# Accelerated Accurate Timing Yield Estimation Based on Control Variates and Importance Sampling

Alp Arslan Bayrakci, Member, IEEE

Abstract—Extensive research has been conducted on a statistical timing analysis of digital integrated circuits in the existence of statistical parameter variations. However, the proposed methods either lack accuracy or efficiency, which avoids coming up with an industry standard tool. Despite this fact, there is a certain consensus that Monte Carlo (MC) methods are accurate, so that they are called golden. In this paper, we propose novel techniques to combine control variates with importance sampling in order to come up with a new timing yield estimator as accurate as SPICE-based standard MC (STD-MC) but much faster. The performance of three different estimators, two of which are proposed in this paper, is compared through experiments, and the results show that the precise SPICE simulation-based STD-MC method can be accelerated about  $260 \times$  on the average without sacrificing any accuracy.

*Index Terms*—Control variates (CV), importance sampling (IS), Monte Carlo (MC), statistical timing analysis, timing yield, variance reduction.

## I. INTRODUCTION

**S** TATISTICAL timing analysis aims at estimating timing yield that is the fraction of chips that satisfy the timing constraints in the presence of statistical parameter variations. Accurate estimation of timing yield is crucial for integrated circuit (IC) industry for the decision of further optimization or manufacturing. There are various challenges for the estimation of timing yield. A large number of random variables, interdie, and intradie variations with spatial and topological correlations, the necessity to work with non-Gaussian random variables, and the nonlinear and differential relationship between random circuit parameters and performance metrics such as circuit delay are the main difficulties of timing yield estimation problem.

For the last decade, extensive research has been conducted on the statistical timing analysis, with an emphasis on a statistical static timing analysis (SSTA), which is the statistical extension of static timing analysis. A comprehensive review of the statistical timing analysis literature is presented in [1] and [2]. A survey over IC industry professionals shows the necessity of variation-aware design and SPICE simulations

The author is with the Department of Computer Engineering, Gebze Technical University, Gebze 41400, Turkey (e-mail: abayrakci@gtu.edu.tr). Color versions of one or more of the figures in this paper are available

online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2016.2521541

considering parameter variations [3]. The inclusion of the transistor-level (TL) path-based statistical timing analysis into the reference design flow by Taiwan Semiconductor Manufacturing Company is another sign for the necessity of the statistical analysis [4].

The widely accepted golden method to estimate timing yield is Monte Carlo (MC) estimation, especially when it is based on precise TL SPICE simulations. MC estimation dates back to 1946 [5] and formally proposed in [6]. Today, this golden method is used as reference by most of the statistical timing analysis papers, yet its computational complexity makes it impossible to be used solely. There is considerable effort for accelerating the MC method using variance reduction techniques, such as importance sampling (IS) [7] and control variates (CV) [8], because the accuracy of MC-based estimation is crucial and very difficult to be reached by its alternatives. There are plenty of books describing the variance reduction techniques even by researchers of different academic disciplines [9]-[13]. Singhee et al. [14] state that the MC-based statistical timing analysis is believed to be too slow despite the lack of rigorous studies to support this belief. This paper refutes this idea and shows that MC need not be slow similar to the argument in [15].

This paper is definitely not the first one in the literature that uses variance reduction techniques in order to increase the efficiency of MC analysis of statistical phenomena in electronic circuits. The initial efforts were made for estimating the yield of analog circuits. One of the earliest papers published in [16] separately implements IS, stratified sampling, and CV techniques to estimate general yield of analog circuits, such as low-pass filter. Around that time, Soin and Rankin [17] present a classical CV-based method employing a shadow model for estimating yield of analog circuits. Recent papers [18], [19] on the same topic propose CV and a sample reduction technique called latin hypercube sampling (LHS), respectively.

The problem of timing yield estimation and an incremental timing analysis of digital ICs can also be dealed with the application of variance reduction techniques to MC-based SSTA. A comprehensive work in [20] employs different variance reduction techniques for this purpose. The first method in this paper is an enhanced version of the quasi-MC (QMC) method [21] with about six times speedup. The disadvantage of QMC/LHS approach is the inability to handle high dimensionality and almost negligible speedups it offers. The second method in this paper is the order statistics-based CV similar

1063-8210 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications\_standards/publications/rights/index.html for more information.

Manuscript received August 26, 2015; revised November 30, 2015; accepted January 11, 2016. Date of publication February 11, 2016; date of current version July 22, 2016. This work was supported by The Scientific and Technological Research Council of Turkey under Grant 112E237.

to the one in [22]. However, this technique is prone to the underestimation of the timing yield and also the speedup it offers is almost similar to the QMC-based method. The last method of this paper is based on classical CV, but this time CV is used to estimate mean, variance, and skewness of the critical delay probability distribution as in [23]. Another work in [24] combines QMC and LHS to obtain the statistical distribution of circuit delay, whereas the work in [25] further combines stratified sampling with QMC and LHS for the same purpose, also addressing incremental timing analysis. This class of papers aims to improve the speed of the gatelevel static timing analysis-based MC to compete against fast SSTA methods. This brings the necessity to tolerate some errors and makes some assumptions on the circuitlike Gaussian arrival times and linearity, which contradicts with some of the major reasons behind using an MC estimator and result in inherently biased estimators regardless of sample size [20].

Another type of problem that variance reduction techniques are extensively used is SRAM failure analysis. Similar to [26], Kanj et al. [27] apply a technique called mixture IS to estimate SRAM failure probabilities. Singhee and Rutenbar [28] argue that QMC is better than LHS, where LHS is better than conventional MC. They test their method on SRAM or flip-flop circuits, resulting in  $2 \times$  to  $8 \times$  speedup. A machine learningbased sample reduction method called statistical blockade is used to accelerate the SRAM failure analysis in [29], whereas IS is used to calculate SRAM robustness resulting in a seven-order run-time improvement [30]. SRAM failure is a relatively low-dimensional problem [22]. Problems, such as the difficulty to adapt more complex digital circuits and the inability to cope with the curse of dimensionality resulting in low speedups, must be solved to further apply these methods to more complex digital circuits.

Our aim is to accelerate timing yield estimation of digital ICs without sacrificing the accuracy of TL SPICE simulationsbased standard MC (STD-MC). Similar to the statement in [31], the approach proposed in this paper is based on the premise that, given the magnitude of process parameter variations and the nonlinear dependence of gate and circuit delay on these variations, the only sufficiently reliable and accurate method for final timing yield verification before sign off is the TL circuit simulation. However, we realize that the TL MC estimation of timing yield may never become efficient enough for use in a loop for timing optimizations. As such, the MC timing yield estimation method that we propose in this paper is meant not as a replacement for fast block-based SSTA methods, but rather, as a complement to them.

In [31], we devise a methodology to estimate timing yield of digital ICs as accurate as SPICE-based STD-MC. The method relies on IS in conjunction with an approximate gate delay model called a polynomial delay model (PDM). Although the resultant IS estimator obtains about  $150 \times$  speedups with respect to STD-MC, the efficiency of the IS estimator decreases by decreasing timing yield values as expected.

This paper is a significant improvement over the work in [31]. We first construct a timing yield estimator based on the well-known CV technique in conjunction with PDM and TL SPICE simulations. Then, using this CV estimator as the first step, we combine CV and IS in a novel manner to propose a new timing yield estimator called CV with IS estimator. Besides the combination of two different variance reduction techniques, this paper solves two crucial issues to obtain an efficient estimator: a selection of a good biasing distribution and the development of a heuristic algorithm to detect margins. The proposed solutions to these issues and the way CV and IS are combined define the performance of the resultant CVIS estimator. The experiments verify that the accuracy of the theoretical derivations and CVIS estimator achieves a speedup of  $260 \times$  on the average with respect to STD-MC.

The rest of this paper is organized as follows. Section II formulates the problem and reminds the IS estimator. The theory of the estimators is presented in Section III. The section starts with the CV estimator, it continues with the CVIS estimator, and finally, it presents the margin detection heuristic. Section IV shows the results, where the performance of the estimators is reported with respect to STD-MC estimator. The theoretical variance derivations of this paper are also verified in that section. Finally, the conclusion is drawn in Section V.

#### **II. BACKGROUND AND PRELIMINARIES**

#### A. Problem Formulation

Satisfying timing requirements for a circuit corresponds to having smaller delays than a predefined timing constraint value, which is shown as  $T_c$  throughout this paper. X is a random variable vector, each element of which corresponds to a random variable in the circuit. The joint probability density function (pdf) corresponding to X is represented by f(X).

