voltage variation at 16 nm.

TABLE 4. PERCENTAGE FREQUENCY REDUCTION CAUSED BY THRESHOLD VOLTAGE VARIABILITY

| Architecture | 16 nm | 45 nm |
|----------------|---------|--------|
| Knowles | 11.3187 | 6.4507 |
| Kogge-Stone | 11.0296 | 6.4913 |
| Sklansky | 12.1219 | 6.4976 |
| Brent-Kung | 10.2750 | 5.5160 |
| Han-Carlson | 9.9175 | 6.0042 |
| Ladner-Fischer | 9.9998 | 5.7173 |
| New | 11.1043 | 6.0783 |
| Borrow-Save | 11.6787 | 7.8892 |

TABLE 5. KULLBACK-LEIBLER DIVERGENCE FOR PROPOSED GAUSSIAN AND FITTED GAUSSIAN/STABLE PDFS

| Architecture | Proposed-Fitted Gaussian | | Proposed-Fitted Stable | |
|----------------|--------------------------|--------|------------------------|--------|
| | 16 nm | 45 nm | 16 nm | 45 nm |
| Knowles | 0.0011 | 0.0013 | 0.0034 | 0.0013 |
| Kogge-Stone | 0.0537 | 0.1570 | 0.0651 | 0.1611 |
| Sklansky | 0.0359 | 0.0560 | 0.0374 | 0.0574 |
| Brent-Kung | 0.0004 | 0.0071 | 0.0004 | 0.0108 |
| Han-Carlson | 0.0616 | 0.0010 | 0.0655 | 0.0020 |
| Ladner-Fischer | 0.0324 | 0.0131 | 0.0325 | 0.0135 |
| New | 0.0162 | 0.0030 | 0.0182 | 0.0030 |

TABLE 6. DELAY YIELD ESTIMATION DETERMINED BY $\mu + 3\sigma$

| Architecture | 16 nm | | | 45 nm | | |
|----------------|-----------------|---------------------|-----------|-----------------|---------------------|-----------|
| | M.C. SPICE (ns) | Proposed Model (ns) | Error (%) | M.C. SPICE (ns) | Proposed Model (ns) | Error (%) |
| Knowles | 0.2568 | 0.2570 | -0.0941 | 0.4478 | 0.4474 | 0.089 08 |
| Kogge-Stone | 0.2076 | 0.2089 | -0.6329 | 0.3566 | 0.3492 | 2.086 08 |
| Sklansky | 0.2770 | 0.2733 | 1.3246 | 0.4792 | 0.4794 | -0.043 77 |
| Brent-Kung | 0.3329 | 0.3324 | 0.1496 | 0.5768 | 0.5778 | -0.175 46 |
| Han-Carlson | 0.2448 | 0.2443 | 0.1852 | 0.4202 | 0.4210 | -0.184 89 |
| Ladner-Fischer | 0.2948 | 0.2950 | -0.0667 | 0.5094 | 0.5104 | -0.201 80 |
| New | 0.2954 | 0.2921 | 1.1324 | 0.5105 | 0.5096 | 0.176 98 |

TABLE 7. MEAN COMPUTATIONAL TIME OF M.C. SIMULATIONS FOR THE CONVENTIONAL AND PROPOSED METHODS

| Method | Conventional | Proposed |
|----------|--------------|----------|
| Run Time | 49 h | 4.5 h |

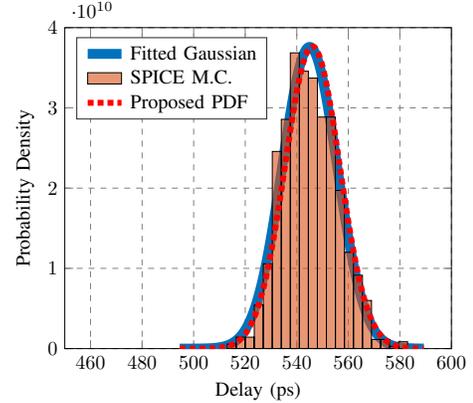
Moreover, even if individual cell delays are described by delay distributions other than a Gaussian, the path delays are assumed to follow a Gaussian distribution for sufficiently long paths. The proposed model is a path-based model as it relies on the NMD critical path of each adder architecture and on the extraction of statistical delay moments from the simulations of only one path. As a first step, Monte-Carlo Simulations under V_{th} variations are performed only for the NMD critical path for each adder. A Gaussian distribution is used to approximate the maximum delay PDF, using the mean delay and standard deviation obtained via the Monte-Carlo Simulations of NMD critical path. The mean delay of NMD critical path under the assumed variation mechanism is expected to equal the nominal delay of the respective path, hence, it can be approximated as

