# Stochastic Behavioral Modeling and Analysis for Analog/Mixed-Signal Circuits

Fang Gong, *Student Member, IEEE,* Sina Basir-Kazeruni, *Student Member, IEEE,* Lei He, *Senior Member, IEEE,* and Hao Yu, *Member, IEEE*

***Abstract*—It has become increasingly challenging to model the stochastic behavior of analog/mixed-signal (AMS) circuits under large-scale process variations. In this paper, a novel moment-matching-based method has been proposed to accurately extract the probabilistic behavioral distributions of AMS circuits. This method first utilizes Latin hypercube sampling coupling with a correlation control technique to generate a few samples (e.g., sample size is linear with number of variable parameters) and further analytically evaluate the high-order moments of the circuit behavior with high accuracy. In this way, the arbitrary probabilistic distributions of the circuit behavior can be extracted using moment-matching method. More importantly, the proposed method has been successfully applied to high-dimensional problems with linear complexity. The experiments demonstrate that the proposed method can provide up to 1666X speedup over crude Monte Carlo method for the same accuracy.**

***Index Terms*—Behavioral modeling, Latin hypercube sampling, Monte Carlo method, process variation.**

## I. Introduction

LARGE-SCALE process variations are inevitable in the nanotechnology era [1], [2] and significantly change the behavior of analog/mixed-signal (AMS) circuits (e.g., voltage swing, timing delay, and clock frequency) [3]–[9]. Therefore, it is urgently sought to accurately extract the probabilistic behavioral distribution of AMS circuits under process variations.

In general, there are two types of process variation sources: systematic global variation and local random variation. In this paper, we focus on the local variation that is purely random and more difficult to model. The most straightforward approach is crude Monte Carlo (MC) method [10], which utilizes massive samples and expensive simulation program with integrated circuit emphasis (SPICE) simulations to evaluate the probabilistic distributions [e.g., probability density function (PDF) and cumulative distribution function (CDF)] of circuit behavior. MC method can be easily applied to any variable parameter and circuit behavior with arbitrary distributions. However, it is too time consuming and not affordable.

Many statistical methods have been developed in the past few years, including the linear regression method [1], which models the circuit behavior as a linear function of a number of normally distributed process variables and thus becomes inaccurate for strongly nonlinear circuits. The work in [9] and [11] can estimate the unknown distribution of circuit behavior with stochastic orthogonal polynomials but requires prior knowledge of the distribution type that is unavailable in practice.

In addition, asymptotic waveform evaluation (AWE) [12] approximates the "arbitrary" circuit behavioral distribution with the impulse response of a linear time-invariant (LTI) system by matching a few high-order moments. This approach requires no prior knowledge of the circuit behavioral distribution but needs expensive computational efforts to evaluate the high-order moments. Note that our work in this paper is based on AWE framework but proposes a novel method to calculate high-order moments efficiently with high accuracy.

To resolve this issue of AWE [12], response surface method (RSM)-based methods [13], [14] have been proposed to model the circuit behavior as a polynomial function of all variable process parameters and further evaluate the high-order moments. For example, asymptotic probability extraction [13] evaluates the RSM model using ordinary least-square (OLS) regression method so that the number of needed SPICE simulations equals the total number of unknown coefficients in the polynomial function of the RSM model. Moreover, a novel approach has been proposed [14] to extract all unknown coefficients of RSM model from a small number of samples with regularization-based regression method. In fact, this approach finds a unique sparse solution of an underdetermined equation system using $L_0$-norm regularization [15].

However, these RSM-based methods have been plagued by the following issues. First, fully nonlinear AMS circuits tend to need higher order RSM models (e.g., strongly nonlinear functions of random process variables) where the number of unknown coefficients and required SPICE simulations in OLS-based RSM method can increase exponentially, thereby, making the OLS-based RSM method infeasible. Second, the regularization-based regression method [14] suffers from bias-variance tradeoff [16] that can potentially degrade the accuracy and robustness of the extracted RSM models. Third, when a large number of process variables are considered, the

Manuscript received August 29, 2011; revised May 11, 2012; accepted July 29, 2012. Date of current version December 19, 2012. This paper was recommended by Associate Editor A. Pick.

F. Gong, S. Basir-Kazeruni, and L. He are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: fang08@ee.ucla.edu; sinabk@ee.ucla.edu; lhe@ee.ucla.edu).

H. Yu is with the School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore (e-mail: haoyu@ntu.edu.sg).

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/TCAD.2012.2217961

RSM model becomes highly complicated and requires more computational efforts. Therefore, an efficient and accurate method to evaluate high-order moments and further extract the stochastic behavior of AMS circuits is still urgently sought.

In this paper, we propose a novel and efficient algorithm to accurately predict the arbitrary probabilistic distributions of circuit behavior based on AWE [12]. This approach first utilizes Latin hypercube sampling (LHS) method along with correlation control technique to generate a few samples (e.g., sample size is linear with number of variable parameters) and further evaluates the high-order moments accurately with analytical formulas. Then, the PDF/CDF of stochastic circuit behavior can be recovered using conventional moment-matching method in AWE. In addition, a normalized PDF function is introduced to enhance the accuracy by eliminating the potential numerical errors. The experiments demonstrate that our proposed method provides very high accuracy along with up to 1666X speedup when compared with MC.