 $d_C^M(X)$  represents maximum circuit delay, computed by a method M and when the random variables inside the circuit are set according to the values given by X. The indicator variable  $I^M(T_c, X)$  is defined as

$$I^{M}(T_{c}, X) = \begin{cases} 1, & \text{if } d^{M}_{C}(X) > T_{c} \\ 0, & \text{if } d^{M}_{C}(X) \le T_{c}. \end{cases}$$
(1)

The most accurate circuit delay computation method is the industry standard TL SPICE simulation. If the circuit delay is computed by TL SPICE simulations, we put TL instead of M in the definitions. Using these definitions, the best estimation for timing yield, which is equal to the probability that a manufactured chip will satisfy the timing requirements, can be defined as one minus the expected value of the indicator variable, as shown in

$$\text{Yield}^{\text{TL}} = 1 - \text{Loss}^{\text{TL}} = 1 - \int_{\Omega} I^{\text{TL}}(T_c, X) f(X) dX \quad (2)$$

where  $\Omega$  is the whole random parameter space, and Loss<sup>TL</sup> is the probability that a manufactured chip will not satisfy the timing constraint. We call Yield<sup>TL</sup> the actual yield and Loss<sup>TL</sup> the actual loss, because they depend on accurate TL SPICE simulations. However, it is not possible to compute this integral analytically. The MC method is widely used to estimate such integrals. In (3), MC estimator for  $Loss^{TL}$  in (2) can be seen

$$\operatorname{Loss}_{N}^{\mathrm{TL}} = \frac{1}{N} \sum_{i=1}^{N} I^{\mathrm{TL}}(T_{c}, X_{i})$$
(3)

where  $X_i$  is a sample drawn from the joint pdf f(X) and N is the number of sample points used in estimation. This estimator is called STD-MC estimator throughout this paper. It is well known with its accuracy, where it requires almost no approximations in estimating the timing loss value. However, it requires a large number of samples to be accurate. This, in turn, means too many TL SPICE simulations each of which is computationally too costly resulting in very slow estimation performance.

Expected value of an MC estimator is the actual value the estimator converges to, when an infinite number of sample points are used for the estimation. It equals to the actual timing loss (Loss<sup>TL</sup>) for an unbiased estimator as in (3). Variance of an MC estimator is the expected squared distance between the estimator's resultant estimate and the actual value. For timing loss estimation problem and for an unbiased MC estimator using method M inside, it becomes as shown in (4). Variance value shows the convergence rate of the estimator over another one as will be discussed later. It is also shown in [31] that the absolute difference between the estimator's resultant estimate and the actual loss value cannot be higher than two times the square root of the estimator's variance with more than 95% confidence

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{M}\right\} = E\left\{\left(\operatorname{Loss}_{N}^{M} - \operatorname{Loss}^{\mathrm{TL}}\right)^{2}\right\}.$$
(4)

The variance of the well-known STD-MC estimator is shown in (5) [31]. This equation also shows the dependence of variance on the number of samples

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\operatorname{TL}}\right\} = \frac{\operatorname{Loss}^{\operatorname{TL}} \cdot \operatorname{Yield}^{\operatorname{TL}}}{N}.$$
(5)

## B. Importance Sampling Estimator

IS can be used to speed up the estimator in (3) without losing accuracy. According to this method, Bayrakci *et al.* [31] propose the IS estimator shown by (6), to estimate timing loss

$$\operatorname{Loss}_{N}^{\mathrm{IS}} = \frac{\operatorname{Loss}^{\mathrm{ADM},\varepsilon}}{N} \sum_{i=1}^{N} I^{\mathrm{TL}}(T_{c}, X_{i})$$
(6)

where all sample points are drawn from another distribution  $\tilde{f}(X)$ , shown in (7), instead of f(X)

$$\tilde{f}(X) = \frac{I^{\text{ADM}}(T_c^{\varepsilon}, X_i) f(X)}{\text{Loss}^{\text{ADM}, \varepsilon}}.$$
(7)

ADM represents a fast but approximate delay model used to approximately capture the variations of gate delays with respect to the circuit parameter variations. Accordingly,  $\text{Loss}^{\text{ADM},\varepsilon}$  is the loss value, estimated based on delays computed by ADM and using a timing constraint with  $\varepsilon$  margin, i.e.,  $T_c^{\varepsilon}$ . This loss value is not very accurate instead it is a loose approximation, computed by the MC estimator in (8), where fast ADM is used instead of TL circuit simulations enabling the use of a huge number of N

$$\operatorname{Loss}^{\operatorname{ADM},\varepsilon} = \frac{1}{N} \sum_{i=1}^{N} I^{\operatorname{ADM}} \left( T_{c}^{\varepsilon}, X_{i} \right)$$
(8)

where *N*, in the left-hand side of the equation, is omitted as it is a huge number. The  $T_c^{\varepsilon}$  in these equations represents  $T_c - \varepsilon$ , where  $\varepsilon$  is a safety margin computed by an adaptive heuristic algorithm again proposed in [31]. This margin is needed in order to satisfy the condition that for every sample point, *X*, if  $I^{\text{TL}}(T_c, X)$  is 1, then  $I^{\text{ADM}}(T_c^{\varepsilon}, X)$  is also 1. This condition is necessary to have an unbiased loss estimator based on IS, as shown in (6). The expected value of the IS estimator is shown to be equal to the actual loss, i.e.,  $\text{Loss}^{\text{TL}}$ . For this estimator, the theoretical variance, which represents the convergence rate of the estimator, is shown in [31]

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\mathrm{IS}}\right\} = \frac{\operatorname{Loss}^{\mathrm{TL}} \cdot (\operatorname{Loss}^{\mathrm{ADM},\varepsilon} - \operatorname{Loss}^{\mathrm{TL}})}{N}.$$
 (9)

When the variance of the IS estimator is compared with the variance of the STD-MC estimator in (5), it is seen that the numerator in (9) has  $\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{TL}}$  instead of Yield<sup>TL</sup> which is equal to  $1 - \text{Loss}^{\text{TL}}$  in (5). As a result, the variance of the IS estimator has to be smaller than or equal to the STD-MC estimator as  $\text{Loss}^{\text{ADM},\varepsilon}$  is smaller or equal to 1 theoretically. Whereas, practically  $\text{Loss}^{\text{ADM},\varepsilon}$  is closer to the actual loss as it is the loss value estimated by (8), which explains why IS estimator has much smaller variance than the STD-MC estimator for a fixed number of samples (*N*).

## **III. PROPOSED ESTIMATORS**

STD-MC estimator is the most accurate statistical timing loss estimator provided that an enough number of samples are used with the estimator. The main problem with the STD-MC estimator is that it requires too many costly TL SPICE simulations to converge to the actual loss, Loss<sup>TL</sup>. Therefore, converging to an accurate value is not enough for an estimator to be good. The goodness of a loss estimator depends on two quantities: the accuracy of the value it converges to and its convergence rate to that value.

- Mean of the Estimator: The estimator must converge to the actual value when a larger number of samples are used. That means that the expected value of the estimator must be equal to the actual value it estimates, i.e., Loss<sup>TL</sup> in our case. Such estimators are called unbiased. This condition is very important, because if an estimator has a large bias, no matter how many sample points are utilized, and the resultant estimates are far from the actual result.
- 2) Variance of the Estimator: The convergence rate of the estimator must be high, so that even with relatively small number of sample points (N) used inside the MC estimator, the estimates are not far from the actual value. This means that the variance (error) of the estimator using a fixed N must be low. In other words, an estimator with a smaller variance requires less number of samples (simulations) to estimate actual value with

a fixed error tolerance. Therefore, decreasing the variance of an estimator transforms into a speedup in the estimation.

Accurate estimates can be obtained with fewer samples (simulations) only if the estimator is both unbiased and has a higher convergence rate. A biased estimator results in the lack of accuracy, whereas slow convergence results in a need for more samples.

In this section, CV estimator, which is based on CV solely is proposed as an initial step. Then, we present the prime CVIS estimator that is based on a novel combination of CV with IS. The aim is to increase the convergence rate of the STD-MC estimator without introducing bias and much extra cost. Throughout this section, the presented estimators are theoretically analyzed in terms of mean and variance to see whether the estimates converge to the actual loss and at what rate they converge.

# A. Control Variates Estimator

The CV estimator is constructed utilizing an approximate but very efficient gate delay model (ADM) like the PDM [31] in accordance with the CV variance reduction technique. We start with Loss<sup>TL</sup> integral in (2) and apply it CV by adding and subtracting  $I^{\text{ADM}}(T_c, X)$  as shown in

$$\operatorname{Loss}^{\mathrm{TL}} = \int_{\Omega} (I^{\mathrm{TL}}(T_{c}, X) - I^{\mathrm{ADM}}(T_{c}, X) + I^{\mathrm{ADM}}(T_{c}, X)) f(X) dX$$
$$= \int_{\Omega} I^{\mathrm{ADM}}(T_{c}, X) f(X) dX$$
$$+ \int_{\Omega} (I^{\mathrm{TL}}(T_{c}, X) - I^{\mathrm{ADM}}(T_{c}, X)) f(X) dX. \quad (10)$$

The first integral is equal to Loss<sup>ADM</sup> by definition. For the second integral, MC estimation is applied, as shown in

$$\text{Loss}_{N}^{\text{CV}} = \text{Loss}^{\text{ADM}} + \frac{1}{N} \sum_{i=1}^{N} (I^{\text{TL}}(T_{c}, X_{i}) - I^{\text{ADM}}(T_{c}, X_{i})).$$
(11)

To compute Loss<sup>ADM</sup>, (8) is used but with no margin, i.e.,  $\varepsilon = 0$  and with a huge number of sample points (*N*). The computation of Loss<sup>ADM</sup> is not computationally costly, because the computation of  $I^{\text{ADM}}(T_c, X_i)$  does not depend on SPICE TL simulations, and instead, it relies on approximate delay models with very low evaluation costs.