$$t_{NMD} = t_{PG} + \sum_{i=1}^L t_{group\ PG,i}(F, WT) + t_{XOR}, \quad (1)$$

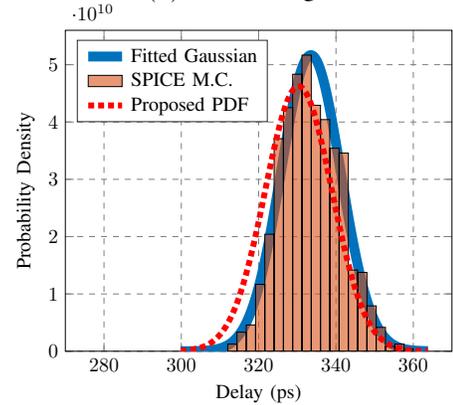
where t_{NMD} is the delay of the NMD critical path, t_{PG} is the delay of first-stage PG cells, $t_{group\ PG,i}$ is the delay of group PG cells of i -th logic level, and t_{XOR} is the delay of XOR gate used to compute sum. Furthermore, propagation delays of a PG network is a function dependent on the number of logic levels of the PG network, L , Fanout, F , and Wiring Tracks, WT [25]. Similarly, standard deviation of the NMD critical path delay is estimated as

$$\sigma_{NMD} = \sqrt{\sigma_{PG}^2 + \sum_{i=1}^L \sigma_{group\ PG,i}(F, WT)^2 + \sigma_{XOR}^2}, \quad (2)$$

assuming negligible correlation between variations manifested in each cell thus permitting the expression of path delay variance as the sum of variances of cells along the NMD critical path. Cell nominal delays and standard deviations can be characterized once per technology node and variation mechanism, and, thus, used for the delay estimation of more elaborate paths. The proposed modeling methodology assumes that the maximum delay PDF can be approximated



(a) Brent-Kung



(b) Kogge-Stone

Figure 6. Monte-Carlo histograms, Fitted Gaussian and Proposed Probability Density Functions for (a) Brent-Kung and (b) Kogge-Stone at 45 nm.

as $N(\mu_{NMD}, \sigma_{NMD})$, where $N(\mu, \sigma)$ denotes a normal distribution with mean value μ and standard deviation σ .

3. Evaluation of Proposed Delay Model

3.1. Experimental Framework

As variations are manifested in a no fully correlated way for the case of intra-die variations, every path has a probability of becoming the maximum-delay critical path. Thus, in a circuit subjected to delay variations, an input pattern which sensitizes the NMD critical path, does not

necessarily sensitize the maximum delay of the particular circuit instance. For this reason a pre-processing step is required in order to identify input patterns that sensitize a set of paths that are likely to become critical. Resembling most Monte-Carlo methods [1], identification of input patterns is performed under nominal conditions. Our analysis involves such input patterns, able to sensitize paths whose nominal delay is as much as 75% of that of an NMD critical path, *i.e.*, a path with a propagation delay from the LSB to the MSB is sensitized. Thus, we assume that the maximum-delay PDF of the adders under study, can be adequately described by utilizing the aforementioned sensitization scenario in order to restrict the computational complexity of Monte-Carlo Simulations.