It is worth noting the benefits that the proposed work can offer.

1) This method does not need RSM models to estimate the moments and, therefore, avoids the aforementioned exponential complexity and bias-variance tradeoff.
2) The proposed method can handle strongly nonlinear AMS circuits and high-dimensional problems with a large number of random process variables.
3) This approach can achieve nearly linear complexity while providing highly accurate behavioral distributions.

The remainder of this paper is organized as follows. In Section II, we present the necessary background knowledge. Section III describes the high-order moments estimation. The PDF/CDF calculation is presented in Section IV. Section V summarizes the overall algorithm. The experiments are provided in Section VI. This paper is concluded in Section VII.

## II. Background

### A. Mathematical Formulation

Assume $\vec{x} = [x_1, x_2, \cdots]$ is a vector of random variables (e.g., threshold voltage, channel length, and gate oxide capacitance) and can be characterized by a sequence of probabilistic distributions $[\text{pdf}(x_1), \text{pdf}(x_2), \ldots]$, where $\text{pdf}(x_i)$ is the PDF function associated with the element $x_i$ of $\vec{x}$. These random variables can be fed into SPICE simulator engines, and the output is the circuit behavior $y$ (e.g., voltage, bandwidth, and power) as follows.

$$\underbrace{\vec{x}}_{\text{variable}} \Rightarrow \boxed{\text{SPICE simulators}} \Rightarrow \underbrace{y}_{\text{circuit behavior}}. \tag{1}$$

Clearly, there exist two spaces: "parameter space" that contains all possible values of $\vec{x}$ and "performance space" that has all possible values of $y$. In fact, there is a mapping from the parameter space to the performance space so that each sample of $\vec{x}$ has its corresponding $y$. Mathematically, the mapping can be viewed as an implicit function $y = f(\vec{x})$, which, unfortunately, has no analytical formula. Therefore, our aim is to determine the unknown probabilistic distribution of $y$ that results from the uncertainties in $\vec{x}$.

To this purpose, the high-order moments of $y$ need to be evaluated and then the probabilistic distributions of $y$ can be recovered by AWE method as proposed in [12]. According to probability theory [17], [18], the $p$th-order probabilistic moments of $y$ can be defined as

$$m_y^p = E(y^p) = \int_{-\infty}^{+\infty} (y^p \cdot \text{pdf}(y))dy \tag{2}$$

where $\text{pdf}(y)$ is the PDF function of $y$ and $m_y^p$ is the $p$th probabilistic moment of $y$.

For illustration purposes, we introduce a significant observations of AWE method [12] as follows.

*Property 1:* The low-order moments are more important to achieve high accuracy when moments $m_y^p$ $(p = 1, \ldots, +\infty)$ are used to recover $\text{pdf}(y)$.

*Proof:* Let $\Phi(\omega)$ be the Fourier transform of $\text{pdf}(y)$ as (detailed derivation can be referred to [19])

$$\Phi(\omega) = \sum_{p=0}^{+\infty} \frac{(-j\omega)^p}{p!} \cdot m_y^p. \tag{3}$$

Equation (3) is equivalent to the Taylor expansion of $\Phi(\omega)$ at the expansion point $\omega = 0$ [13], and the high-order moments are related to the coefficients of the Taylor expansion. It is well known that Taylor expansion linearizes the function around the expansion point and thus (3) provides high accuracy around $\omega = 0$. As such, the low-order moments (around $\omega = 0$) are more important for the accuracy of both the expansion in (3) and approximated $\text{pdf}(y)$.

From another point of view [13], the magnitude of moments (coefficients) reaches its maximum at $\omega = 0$ as $\Phi(0) = 1$ and decays as $\omega$ increases, which behaves as a low-pass filter. Therefore, the low-frequency band is much more important for approximation accuracy that is mainly determined by the low-order moments ($\omega$ around 0). ■

### B. Moment Matching for PDF Calculation

We will briefly review the method to extract $\text{pdf}(y)$ with probabilistic moments $\{m_y^p\}$ [12], [13]. First, "time moments" for $y$ can be defined as

$$\hat{m}_y^k = \frac{(-1)^k}{k!} \cdot \int_{-\infty}^{+\infty} y^k \cdot \text{pdf}(y)dy. \tag{4}$$

It is clear that $\hat{m}_y^k$ is different from $m_y^k$ in (2) due to a scaling factor $(-1)^k/k!$. In the mean time, consider a LTI system $H$ whose time moments defined as [12]

$$\hat{m}_t^k = \frac{(-1)^k}{k!} \cdot \int_{-\infty}^{+\infty} t^k \cdot h(t)dt \tag{5}$$

where $t$ is the time variable and $h(t)$ is the impulse response of the LTI system $H$.