We could instead add and subtract  $\beta \cdot I^{\text{ADM}}(T_c, X)$  in (10), where  $\beta$  is to be defined optimally, as shown in [17]. However, it is impossible to compute this optimum value, because it is a function of the covariance of  $I^{\text{ADM}}(T_c, X)$  and  $I^{\text{TL}}(T_c, X)$ , which is unknown *a priori* as  $I^{\text{TL}}(T_c, X)$  requires costly TL simulations. Estimating this value using a small test set is a method, but using different  $\beta$  values for different estimation runs causes a bias, which lowers the speedup of the CV estimator. Optimum  $\beta$  is unknown *a priori* but can be computed if a huge number of TL simulations are performed to collect the indicator variable values. The resultant optimum  $\beta$  values for ISCAS-85 circuits are almost equal to 1, where the most distant value to 1 comes out to be 0.989. Therefore, similar to [17] and [20],  $\beta$  is taken to be one in this paper.

1) Mean of the CV Estimator: As it is discussed in the beginning of this section, the mean of an estimator is an indicator of its accuracy. Below, we compute the mean of the CV estimator.

*Theorem 1:* The expected value of the CV estimator in (11) is equal to the actual loss value

$$E\{\operatorname{Loss}_{N}^{\operatorname{CV}}\} = \operatorname{Loss}^{\operatorname{TL}}.$$
(12)

*Proof:* Considering constants in (11), expected value becomes

$$E\{\text{Loss}_{N}^{\text{CV}}\} = \text{Loss}^{\text{ADM}} + \frac{1}{N} \sum_{i=1}^{N} (E\{I^{\text{TL}}(T_{c}, X_{i})\} - E\{I^{\text{ADM}}(T_{c}, X_{i})\}).$$
(13)

From (2), we know that

$$E\{I^{\mathrm{TL}}(T_c, X)\} = \int_{\Omega} I^{\mathrm{TL}}(T_c, X) f(X) dX$$
  
= Loss<sup>TL</sup> (14)

where  $\Omega$  is the whole feasible parameter space. Similarly

$$E\{I^{\text{ADM}}(T_c, X)\} = \int_{\Omega} I^{\text{ADM}}(T_c, X) f(X) dX$$
  
= Loss<sup>ADM</sup>. (15)

Combining these together, we get

$$E\{\text{Loss}_{N}^{\text{CV}}\} = \text{Loss}^{\text{ADM}} + \text{Loss}^{\text{TL}} - \text{Loss}^{\text{ADM}} = \text{Loss}^{\text{TL}}$$
(16)

 $\square$ 

which completes the proof.

This concludes that the CV estimator is unbiased, which means that while the number of sample points (N) increases, the result of the CV estimator converges to the actual loss. Now, the question is how fast it converges, which is answered next.

2) Variance of the CV Estimator: In this section, the variance of the CV estimator for a fixed number of sample points is derived, so that it can be compared with other estimators in terms of converging rate.

*Theorem 2:* The variance of the CV estimator in (11) has a lower bound shown with

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{CV}\right\} \geq \frac{|\operatorname{Loss}^{TL} - \operatorname{Loss}^{ADM}| \cdot (1 - |\operatorname{Loss}^{TL} - \operatorname{Loss}^{ADM}|)}{N}.$$
 (17)

*Proof:* There is no variance on Loss<sup>ADM</sup>, because it is computed using a huge number of sample points for once, and this Loss<sup>ADM</sup> value is used for all different estimations.

Since the summation terms in (11) are independent, the variance of the CV estimator can be written as

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\operatorname{CV}}\right\} = \operatorname{Var}\left\{\frac{1}{N}\sum_{i=1}^{N} (I^{\operatorname{TL}}(T_{c}, X_{i}) - I^{\operatorname{ADM}}(T_{c}, X_{i}))\right\}$$
$$= \frac{1}{N}\operatorname{Var}\left\{I^{\operatorname{TL}}(T_{c}, X) - I^{\operatorname{ADM}}(T_{c}, X)\right\}$$
$$= \frac{1}{N}(\operatorname{Var}\left\{I^{\operatorname{TL}}(T_{c}, X)\right\} + \operatorname{Var}\left\{I^{\operatorname{ADM}}(T_{c}, X)\right\}$$
$$- 2\operatorname{Cov}\left\{I^{\operatorname{TL}}(T_{c}, X), I^{\operatorname{ADM}}(T_{c}, X)\right\}).$$
(18)

We must compute two variance values and the covariance in (18). Using (14), we get

$$Var\{I^{TL}(T_c, X)\} = E\{(I^{TL}(T_c, X) - Loss^{TL})^2\} = E\{I^{TL}(T_c, X)\} - Loss^{TL^2} = Loss^{TL}(1 - Loss^{TL}).$$
(19)

Using (15), we get

$$\operatorname{Var}\{I^{\operatorname{ADM}}(T_c, X)\} = E\{(I^{\operatorname{ADM}}(T_c, X) - \operatorname{Loss}^{\operatorname{ADM}})^2\}$$
$$= E\{I^{\operatorname{ADM}}(T_c, X)\} - \operatorname{Loss}^{\operatorname{ADM}^2}$$
$$= \operatorname{Loss}^{\operatorname{ADM}}(1 - \operatorname{Loss}^{\operatorname{ADM}}).$$
(20)

Using (14) and (15), we can write

$$\operatorname{Cov}\{I^{\mathrm{TL}}(T_c, X), I^{\mathrm{ADM}}(T_c, X)\} = E\{I^{\mathrm{TL}}(T_c, X) \cdot I^{\mathrm{ADM}}(T_c, X)\} - \operatorname{Loss}^{\mathrm{TL}} \operatorname{Loss}^{\mathrm{ADM}}.$$
 (21)

 $I^{\text{TL}}(T_c, X)$  and  $I^{\text{ADM}}(T_c, X)$  can be either 0 or 1. The expected value in (21) is the integral of f(X) over the region, where both the indicator variables are 1. Therefore, using (14) and (15), we can write the relationship below

$$E\{I^{\mathrm{TL}}(T_c, X) \cdot I^{\mathrm{ADM}}(T_c, X)\} \le \min(\mathrm{Loss}^{\mathrm{TL}}, \mathrm{Loss}^{\mathrm{ADM}}).$$
(22)

Combining (19)–(22) as shown in (18) and after some trivial computations, we get the relationship shown in Theorem 17.

Theorem 17 gives a lower bound for the variance of the CV estimator. The greater or equal operator in (17) becomes an equality, if for every sample point  $I^{\text{ADM}}(T_c, X)$  is 1, then  $I^{\text{TL}}(T_c, X)$  is 1 or alternatively if the reverse is true.

CV estimator is advantageous in the sense that it does not need any margin detection heuristic to be unbiased in contrast to IS estimator. However, when the variance lower bound of the CV estimator shown by (17) is compared with the variance of the previously proposed IS estimator in (9), it can be deduced that the CV estimator is not as good as the IS estimator, because, in practice,  $\text{Loss}^{\text{TL}}$  is expected to be smaller than  $(1 - |\text{Loss}^{\text{TL}} - \text{Loss}^{\text{ADM}}|)$ .

# B. Control Variates With Importance Sampling Estimator

In order to have a faster estimator than the IS estimator, we combine IS and CV techniques in a novel manner to construct the new CVIS estimator. We start with (10). For the difference integral in this equation, we apply IS technique by introducing

the biasing distribution  $\tilde{f}(X)$ . The first integral is known to be equal to  $\text{Loss}^{\text{ADM}}$ . As a result, the following equation comes out:

$$Loss^{TL} = Loss^{ADM} + \int_{\Omega} (I^{TL}(T_c, X) - I^{ADM}(T_c, X)) \frac{f(X)}{\tilde{f}(X)} \tilde{f}(X) dX.$$
(23)

The integral in (23) can be estimated using the MC method. The resultant CVIS timing loss estimator, which utilizes a combination of CV and IS, can be constructed, as shown in

$$\operatorname{Loss}_{N}^{\operatorname{CVIS}} = \operatorname{Loss}^{\operatorname{ADM}} + \frac{1}{N} \sum_{i=1}^{N} (I^{\operatorname{TL}}(T_{c}, \tilde{X}_{i}) - I^{\operatorname{ADM}}(T_{c}, \tilde{X}_{i})) \frac{f(X)}{\tilde{f}(X)}$$

$$(24)$$

where  $\tilde{X}_i$  represents the sample point drawn from  $\tilde{f}(X)$  instead of f(X). Here, the key issue is the selection of a good  $\tilde{f}(X)$ . This new distribution must have the following properties to be a regular pdf resulting in an unbiased loss estimator.

- 1) Integral of  $\tilde{f}(X)$   $(\int_{\Omega} \tilde{f}(X) dX)$  over the whole parameter space must be equal to 1 as it is a pdf.
- 2)  $\tilde{f}(X)$  must be nonzero everywhere  $I^{\text{TL}}(T_c, X) I^{\text{ADM}}(T_c, X)$  is nonzero. Otherwise, (23) cannot be equal to (10).