3.2. Proposed Path-Based Model and Analysis

In the following analysis, the $\mu + 3\sigma$ delay is used as a measure of worst-case delay and, thus, determines the delay yield for the particular adders. Tables 2 and 3 present the delay estimators for the examined adder architectures at 45-nm and 16-nm nodes, respectively. At both technology nodes, the Kogge-Stone architecture demonstrates the best delay performance in terms of $\mu + 3\sigma$. On the contrary, Brent-Kung adder has the worst performance. These delay figures correspond to about 3.004 GHz and 1.733 GHz datapath frequency for Brent-Kung adder at 16-nm and 45-nm nodes, respectively. As for the fastest parallel-prefix adder, Kogge-Stone presents a 4.817 GHz and a 2.804 GHz datapath frequency at 16-nm and 45-nm node, respectively, while the difference between the second-fastest adder (Han-Carlson) ranges from 37 ps to 63 ps.

Comparing Tables 2 and 3, mean delays indicate a delay decrease of almost 1.8 for all adder cases when moving from 45-nm to 16-nm technology node. Sklansky adder is found to exhibit the greatest delay variability introduced by the variations of threshold voltage, quantified by the σ/μ ratio, at both 16-nm and 45-nm technology nodes. On the other hand, a Han-Carlson adder demonstrates the smallest variability at 16 nm and Brent-Kung at 45 nm among the studied parallel-prefix adders. Values of σ/μ for greater variation magnitudes at 16 nm reported in Fig. 5 reveal a similar trend. The variability metric is primarily determined by the values of mean maximum delay, as inherent averaging mechanisms are more effective in circuits with longest paths than in those with shorter ones, and, consequently, variability is smaller for larger paths.

Delay performances of the particular Parallel-Prefix Adders are also compared against the ultra-fast architecture of Borrow-Save Adder build using Mirror Full-Adder cells. BSA demonstrates attractive characteristics for variation-tolerant designs as critical path length is independent of word length and its worst-case delay increases marginally under delay variations [15]. The performance of Borrow-Save Adder outperforms the fastest adder (Kogge-Stone) by 33.38% and 50.87% at 16 nm and 45 nm, respectively, in the presence of an identical variation mechanism.

Table 4 reports the datapath operation frequency reduction, as determined by the ratio $3\sigma/(\mu+3\sigma)$, for each adder and technology node caused by only V_{th} variability. Values of percentage frequency reduction are consistent with the values of delay variability reported in Tables 2 and 3, and show that a Sklansky adder undergoes the greatest frequency degradation among the parallel-prefix adders under study, while Han-Carlson (16 nm) and Brent-Kung (45 nm) the smallest, as indicated also by the σ/μ metric. Borrow-Save Adder presents the greatest frequency reduction at 45 nm and the second greatest at 16 nm among all adders, which can be attributed to its small logic depth that renders the particular adder sensitive to variations.

The accuracy of the proposed model with respect to simulation data is quantified in Table 5. Table 5 shows the Kullback-Leibler Divergence (KLD) between the proposed path-based delay model and a Gaussian Delay Model, fitted to delays extracted from Monte-Carlo Simulations considering a number of near maximum-delay critical paths, as described in Section 3.1. Close-to-zero values of the aforementioned metric indicate a good fitness accuracy. It is revealed that the best accuracy is achieved for the Brent-Kung and Knowles adders for 16-nm and 45-nm nodes, respectively, as they present the minimum KLD. On the contrary, the worst fitting appears for the Han-Carlson adder at 16-nm node.

It is investigated whether a Stable distribution can approximate more accurately the skewed behavior of maximum delay PDF in the general case and the accuracy of proposed model is also assessed compared to Fitted Stable distribution using the same metric. The KLD values reported in Table 5 for the Stable distribution are similar as those achieved by the Gaussian. Furthermore, certain values of KLD for Stable and the Proposed distribution show a marginally better fitting accuracy than that of Gaussian and the Proposed one, such as in the case of Brent-Kung adder at 16 nm.

The delay yield has been assessed by the proposed model and Monte-Carlo Simulations, and the related values are reported in Table 6 for both technology nodes. Fig. 6 shows the Proposed PDF for Brent-Kung and Kogge-Stone adders at 45 nm. As expected, the proposed model demonstrates better accuracy at the right tail of the PDF compared with to Monte-Carlo Simulations. Accuracy in the right tail is particularly important as it refers to the higher values that may be assumed by the delay under variations that determine the worst-case performance.