We propose the following  $\tilde{f}(X)$ :

$$\tilde{f}(X) = \frac{(I^{\text{ADM}}(T_c - \varepsilon, X) - I^{\text{ADM}}(T_c + \delta, X))f(X)}{\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{ADM},\delta}}$$
(25)

where  $\varepsilon$  and  $\delta$  are margins that are determined by a heuristic introduced in Section III-C. Loss<sup>ADM, $\varepsilon$ </sup> and Loss<sup>ADM, $\delta$ </sup> are defined in

$$\operatorname{Loss}^{\operatorname{ADM},\varepsilon} = \int_{\Omega} I^{\operatorname{ADM}}(T_c - \varepsilon, X) f(X) dX \qquad (26)$$

$$\operatorname{Loss}^{\operatorname{ADM},\delta} = \int_{\Omega} I^{\operatorname{ADM}}(T_c + \delta, X) f(X) dX.$$
(27)

Let us look if  $\tilde{f}$  in (25) satisfies the above two properties

$$\int_{\Omega} \tilde{f}(X) dX$$

$$= \frac{\int_{\Omega} (I^{\text{ADM}}(T_c - \varepsilon, X) - I^{\text{ADM}}(T_c + \delta, X)) f(X) dX}{\text{Loss}^{\text{ADM}, \varepsilon} - \text{Loss}^{\text{ADM}, \delta}}$$

$$= \frac{\text{Loss}^{\text{ADM}, \varepsilon} - \text{Loss}^{\text{ADM}, \delta}}{\text{Loss}^{\text{ADM}, \varepsilon} - \text{Loss}^{\text{ADM}, \delta}} = 1.$$
(28)

Therefore, property 1) above holds. This makes  $\tilde{f}$  in (25) a regular pdf.

For property 2) to hold, the following condition must be satisfied. For every where  $I^{\text{TL}}(T_c, X) - I^{\text{ADM}}(T_c, X)$  is nonzero, then  $I^{\text{ADM}}(T_c - \varepsilon, X) - I^{\text{ADM}}(T_c + \delta, X)$  must be nonzero. This is why we introduce the margins  $\varepsilon$  and  $\delta$ . The margins must be large enough to satisfy property 2). This can be done by satisfying two conditions.

- 1) For every point  $I^{TL}(T_c, X)$  is 1,  $I^{ADM}(T_c \varepsilon, X)$  is also 1.
- 2) For every point  $I^{\text{ADM}}(T_c + \delta, X)$  is 1,  $I^{\text{TL}}(T_c, X)$  is also 1.

These two conditions are called margin conditions throughout this paper. There is a tradeoff in margin selection. Selecting very large  $\varepsilon$  and  $\delta$  margins surely satisfies the margin conditions. However, it slows down the convergence rate of the resultant CVIS estimator. The reason for that will be clarified later in this section. Ideally, the margins must be the minimum values satisfying the margin conditions. Heuristic proposed in Section III-C is used to smartly select these margins sufficient enough to satisfy margin conditions.

After f(X) in (25) is substituted into (24), the resultant CVIS estimator is shown in

$$\operatorname{Loss}_{N}^{\operatorname{CVIS}} = \operatorname{Loss}^{\operatorname{ADM}} + \frac{\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta}}{N} \times \sum_{i=1}^{N} (I^{\operatorname{TL}}(T_{c}, \tilde{X}_{i}) - I^{\operatorname{ADM}}(T_{c}, \tilde{X}_{i})) \quad (29)$$

where N sample points are drawn according to the  $\tilde{f}(X)$  distribution instead of f(X). The Loss<sup>ADM</sup>, Loss<sup>ADM, $\varepsilon$ </sup> and Loss<sup>ADM, $\delta$ </sup> in (29) are computed using a huge number of samples and (8) with no margin,  $\varepsilon$  margin, and  $\delta$  margin, respectively. It should be noted that having computed the circuit delays according to ADM for all samples, it is very easy and fast to compute these three loss values using (8) with different margins. In order to measure the goodness of the CVIS estimator, the mean and the variance of the estimator in (29) are derived next.

1) Mean of the CVIS Estimator: CVIS estimator is an unbiased estimator eventually converging to the actual loss provided that the margin conditions hold. The detailed derivation of the expected value for this estimator is given below.

*Theorem 3:* Provided that the margin conditions hold, the expected value of the CVIS estimator in (29) is equal to the actual loss value

$$E\{\operatorname{Loss}_{N}^{\operatorname{CVIS}}\} = \operatorname{Loss}^{\operatorname{TL}}.$$
(30)

*Proof:* Considering constants in (29), expected value becomes

$$E\{\text{Loss}_{N}^{\text{CVIS}}\} = \text{Loss}^{\text{ADM}} + \frac{\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{ADM},\delta}}{N} \times \sum_{i=1}^{N} E\{I^{\text{TL}}(T_{c}, \tilde{X}_{i})\} - E\{I^{\text{ADM}}(T_{c}, \tilde{X}_{i})\}$$
(31)

where N samples are drawn from  $\tilde{f}(X)$ . Therefore, we can write

$$E\{I^{\mathrm{TL}}(T_c, X)\} = \int_{\Omega} I^{\mathrm{TL}}(T_c, X) \tilde{f}(X) dX$$
$$= \frac{\int_A^B I^{\mathrm{TL}}(T_c, X) f(X) dX}{\mathrm{Loss}^{\mathrm{ADM}, \varepsilon} - \mathrm{Loss}^{\mathrm{ADM}, \delta}}.$$
(32)

When  $\tilde{f}(X)$  in (25) is substituted into (32), the boundaries can be changed considering the nonzero regions of  $\tilde{f}(X)$ . A is the boundary, where  $I^{\text{ADM}}(T_c - \varepsilon, X)$  changes from 0 to 1, and *B* is the boundary where  $I^{\text{ADM}}(T_c + \delta, X)$  changes from 0 to 1. Provided that the margin conditions explained above are satisfied, the integral in the numerator of (32) becomes equal to  $\text{Loss}^{\text{TL}} - \text{Loss}^{\text{ADM},\delta}$ 

$$E\{I^{\text{ADM}}(T_c, X)\} = \int_{\Omega} I^{\text{ADM}}(T_c, X) \tilde{f}(X) dX$$
$$= \frac{\int_{A}^{B} I^{\text{ADM}}(T_c, X) f(X) dX}{\text{Loss}^{\text{ADM}, \varepsilon} - \text{Loss}^{\text{ADM}, \delta}}.$$
 (33)

Similar to (32), the integral in (33) with boundaries given by A and B becomes equal to  $\text{Loss}^{\text{ADM}} - \text{Loss}^{\text{ADM},\delta}$ . When we substitute the integrals in (32) and (33) and then substitute these two equations into (31), we get

$$E\{\text{Loss}_{N}^{\text{CVIS}}\} = \text{Loss}^{\text{ADM}} + \text{Loss}^{\text{TL}} - \text{Loss}^{\text{ADM}}$$
(34)

 $\square$ 

which finalizes the proof.

This theorem means that if the number of sample points N is increased, the estimate of the CVIS estimator converges to the actual loss value eventually. This is about the bias of the estimator. The second most important question is how fast it converges to the actual loss, which is answered next.

2) Variance of the CVIS Estimator: Below, we derive the variance of the CVIS estimator as a measure of the convergence rate of the estimator to its mean, which is the actual loss as proved above.

*Theorem 4:* Provided that the margin conditions hold, the variance of the CVIS estimator in (29) has a lower bound shown with

$$\begin{aligned}
\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\operatorname{CVIS}}\right\} &\geq |\operatorname{Loss}^{\operatorname{TL}} - \operatorname{Loss}^{\operatorname{ADM}}| \\
\times \frac{\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta} - |\operatorname{Loss}^{\operatorname{TL}} - \operatorname{Loss}^{\operatorname{ADM}}|}{N}.
\end{aligned}$$
(35)

*Proof:* Similar to what we explained inside the proof of the CV estimator variance,  $Loss^{ADM}$ ,  $Loss^{ADM,\varepsilon}$ , and  $Loss^{ADM,\delta}$  are fixed constants after computed for once. From (29) and independency of the summation terms, it can be deduced that

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\operatorname{CVIS}}\right\} = \frac{(\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta})^{2}}{N^{2}} \times N$$
$$\times \operatorname{Var}\left\{I^{\operatorname{TL}}(T_{c}, X) - I^{\operatorname{ADM}}(T_{c}, X)\right\} \quad (36)$$

where X is shown from  $\tilde{f}(X)$  in (25) and

$$\operatorname{Var}\{I^{\mathrm{TL}}(T_{c}, X) - I^{\mathrm{ADM}}(T_{c}, X)\} = \underbrace{\operatorname{Var}\{I^{\mathrm{TL}}(T_{c}, X)\}}_{\tilde{\sigma}_{I}\mathrm{TL}} + \underbrace{\operatorname{Var}\{I^{\mathrm{ADM}}(T_{c}, X)\}}_{\tilde{\sigma}_{I}\mathrm{ADM}} - 2\underbrace{\operatorname{Cov}\{I^{\mathrm{TL}}(T_{c}, X), I^{\mathrm{ADM}}(T_{c}, X)\}}_{\tilde{\sigma}_{I}\mathrm{TL}, I^{\mathrm{ADM}}}.$$
(37)

In the remaining part of the proof, the three expressions in (37) will be derived and then combined, but before that we write down the expected values for  $I^{\text{TL}}(T_c, X)$  and  $I^{\text{ADM}}(T_c, X)$ ,

where the random variable X is from f(X) in (25). Directly from (32) and (33)

$$\underbrace{E\{I^{\text{TL}}(T_c, X)\}}_{\tilde{\mu}_{I^{\text{TL}}}} = \frac{\text{Loss}^{\text{TL}} - \text{Loss}^{\text{ADM},\delta}}{\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{ADM},\delta}}$$
(38)

$$\underbrace{E\{I^{\text{ADM}}(T_c, X)\}}_{\tilde{\mu}_{I^{\text{ADM}}}} = \frac{\text{Loss}^{\text{ADM}} - \text{Loss}^{\text{ADM},\delta}}{\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{ADM},\delta}}.$$
 (39)

Now, the unknown quantities  $\tilde{\sigma}_{I^{\text{TL}}}$ ,  $\tilde{\sigma}_{I^{\text{ADM}}}$ , and  $\tilde{\sigma}_{I^{\text{TL}},I^{\text{ADM}}}$ in (37) can be computed

$$\tilde{\sigma}_{I^{\mathrm{TL}}} = \tilde{\mu}_{I^{\mathrm{TL}}} - \tilde{\mu}_{I^{\mathrm{TL}}}^2 \tag{40}$$

$$\tilde{\sigma}_{I}_{ADM} = \tilde{\mu}_{I}_{ADM} - \tilde{\mu}_{I}^{2}_{ADM} \tag{41}$$

$$\tilde{\sigma}_{I^{\mathrm{TL}},I^{\mathrm{ADM}}} = E\{(I^{\mathrm{TL}}(T_{c},X) - \tilde{\mu}_{I^{\mathrm{TL}}})(I^{\mathrm{ADM}}(T_{c},X) - \tilde{\mu}_{I^{\mathrm{ADM}}})\}$$

$$= \underbrace{E\{I^{\mathrm{TL}}(T_c, X)I^{\mathrm{ADM}}(T_c, X)\}}_{\tilde{\mu}_{I^{\mathrm{TL}}I^{\mathrm{ADM}}}} - \tilde{\mu}_{I^{\mathrm{TL}}}\tilde{\mu}_{I^{\mathrm{ADM}}}$$
(42)

$$\tilde{\mu}_{I^{\mathrm{TL}}I^{\mathrm{ADM}}} = \int_{\Omega} I^{\mathrm{TL}}(T_c, X) I^{\mathrm{ADM}}(T_c, X) \tilde{f}(X) dX$$
(43)

$$=\frac{\int_{A}^{B} I^{\mathrm{TL}}(T_{c}, X) I^{\mathrm{ADM}}(T_{c}, X) f(X) dX}{\mathrm{Loss}^{\mathrm{ADM}, \varepsilon} - \mathrm{Loss}^{\mathrm{ADM}, \delta}}$$
(44)

where  $I^{TL}(T_c, X)$  and  $I^{ADM}(T_c, X)$  can be either 0 or 1. The integral in (44) is the integral of f(X) over the region where both the indicator variables are 1. Provided that the margin conditions hold, we can write the relationship below

$$\tilde{\mu}_{I^{\text{TL}}I^{\text{ADM}}} \leq \frac{\min(\text{Loss}^{\text{TL}}, \text{Loss}^{\text{ADM}}) - \text{Loss}^{\text{ADM},\delta}}{\text{Loss}^{\text{ADM},\varepsilon} - \text{Loss}^{\text{ADM},\delta}}.$$
 (45)

We substitute  $\tilde{\sigma}_{I^{\text{TL}}}$ ,  $\tilde{\sigma}_{I^{\text{ADM}}}$ , and  $\tilde{\sigma}_{I^{\text{TL}}I^{\text{ADM}}}$  computed above into (37)

$$\operatorname{Var}\{I^{\mathrm{TL}}(T_{c}, X) - I^{\mathrm{ADM}}(T_{c}, X)\}$$

$$\geq \tilde{\mu}_{I^{\mathrm{TL}}} - \tilde{\mu}_{I^{\mathrm{TL}}}^{2} + \tilde{\mu}_{I^{\mathrm{ADM}}} - \tilde{\mu}_{I^{\mathrm{ADM}}}^{2}$$

$$- 2(\tilde{\mu}_{I^{\mathrm{TL}}I^{\mathrm{ADM}}} - \tilde{\mu}_{I^{\mathrm{TL}}}\tilde{\mu}_{I^{\mathrm{ADM}}}).$$
(46)

Substituting the variables in (46) and combining with (36), we get

$$\operatorname{Var}\left\{\operatorname{Loss}_{N}^{\operatorname{CVIS}}\right\} \geq \frac{(\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta})^{2}}{N} \times \left(\frac{\operatorname{Loss}^{\operatorname{TL}} + \operatorname{Loss}^{\operatorname{ADM}} - 2\operatorname{min}(\operatorname{Loss}^{\operatorname{TL}}, \operatorname{Loss}^{\operatorname{ADM}})}{\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta}} - \frac{(\operatorname{Loss}^{\operatorname{TL}} - \operatorname{Loss}^{\operatorname{ADM}})^{2}}{(\operatorname{Loss}^{\operatorname{ADM},\varepsilon} - \operatorname{Loss}^{\operatorname{ADM},\delta})^{2}}\right).$$
(47)

Whether the minimum of the duple is Loss<sup>TL</sup> or Loss<sup>ADM</sup>, this expression turns into (35). 

Previously in this section, it was stated that the margin conditions require larger margins to be satisfied, whereas the convergence rate is affected negatively with the increasing margins. This can be observed from (35). When the margins are larger, the difference between the

# Algorithm 1 detectMargins (NS, SM, $T_c$ )

- 1. Generate NS sample points  $\{X_1, X_2, X_3, ..., X_{NS}\}$  from f(X).
- 2. For each  $X_i$ , compute  $d_C^{ADM}(X_i)$ .
- 3. Let  $X = \{Y_1, Y_2, Y_3, ..., Y_{NS}\}$  be the NS samples in decreasing order of  $d_C^{ADM}(Y_i)$ .
- 4. Let  $Y_{NP}$  is the element of X, which has the nearest  $d_C^{ADM}(Y_{NP})$  value to  $T_c$ .

5. i = NP - 1,  $SafeMargin_{\varepsilon} = 0$ 

- 6. j = NP + 1, SafeMargin<sub> $\delta$ </sub> =  $\emptyset$
- 7. //First compute the  $\varepsilon$  margin
- 8. while ( $|SafeMargin_{\varepsilon}| \leq SM$ ) do
- i = i + 19.
- if  $(d_C^{TL}(Y_i) < T_c)$  then 10.

11. 
$$SafeMargin_{\varepsilon} = SafeMargin_{\varepsilon} \cup \{Y_i\}$$

- 12. else
- 13.  $SafeMargin_{\varepsilon} = \emptyset$
- end if 14.
- 15. end while

16. 
$$i_{\varepsilon} = i - |SafeMargin_{\varepsilon}|$$
  
17.  $T_{\varepsilon} = 0.5(d_C^{ADM}(Y_{i_{\varepsilon}}) + d_C^{ADM}(Y_{i_{\varepsilon}+1}))$   
18.  $\varepsilon = T_c - T_{\varepsilon}$ 

18. 
$$\varepsilon = T_c - T_c$$

- 19. //Similarly compute the  $\delta$  margin
- 20. while ( $|SafeMargin_{\delta}| \leq SM$ ) do
- i = i 121.
- if  $(d_C^{TL}(Y_j) > T_c)$  then 22.

23. 
$$SafeMargin_{\delta} = SafeMargin_{\delta} \cup \{Y_j\}$$

- 24.
- $SafeMargin_{\delta} = \emptyset$ 25.
- 26. end if
- 27. end while
- 28.  $i_{\delta} = j + |SafeMargin_{\delta}|$ 29.  $T_{\delta} = 0.5(d_{C}^{ADM}(Y_{i_{\delta}}) + d_{C}^{ADM}(Y_{i_{\delta}-1}))$ 30.  $\delta = T_{\delta} - T_c$
- Loss<sup>ADM, $\varepsilon$ </sup> and Loss<sup>ADM, $\delta$ </sup> gets bigger, which causes an

increase in the variance of the CVIS estimator for a fixed N.

# C. Detection of $\varepsilon$ and $\delta$ Margins

We must compute a minimum  $\varepsilon$ -margin and a minimum  $\delta$ -margin that satisfy the two margin conditions, explained in Section III-B. In addition, this must be done without introducing much extra cost. In order to heuristically compute these two margins, we propose *detectMargins*, the pseudocode of which is shown in Algorithm 1.