In most cases, yield values of proposed model agree with that of Monte-Carlo Simulations in the first two digits. Error values indicate that in most cases the yield is underestimated by the proposed model and the error is less than 2.1% indicating sufficient agreement between the proposed model and experimental SPICE-level evaluation. Thus, delay yield estimations for Parallel-Prefix Adders can be accurately performed based on NMD critical path with relatively minimum accuracy loss compared to the case that considers a set of potentially critical paths.

Regarding the computational overhead of simulations,

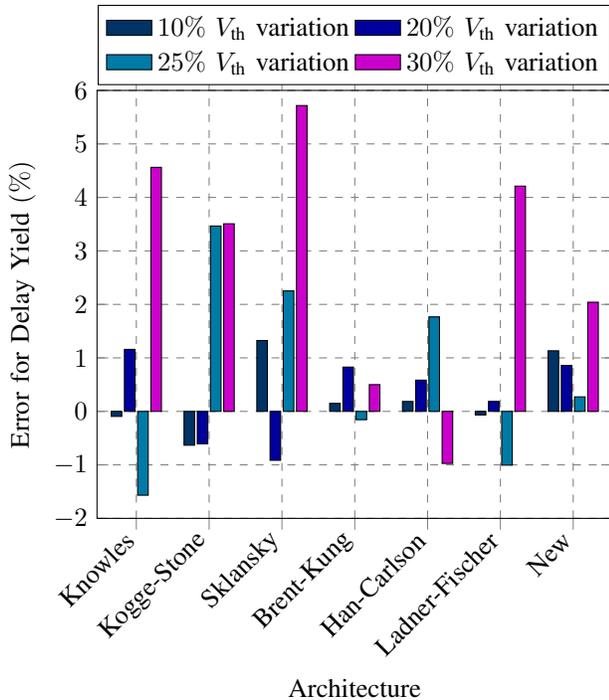


Figure 7. Delay yield error for certain percentages of threshold voltage variation.

Table 7 presents the mean computational time of Monte-Carlo Simulations for a set of highly probable critical paths (Conventional) and that of proposed model for 1000 iterations. A 10.8-times shorter simulation time for the delay yield estimation of the examined Parallel-Prefix Adders is achieved. This means that the simulation run time for the characterization of delay yield for the examined architectures can be significantly reduced based on the analysis of only NMD critical path, given that the proposed model introduces an error that is lower than 2%.

To further investigate the accuracy of the proposed model for increasing values of the examined variation mechanism, delay yield estimations are assessed for 16-bit adders at 16 nm and introduced variations as large as 30% of nominal delay. Specifically, Fig. 7 demonstrates the delay yield error between the proposed model and the respective M.C. simulations considering a number of critical paths. The absolute error is lower than 6% in all the examined cases, and it is the greatest at 30% V_{th} variations, as the probability of observing a non-critical path as the maximum-delay critical increases compared to smaller percentage variations. In all investigated cases, model still provides acceptable error for an early-stage delay exploration for a range of variation magnitudes.

Given the relatively large error introduced for the 30% variation case, it is investigated whether the error is reduced when a few more potentially critical paths, selected under nominal circumstances, are considered for the model derivation. Indicatively, Fig. 8 presents the positive skewed shape that the maximum delay PDF acquires for such a

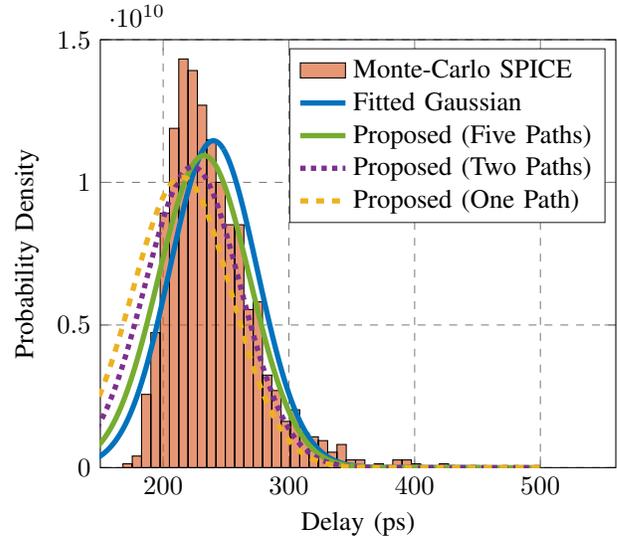


Figure 8. Positive-skewed PDF of Kogge-Stone Adder for 30% nominal value V_{th} fluctuations and proposed models.