*detectMargins* first starts by drawing NS sample points  $(X_i)$ from the joint probability distribution f(X). For each sample point  $X_i$ , the delay of the circuit is computed using fast ADM. After that it sorts the sample points according to their respective circuit delays computed by ADM. Here, note that loss point means that the circuit delay corresponding to that sample point exceeds the  $T_c$  value and a yield point is vice versa. As we utilize two different circuit delay computation methods, namely, ADM and TL simulations, a sample point may be



Fig. 1. Imaginary 15 point sample set example.

a loss point for a method but may be a yield point for another method.

Without loss of generality, assume that a sorted list of an imaginary 15 sample point set is shown in Fig. 1. The points are sorted according to their respective circuit delays computed by ADM. In addition, assume that NP (nearest point index) is seven for this imaginary set, and therefore,  $Y_7$  is the sample point whose ADM circuit delay is closest to the timing constraint  $T_c$ . The algorithm starts with  $Y_7$  computes the circuit delay using the TL simulation and then looks whether  $Y_7$  is a loss point for the TL simulation. It proceeds rightward in Fig. 1 and counts the number of consecutive sample points that are yield points for both ADM and TL simulations. This number corresponds to the number of elements in the  $SafeMargin_{\varepsilon}$ set. If this number exceeds SM, then the algorithm infers without going further that it reached inside the yield region where both  $I^{TL}(T_c, X)$  and  $I^{ADM}(T_c, X)$  are equal to 0. The algorithm comes back SM sample points and fixes  $T_{\varepsilon}$  at the middle of the delays corresponding to two sample points staying at the boundary. The  $\varepsilon$  value is determined according to this  $T_{\varepsilon}$  value, where  $T_{\varepsilon} = T_c - \varepsilon$ . After detecting  $\varepsilon$  margin, algorithm turns back to  $Y_7$  but this time it moves leftward and counts the number of consecutive points, which are loss points for both ADM and TL simulations. Similarly, when  $SafeMargin_{\delta}$  set exceeds SM points, the algorithm infers that the loss region, where both  $I^{TL}(T_c, X)$  and  $I^{ADM}(T_c, X)$ are equal to 1, is reached. Then, the algorithm computes  $T_{\delta}$ , which is equal to  $T_c + \delta$  as it does for  $T_{\varepsilon}$ . Fig. 1 shows the corresponding  $T_{\varepsilon}$  and  $T_{\delta}$  values for the imaginary 15 point sample set.

It can be seen that if the  $\varepsilon$  and  $\delta$  margins are computed by this heuristic algorithm, at least for the drawn sample points, both margin conditions in Section III-B can hold. Setting a bigger SM value increases the level of confidence. However, increasing SM may affect the performance of the algorithm, because this algorithm requires additional  $2 \times SM$ TL simulations, which will not be used in loss estimation, instead they are used for margin detection as explained. The remaining TL simulations are not additional cost, because they are required by the CVIS loss estimator. This is due the fact that the points used in the summation of (29) are drawn from  $\tilde{f}(X)$  in (25) and this distribution is nonzero only in between  $T_{\delta}$  and  $T_{\varepsilon}$ . We experimentally see that even for SM = 1, the results are promising, which will be quantified in Section IV.

Selection of a larger NS results in better loss estimates as expected. Only a portion of NS sample points requires TL simulations. Our experiments practically indicate that when PDM is used as an approximate delay model, this portion is about one-tenth of the selected NS for all benchmark circuits. One way similar to the devised in [31] would be to select NS, such that it is ten times the maximum number of affordable TL simulations. Alternatively, when there is a strict limit on the number of affordable TL simulations, an adaptive method to decide upon NS can be developed as follows. Initially, NS is set to be equal to the maximum number of affordable TL simulations. Until the number of TL simulations reaches this affordable amount, NS can adaptively be increased by drawing new sample points. The number of additional sample points can be determined by observing the ratio of the TL simulations and NS value at each iteration of the algorithm. This would require no repeated costly TL simulations provided that the TL simulation results for the previous sample points are stored.

# IV. RESULTS

In this section, the traditional STD-MC estimator and the previously proposed IS estimator [31] are compared with the CV and CVIS estimators proposed in this paper. Different experiments are carried out to record the convergence rate of the estimators. In addition, the theoretical variances derived in this paper, and the proposed heuristic algorithm is analyzed empirically.

# A. Experimental Setup

Among all process parameters, the most significant ones are gate length (L) and transistor threshold voltage  $(V_{\text{th}})$  [32]. The variation amounts are set according to the International Technology Roadmap for Semiconductors 2011 report [33], in which the effective gate length  $3\sigma/\mu$  ratio is 12%, whereas the threshold voltage  $3\sigma/\mu$  ratio is 20% for the 45-nm technology. Half of the variation comes from interdie and the other half from intradie components [34]. Spatial correlations are considered by using a four-level quad tree model [35], which models variations with spatial correlation utilizing 170 independent random variables for both L and  $V_{\text{th}}$ . The ISCAS-85 benchmark circuits [36] are synthesized using NanGate 45-nm cell library [37]. A modified version of NgSpice [38] simulator is used for TL SPICE simulations. Experiments are performed on a system with two Xeon E5-2620, six-core, 2-GHz processors, and 24 GB of RAM.

As an approximate delay model (ADM) to be used inside IS, CV, and CVIS estimators, we utilize the PDM in [31] that depends on very fast polynomial delay and slope equations. PDM can compute path delays in negligible times when compared with the TL SPICE simulations. Similarly, any fast delay model like the quadratic model in [39] that computes the delay of a gate given the random parameter values for that gate could be used.

For a fixed number of samples, the accuracy of a loss estimator depends on how far its estimates are from the actual loss. The actual loss, i.e., Loss<sup>TL</sup>, is computed using 50 000 sample STD-MC estimator shown in (3), which solely relies on precise

	$Loss^{TL} = 5\%$		$Loss^{TL} = 10\%$			$Loss^{TL} = 15\%$			$Loss^{TL} = 20\%$			
Speedup	CVIS	CV	IS	CVIS	CV	IS	CVIS	CV	IS	CVIS	CV	IS
c432	155.2	1.4	195.7	98.2	1.3	73.4	67.0	1.7	42.2	40.8	1.4	22.8
c499	259.1	1.9	182.7	137.2	2.4	77.9	151.8	1.9	62.5	92.7	2.6	34.5
c880	214.1	1.3	278.9	202.1	1.6	153.6	171.7	1.6	95.9	120.1	1.5	60.8
c1355	188.0	1.3	434.6	147.8	1.2	179.6	101.4	1.2	77.1	80.6	1.4	68.0
c1908	279.8	1.7	315.8	292.2	1.9	170.9	196.8	2.1	117.5	151.7	1.7	70.8
c2670	659.6	7.6	415.2	1237.4	5.9	492.9	980.0	6.6	217.4	1253.8	10.7	209.8
c3540	286.4	1.5	217.6	189.7	1.9	108.4	153.8	2.3	55.2	120.4	1.6	47.7
c5315	235.0	1.7	371.8	224.1	1.8	165.3	222.6	2.0	138.9	189.4	2.3	92.4
c7552	217.5	1.6	296.4	112.6	1.6	79.4	66.2	1.5	43.2	93.7	1.5	51.5
Avg.	277.2	2.2	301.0	293.5	2.2	166.8	234.6	2.3	94.4	238.1	2.7	73.2

TABLE I Speedup Over STD-MC for Different Estimators and Different Actual Loss Values

TL SPICE simulations. After having computed the actual loss value, in order to obtain the variance of an estimator around this actual loss value, we do the following. First, we perform loss estimations repetitively with the same number of samples but each time with a different sample set. We call the number of independent repetitions as R. The results constitute independent estimates of the same estimator using the same number of samples. Then, we compute the variance around the actual loss value over these R estimates. Not only the convergence rate but also the bias of the estimator, if there is any, can be understood by investigating the variance around the actual loss. For the sake of fairness, the estimators, whose variances will be compared with each other, must utilize the same number of sample points (TL simulations) for each estimate. By this way, the effect of the number of sample points can be excluded.

Let us assume two estimators called estimator A and estimator B. After fixing the number of samples/simulations, the ratio of the variance of estimator A over the variance of estimator B tells us how much faster estimator B is than estimator A, because it shows how many times less samples/simulations estimator B requires than estimator A for a fixed accuracy level [31]. Therefore, this ratio can be called sample reduction obtained by utilizing estimator B instead of estimator A. Similarly, we prefer calling it the speedup of estimator B with respect to estimator A. In Table I, we report the Speedup of the estimators with respect to the traditional STD-MC estimator. The speedup that we report in the table for an estimator can be interpreted in two ways.

- 1) It represents the reduction of the required number of TL circuit simulations to achieve the same predefined accuracy (error) with the STD-MC estimator.
- 2) Speedup shows how much smaller the variance of that estimator is than the STD-MC estimator when they both perform the same number of TL simulations.

# B. Performance of the Proposed Estimators

In our experiments, we empirically compare the performance of three estimators: 1) IS estimator in (6); 2) CV estimator in (11); and 3) CVIS estimator in (29). In all experiments, the SM parameter in Algorithm 1 is taken as 1, which is the smallest value that can be given to SM. The number of estimation repetitions, R, is set as 100 for each experiment. Four different experiments are performed at each of which a different timing constraint, and  $T_c$  value is set to four different values, such that the actual loss comes out to be 5%, 10%, 15%, and 20%, respectively. Table I shows the speedup over STD-MC for each experiment, for each estimator, and for each benchmark circuit. That is to say, it shows how many times less samples (TL simulations) than STD-MC are required by the proposed estimators for the same accuracy with the STD-MC estimator. For instance, for c1355 and when the actual loss is 10%, the CVIS estimator requires about 148 times less TL simulations than the traditional STD-MC estimator with the same accuracy. The results show that the proposed CVIS estimator is the fastest estimator with an average of  $260 \times$  speedup. The second best estimator is IS estimator with  $160\times$ , and the worst one is CV estimator with relatively very small speedup values. Considering that the computer setup given in the beginning of this section, 50000 sample STD-MC requires 278 h of computation when averaged over benchmark circuits, whereas if the CVIS estimator is preferred, this time decreases to about 1 h. Certainly, there is an additional cost in each of the IS, CV, and CVIS estimators because of computing  $d_C^{\text{ADM}}(X)$  values for all of the sample points, which is 50000 in our case. It is not considered in Table I, because ADM computations require only very low cost polynomial equation computations, which are negligible when compared with TL simulations and even can be performed in parallel with the simulations [31].

It can be seen from Table I that the performance of the IS estimator significantly decreases, while the actual loss increases. This rapid decrease trend is not observed for the CVIS estimator. CVIS estimator may result in the same, even better speedups with increasing actual loss values. This is another reason to prefer CVIS estimator over IS estimator.

Another observation is that the speedup values are better in some circuits than the others and this nearly holds whether the estimator is IS, CV, or CVIS. The ADM results in approximate delays, yet its accuracy can affect the speedup of the estimator utilizing it. Because each of the three estimators utilizes PDM as its ADM, the speedup follows a similar pattern for each of the three estimators.



Fig. 2. Bar graph of the loss estimated by ADM solely and the mean of the CVIS estimator for each benchmark, where actual loss is 10%.

Comparison of Theoretical Variance and Practical Variance for CV Estimator



Fig. 3. Theoretical versus experimental variance for CV estimator.

One could wonder the performance of an MC estimator that is based on PDM solely. Actually, we have used such an estimator to compute Loss<sup>ADM</sup> that is required for CV and CVIS estimators, as explained in Sections III-A and III-B. This Loss<sup>ADM</sup> estimator is the same one shown in (8) with one difference that there is no  $\varepsilon$  margin, i.e.,  $T_c^{\varepsilon}$  is replaced with  $T_c$ . The Loss<sup>ADM</sup> estimator is very fast as it requires no TL simulations. However, PDM is an approximate method and lacks accuracy in estimating timing loss. This is shown in Fig. 2. Fig. 2 (light gray bars) shows the Loss<sup>ADM</sup> values, computed using the 50 000 sample PDM-based MC estimator. The actual loss in the experiment is 10%. It is clearly seen that the Loss<sup>ADM</sup> estimates that are computed by the PDM-based MC are far off the actual loss, which shows that PDM alone cannot provide accurate results.

Fig. 2 also shows the mean of the CVIS estimator over 100 repetitive estimations for each benchmark circuit and for again 10% actual loss experiment. The mean of the CVIS estimator shown by dark gray bar is almost equal to the actual loss for each of the benchmark circuits. This empirically shows that the CVIS estimator is unbiased and our heuristic algorithm explained in Section III-C worked well even when SM is set to its smallest value, which is 1.

## C. Verification of the Theoretical Variance

In this section, we verify the theoretical variance derivation in Section III with the empirical data. Fig. 3 compares the theoretical variance of CV estimator with the experimental variance results for the 10% actual loss experiment. Fig. 4 is the same plot corresponding to CVIS estimator. The *x*-axis shows the number of samples utilized in the estimation and *y*-axis is the resultant variance. In these figures, the theoretical

TABLE II EFFECT OF SM ON SPEEDUP

	$Loss^{TL} = 10\%$							
Speedup	SM = 1	SM = 2	SM = 3	SM = 4	SM = 5			
c432	98.2	102.4	114.1	107.1	109.2			
c499	137.2	134.2	127.1	112.8	108.5			
c880	202.1	186.5	165.4	146.9	141.6			
c1355	147.8	133.4	128.4	112.5	111.0			
c1908	292.2	280.0	235.8	225.9	203.4			
c2670	1237.4	873.0	724.7	642.5	563.8			
c3540	189.7	170.5	164.6	142.6	124.0			
c5315	224.1	195.0	177.6	165.1	155.5			
c7552	112.6	147.9	143.1	155.3	132.2			
Avg.	293.5	247.0	220.1	201.2	183.3			

variance values of the corresponding estimator are plotted in circles, and the experimental variance values, which are computed by averaging the sum of the euclidian distances of the estimates around the actual loss value, are plotted in asterisks. For CV estimator, we plotted the theoretical and experimental variance values only for c1355 and c3540 circuits, but other benchmark circuits produce similar plots. The plots not only illustrate the inverse relation between the number of samples and the variance of the resultant estimates, but also experimentally verify the accuracy of the theoretical CV and CVIS estimator variances shown by (17) and (35).

## D. Quantification of the Heuristic Algorithm

In this section, the heuristic algorithm and the effect of SM on the resultant speedup of CVIS estimator are investigated. Table II shows the speedup of CVIS estimator over STD-MC for 10% actual loss experiment. It should be reminded that we compute variance around the actual loss for both STD-MC and CVIS estimators to get the speedup values in Tables I and II. Therefore, any bias of the CVIS estimator is taken into consideration for the reported speedup amounts. Table II clearly shows the slow reduction trend of the speedup due to the increasing SM values. This is expected as increasing SM increases the additional overhead of the heuristic algorithm. Another observation is that the speedup very slowly increases for c432 while going from SM = 1 to SM = 2 and SM = 3. The bias of CVIS estimator for c432 and SM = 1experiment is 0.002 as the mean of the CVIS estimator for this experiment is 9.8% instead of 10%. For this case, increasing SM decreases this bias and overcomes the speedup reduction due to SM increase. This rare situation can also be seen for some other experiments like c7552 while switching from SM = 1 to SM = 2 experiment.

The optimal decision of SM value depends on the ability of the selected ADM to capture the variation of circuit delay with respect to the variation of the statistical circuit parameters. In this paper, we select PDM as our approximate delay model and our SM = 1 experiments show that the bias of the CVIS estimator, which is worst (0.002) for c432 circuit, is negligible, as shown in Fig. 2. In addition, the speedup results in Table II experimentally show us that when PDM is used, it is better to set SM to lower values and tolerate the negligible bias introduced by low SM values, instead of increasing SM for removing all bias.



Comparison of Theoretical Variance and Practical Variance for CVIS Estimator

Fig. 4. Theoretical versus experimental variance of CVIS estimator for the benchmark circuits where actual loss is 10%. x-axis is the number of TL simulations used for estimation and y-axis is the amount of variance.

TABLE III Margin Values in Picoseconds Averaged Over 100 Sets and While SM Is Increased

		$Loss^{TL} = 10\%$							
Average ε, δ		SM = 1	SM = 2	SM = 3	SM = 4	SM = 5			
	δ	1.9	1.9	1.9	1.9	1.9			
c432	3	51.5	53.1	53.9	54.5	54.8			
	δ	1.0	1.0	1.0	1.0	1.0			
c499	3	21.7	22.7	23.1	23.3	23.5			
	δ	1.4	1.4	1.4	1.4	1.4			
c880	3	39.8	40.4	40.8	41.0	41.6			
	δ	2.3	2.3	2.3	2.3	2.3			
c1355	3	57.3	57.9	58.1	58.4	58.5			
	δ	1.5	1.5	1.5	1.5	1.5			
c1908	З	46.5	47.2	47.5	47.8	48.4			
	δ	1.3	1.3	1.3	1.3	1.3			
c2670	З	12.1	12.5	12.8	12.8	12.8			
	δ	2.3	2.3	2.3	2.3	2.3			
c3540	З	49.2	51.0	52.4	53.8	54.8			
	δ	2.6	2.6	2.6	2.6	2.6			
c5315	3	67.3	68.4	68.8	69.0	69.1			
	δ	2.7	2.7	2.7	2.7	2.7			
c7552	3	71.4	73.2	73.9	74.6	74.8			

Increasing SM value does not require to repeat the experiment, and only additional simulations must be performed according to SM increase. As a result, the user can implement heuristic algorithm in an adaptive manner, such that SM is increased one by one until a maximum number of affordable simulations is reached or the resultant  $\varepsilon$  and  $\delta$  margins converge. The convergence of the margins does not guarantee but enhances confidence that there is no need to increase SM further. Another improvement would be to use different safety points for  $\varepsilon$  and  $\delta$  margins, so that if the algorithm has already converged in  $\delta$  margin, then all remaining available TL simulations can be used to compute  $\varepsilon$  margin and vice versa.