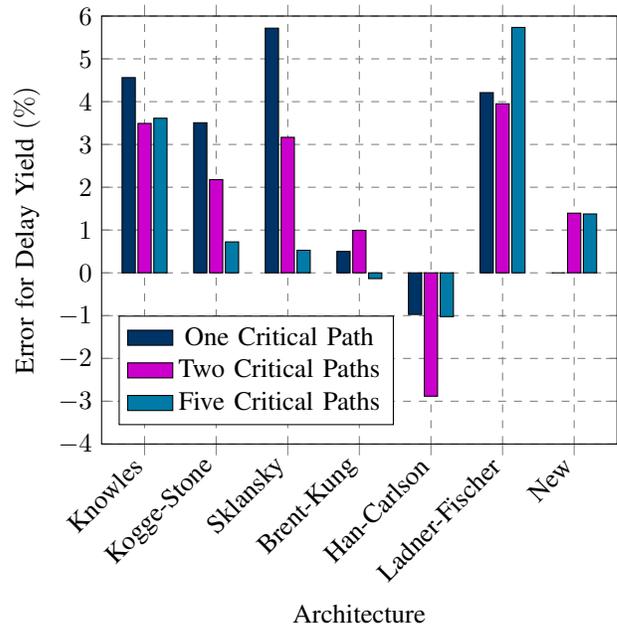


Figure 9. Delay yield error when considering more than one critical paths for the proposed model.

large variation. The analysis further considers the first-two, three, or five NMD critical paths. It is noted that the mean delay and standard deviation of the Gaussian distribution referring to the max of first-two NMD critical paths can be analytically computed by Clark's expressions [26], given the first-two moments of the related paths. In more detail, Fig. 9 presents the errors for the delay yield between the Monte-Carlo SPICE simulations for a number of critical paths and the case of two, three and five NMD critical paths. Fig. 9 reveals that taking also into account the effect of a larger

number of critical paths for the model derivation and delay yield prediction further improves the accuracy of estimation for certain cases. For the remainder cases where the error is not reduced, either small paths may determine delay yield as more sensitive to variations, and more elaborate path selection techniques should be investigated to improve the path-based model accuracy, or the additional paths add extra pessimism (case of New adder). Thus, a tradeoff between the paths considered (computational complexity) and the accuracy of delay yield prediction can be exploited based on the predicted magnitude of introduced variations.

4. Conclusions

In this paper, a path-based model for the PDF of maximum delay of certain Parallel-Prefix Adders is investigated. The proposed model uses fitting parameters extracted from Monte-Carlo Simulations that consider only the NMD critical path of examined adders under intra-die threshold voltage variability. The proposed path-based model shows significant similarity compared to the fitted distributions on Monte-Carlo Simulations that take into account a number of critical paths, as quantified using Kullback-Leibler Divergence. Furthermore, the delay yield as estimated by the proposed approach is compared with that of Monte-Carlo Simulations and presents an error that is lower than 2% for the 10% nominal variation case, leading, furthermore, to a significant run time reduction for M.C. Simulations.

References

- [1] M. Orshansky, S. Nassif, and D. Boning, *Design for Manufacturability and Statistical Design: A Constructive Approach*. New York, NY, USA: Springer, 2007.
- [2] A. Srivastava, D. Sylvester, and D. Blaauw, *Statistical Analysis and Optimization for VLSI: Timing and Power*. New York, NY, USA: Springer, 2006.
- [3] A. B. Kahng, "New Game, New Goal Posts: A Recent History of Timing Closure," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*. IEEE, 2015, pp. 1–6.
- [4] S. S. Sapatnekar, "Overcoming Variations in Nanometer-Scale Technologies," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 1, pp. 5–18, 2011.
- [5] S. Ramprasath, M. Vijaykumar, and V. Vasudevan, "A Skew-Normal Canonical Model for Statistical Static Timing Analysis," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 6, pp. 2359–2368, 2015.
- [6] J. Singh and S. Sapatnekar, "Statistical Timing Analysis with Correlated non-Gaussian Parameters using Independent Component Analysis," in *2006 43rd ACM/IEEE Design Automation Conference*. IEEE, 2006, pp. 155–160.
- [7] D. Mishaghi, E. Koskin, and E. Blokhina, "Path-Based Statistical Static Timing Analysis for Large Integrated Circuits in a Weak Correlation Approximation," in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2019, pp. 1–5.
- [8] F. Frustaci, P. Corsonello, and S. Perri, "Analytical Delay Model Considering Variability Effects in Subthreshold Domain," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 3, pp. 168–172, 2012.

- [9] H. A. Balef, M. Kamal, A. Afzali-Kusha, and M. Pedram, "All-Region Statistical Model for Delay Variation Based on Log-Skew-Normal Distribution," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 9, pp. 1503–1508, 2015.
- [10] R. Goering. Signoff summit: An Update on OCV, AOCV, SOCV, and Statistical Timing. Cadence Blog. Accessed: Sept. 15, 2019. [Online]. Available: https://community.cadence.com/cadence_blogs_8/b/ii/posts/signoff-summit-an-update-on-ocv-aocv-socv-and-statistical-timing
- [11] M. Alioto, G. Scotti, and A. Trifiletti, "A Novel Framework to Estimate the Path Delay Variability on the Back of an Envelope via the Fan-Out-of-4 Metric," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 8, pp. 2073–2085, 2017.
- [12] M. Alioto and G. Palumbo, "Impact of Supply Voltage Variations on Full Adder Delay: Analysis and Comparison," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 14, no. 12, pp. 1322–1335, 2006.
- [13] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Pearson, and N. J. Rohrer, "High-performance CMOS variability in the 65-nm regime and beyond," *IBM Journal of Research and Development*, vol. 50, no. 4.5, pp. 433–449, 2006.
- [14] K. Papachatzopoulos and V. Paliouras, "Low-Power Addition with Borrow-Save Adders under Threshold Voltage Variability," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 5, pp. 572–576, 2018.
- [15] —, "Static Delay Variation Models for Ripple-Carry and Borrow-Save Adders," *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2019.
- [16] K. Papachatzopoulos, I. Kouretas, and V. Paliouras, "Dynamic delay variation behaviour of RNS multiply-add architectures," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2016, pp. 1978–1981.
- [17] M. H. Abu-Rahma and M. Anis, "A Statistical Design-Oriented Delay Variation Model Accounting for Within-Die Variations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 11, pp. 1983–1995, 2008.
- [18] N. H. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*. Noida, India: Pearson Educ., 2015.
- [19] S. Knowles, "A family of adders," in *Proceedings 14th IEEE Symposium on Computer Arithmetic (Cat. No. 99CB36336)*. IEEE, 1999, pp. 30–34.
- [20] P. M. Kogge and H. S. Stone, "A parallel algorithm for the efficient solution of a general class of recurrence equations," *IEEE Transactions on Computers*, vol. 100, no. 8, pp. 786–793, 1973.
- [21] J. Sklansky, "Conditional-sum addition logic," *IRE Transactions on Electronic Computers*, no. 2, pp. 226–231, 1960.
- [22] R. P. Brent and H. T. Kung, "A regular layout for parallel adders," *IEEE Transactions on Computers*, no. 3, pp. 260–264, 1982.
- [23] T. Han and D. A. Carlson, "Fast area-efficient VLSI adders," in *1987 IEEE 8th Symposium on Computer Arithmetic (ARITH)*. IEEE, 1987, pp. 49–56.
- [24] R. E. Ladner and M. J. Fischer, "Parallel prefix computation," *Journal of the ACM (JACM)*, vol. 27, no. 4, pp. 831–838, 1980.
- [25] D. Harris, "A taxonomy of parallel prefix networks," in *The Thirtieth Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2. IEEE, 2003, pp. 2213–2217.
- [26] C. E. Clark, "The greatest of a finite set of random variables," *Operations Research*, vol. 9, no. 2, pp. 145–162, 1961.