Table III contains the  $\varepsilon$  and  $\delta$  margin values in picoseconds corresponding to different SM values. Please note that the timing constraint  $(T_c)$  values are in a range between 700 and 1600 ps for the benchmark circuits. The resultant  $\delta$  margins are the same for all SM values. On the other hand, rising SM results in an increase in  $\varepsilon$  margins, which exhibits a convergence as expected. The unbalanced  $\varepsilon$  and  $\delta$ margins are the result of the fact that PDM, as an approximate delay model, has a delay bias with respect to the circuit delays computed by TL simulations. Our algorithm clearly handles this imbalance. For, especially, large SM values, it is clear from the table that using different numbers of safety points for computing  $\varepsilon$  and  $\delta$  could be advantageous, as suggested above.

## V. CONCLUSION

A novel methodology to combine CV with IS is presented for the solution of timing yield estimation problem, and it is both theoretically and practically verified in this paper. CV and CVIS estimators are constructed based on CV and the combination of CV with IS, respectively. ISCAS-85 benchmark circuits are relatively simple containing about 1200 gates on the average and 3500 gates at most. Using CVIS estimator at the final stage verification still requires hours of computations, but instead of hundreds of hours that STD-MC would require. The experimental results show that the resultant CVIS estimator takes one step further than the IS estimator, and it can accelerate the traditional STD-MC estimator about 260 times on the average. The distinctive method to combine CV with IS and the derivations proposed in this paper can be used in different aspects to gain faster estimators. Moreover, other variance reduction techniques, such as stratified sampling, can further be combined with the techniques in the paper.

#### ACKNOWLEDGMENT

The author would like to thank S. Tasiran and A. Demir for their valuable advice and suggestions.

#### REFERENCES

- D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, "Statistical timing analysis: From basic principles to state of the art," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 4, pp. 589–607, Apr. 2008.
- [2] C. Forzan and D. Pandini, "Statistical static timing analysis: A survey," *Integr., VLSI J.*, vol. 42, no. 3, pp. 409–435, Jun. 2009.
- [3] A. Gupta. Variation Aware Custom IC Design Report 2011. [Online]. Available: http://www.solidodesign.com/files/variation-awarecustom-design-survey-2011.pdf, accessed Feb. 2, 2016.
- M. LaPedus, TSMC Rolls 40-nm Design Flow, EETimes, Jun. 2008.
   [Online]. Available: http://www.eetimes.com/document.asp?doc \_id=1168660
- [5] N. Metropolis, "The beginning of the Monte Carlo method," Los Alamos Sci., vol. 15, no. 584, pp. 125–130, 1987.
- [6] N. Metropolis and S. Ulam, "The Monte Carlo method," J. Amer. Statist. Assoc., vol. 44, no. 247, pp. 335–341, 1949.
- [7] H. Kahn and T. E. Harris, "Estimation of particle transmission by random sampling," *Nat. Bureau Standards Appl. Math. Ser.*, vol. 12, pp. 27–30, Jun. 1951.
- [8] E. C. Fieller and H. O. Hartley, "Sampling with control variables," *Biometrika*, vol. 41, pp. 494–501, Dec. 1954.
- [9] J. M. Hammersley and D. C. Handscomb, Monte Carlo Methods [Methuen's Monographs on Applied Probability and Statistics]. London, U.K.: Methuen Publishing, 1965.
- [10] B. L. Nelson, "On control variate estimators," Comput. Oper. Res., vol. 14, no. 3, pp. 219–225, 1987.
- [11] N. Madras, *Lectures on Monte Carlo Methods* (Fields Institute for Research in Mathematical Sciences Toronto: Fields Institute monographs). Providence, RI, USA: AMS, 2002.
- [12] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*. New York, NY, USA: Wiley, 2008.
- [13] R. Y. Rubinstein and D. P. Kroese, Simulation and the Monte Carlo Method, vol. 707. New York, NY, USA: Wiley, 2011.
- [14] A. Singhee, S. Singhal, and R. A. Rutenbar, "Practical, fast Monte Carlo statistical static timing analysis: Why and how," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2008, pp. 190–195.
- [15] L. Scheffer, "The count of Monte Carlo," in Proc. ACM/IEEE Int. Workshop Timing Issues Specification Synth. Digit. Syst. (TAU), Feb. 2004, pp. 1–6.
- [16] D. E. Hocevar, M. R. Lightner, and T. N. Trick, "A study of variance reduction techniques for estimating circuit yields," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 2, no. 3, pp. 180–192, Jul. 1983.
- [17] R. S. Soin and P. J. Rankin, "Efficient tolerance analysis using control variates," *IEE Proc. G, Electron. Circuits Syst.*, vol. 132, no. 4, pp. 131–142, Aug. 1985.
- [18] P.-F. Desrumaux et al., "An efficient control variates method for yield estimation of analog circuits based on a local model," in Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD), New York, NY, USA, Nov. 2012, pp. 415–421.
- [19] J. Jaffari and M. Anis, "On efficient LHS-based yield analysis of analog circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 1, pp. 159–163, Jan. 2011.
- [20] J. Jaffari and M. Anis, "Advanced variance reduction and sampling techniques for efficient statistical timing analysis," *IEEE Trans. Comput. Aided Design Integr. Circuits Syst.*, vol. 29, no. 12, pp. 1894–1907, Dec. 2010.
- [21] H. Niederreiter, "Quasi-Monte Carlo methods and pseudo-random numbers," Bull. Amer. Math. Soc., vol. 84, no. 6, pp. 957–1041, 1978.
- [22] J. Jaffari and M. Anis, "Timing yield estimation of digital circuits using a control variate technique," in *Proc. Quality Electron. Design (ISQED)*, Mar. 2009, pp. 382–387.
- [23] J. Jaffari and M. Anis, "Practical Monte-Carlo based timing yield estimation of digital circuits," in *Proc. Design, Autom., Test Eur. Conf. Exhibit.*, 2010, pp. 807–812.

- [24] V. Veetil, D. Sylvester, and D. Blaauw, "Efficient Monte Carlo based incremental statistical timing analysis," in *Proc. 45th ACM/IEEE Annu. Design Autom. Conf (DAC)*, Jun. 2008, pp. 676–681.
- [25] V. Veetil, K. Chopra, D. Blaauw, and D. Sylvester, "Fast statistical static timing analysis using smart Monte Carlo techniques," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 6, pp. 852–865, Jun. 2011.
- [26] L. Dolecek, M. Qazi, D. Shah, and A. Chandrakasan, "Breaking the simulation barrier: SRAM evaluation through norm minimization," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2008, pp. 322–329.
- [27] R. Kanj, R. Joshi, and S. Nassif, "Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events," in *Proc. 43rd ACM/IEEE Design Autom. Conf. (DAC)*, Jul. 2006, pp. 69–72.
- [28] A. Singhee and R. A. Rutenbar, "Why quasi-Monte Carlo is better than Monte Carlo or Latin hypercube sampling for statistical circuit analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 29, no. 11, pp. 1763–1776, Nov. 2010.
- [29] A. Singhee and R. A. Rutenbar, "Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 28, no. 8, pp. 1176–1189, Aug. 2009.
- [30] G. Chen, D. Sylvester, D. Blaauw, and T. Mudge, "Yield-driven nearthreshold SRAM design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 11, pp. 1590–1598, Nov. 2010.
- [31] A. A. Bayrakci, A. Demir, and S. Tasiran, "Fast Monte Carlo estimation of timing yield with importance sampling and transistor-level circuit simulation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 29, no. 9, pp. 1328–1341, Sep. 2010.
- [32] D. Sylvester, K. Agarwal, and S. Shah, "Invited paper: Variability in nanometer CMOS: Impact, analysis, and minimization," *Integr., VLSI J.*, vol. 41, no. 3, pp. 319–339, May 2008.
- [33] International Technology Roadmap for Semiconductors (ITRS) Report 2011 Edition. [Online]. Available: http://www.itrs.net/, accessed Feb. 2, 2016.
- [34] J. Cong, P. Gupta, and J. Lee, "Evaluating statistical power optimization," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 29, no. 11, pp. 1750–1762, Nov. 2010.
- [35] A. Agarwal, D. Blaauw, and V. Zolotov, "Statistical timing analysis for intra-die process variations with spatial correlations," in *Proc. Int. Conf. Comput. Aided Design (ICCAD)*, 2003, pp. 900–907.
- [36] F. Brglez and H. Fujiwara, "A neutral netlist of 10 combinational benchmark circuits and a target translator in FORTRAN," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 1985, pp. 695–698.
- [37] NanGate 45 nm Open Cell Library. [Online]. Available: http://www. nangate.com/, accessed Feb. 2, 2016.
- [38] NgSpice Circuit Simulator. [Online]. Available: http://ngspice. sourceforge.net/, accessed Feb. 2, 2016.
- [39] V. Khandelwal and A. Srivastava, "A quadratic modeling-based framework for accurate statistical timing analysis considering correlations," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 15, no. 2, pp. 206–215, Feb. 2007.



Alp Arslan Bayrakci (M'10) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2004, and the Ph.D. degree in computer engineering from Koç University, Istanbul, Turkey, in 2010.

He has been with Gebze Technical University, Gebze, Turkey, since 2011. His current research interests include statistical timing analysis, hardware security, and computer-aided design methodologies.