One important observation is that impulse response $h(t)$ in (5) can be an optimal approximation to $\text{pdf}(y)$ in (4) if we treat $t$ in (5) as $y$ in (4) and make $\hat{m}_t^k$ equal to $\hat{m}_y^k$ (i.e., moment-matching technique).

**Algorithm 1** Overall Algorithm for PDF calculation

1: Input probabilistic distributions of variable parameters.
2: **/* Step 1: Moment Calculation */**
3: Calculate the time moments of observations as $\widehat{m}_y^k$ with (4) in the performance space.
4:
5: **/* Step 2: Moment Matching */**
6: Make $\widehat{m}_y^k$ equal to $\widehat{m}_t^k$ by matching first several moments.
7: Solve the resulting nonlinear equation system in (6) for residues $a_r$ and poles $b_r$.
8:
9: **/* Step 3: PDF Calculation */**
10: Compute impulse response $h(t)$ in (7) with residues $a_r$ and poles $b_r$.
11: Use $h(t)$ as the optimal approximation of $pdf(y)$.

Moreover, according to probability theory [12], [17], [18], the time moments in (5) can be further expressed as

$$\widehat{m}_t^k = -\sum_{r=1}^{M} \frac{a_r}{b_r^{k+1}} \tag{6}$$

where $a_r$ and $b_r$ $(r = 1, \ldots, M)$ are the residues and poles of this LTI system, respectively. As such, the impulse response of the LTI system can be evaluated as

$$h(t) = \begin{cases} \sum_{r=1}^{M} a_r \cdot e^{b_r \cdot t} & (t \geq 0) \\ 0 & (t < 0) \end{cases} \tag{7}$$

The overall algorithm that calculates $h(t)$ as an optimal approximation to pdf$(y)$ can be summarized as follows.

The most challenging step is to evaluate high-order moments $m_y^k$ and time moments $\widehat{m}_y^k$ in the performance space. In this paper, we proposed a novel and efficient algorithm to evaluate these high-order moments without response surface model (RSM) but with high accuracy.

## III. High-Order Moments Estimation

### A. Moments Via Point Estimation

Usually, it is impractical to compute $m_y^k$ as (2) because pdf$(y)$ is unknown. Instead, "point estimation method" proposed in [20], [21] approximates $m_y^k$ by a weighted sum of several sampling values of $y$. For example, assume $x$ is the only variable and $x_j$ $(j = 1, \ldots, p)$ are estimating points of $x$ with weights $P_j$, the $k$th-order probabilistic moment of $y$ can be approximated as

$$m_y^k = \int_{-\infty}^{+\infty} y^k \cdot \text{pdf}(y) dy \approx \sum_{j=1}^{p} P_j \cdot y_j^k = \sum_{j=1}^{p} P_j \cdot f(x_j)^k. \tag{8}$$

The works in [20] and [21] only provide empirical analytical formulas of $x_j$ and $P_j$ for first four moments (e.g., the mean, the variance, the skewness, and the kurtosis) and thus cannot satisfy the requirement of AWE, where higher order moments (e.g., $k \gg 4$) are needed. To this end, an systematical approach has been established in [19] to efficiently calculate the estimating points $x_j$ and weights $P_j$ for arbitrary-order moments.

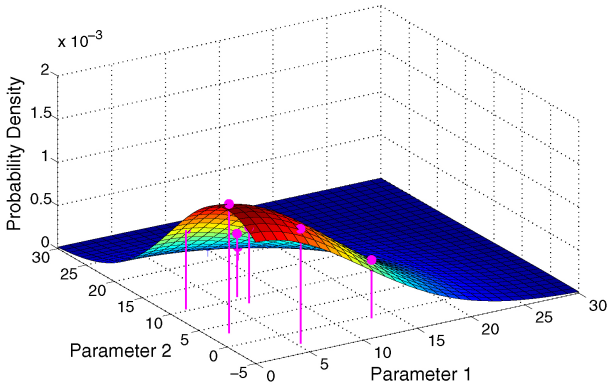



Fig. 1. Probability density map and "representative" sampling points.

However, all these approaches [19]–[21] can only simply handle low-dimensional problems. Therefore, it is significant but remains unknown how to evaluate high-order moments for high-dimensional problems (e.g., tens or hundreds of variables) that is the motivation behind this work.

### B. Basic Idea of Moments Via Sampling Method

The integral of $m_y^k$ in (2) is very difficult to compute because an analytical evaluation is not available. Therefore, it is inevitable to utilize a sampling method. In fact, "point estimation method" is a sampling-like method, which picks a few "representative" samples $y_j$ in (8) and weights them by $P_j$ to approximate the integral value.

Inspired by this observation, our proposed method tries to choose a few samples as "good representatives" of the entire sampling space so that a huge number of samples can be saved. For example, Fig. 1 shows a probability density map consisting of two normal distributed variables. Note that only part of the probability density map is plotted in order to show the interior "representative" sampling points.

Since the integral of $m_y^k$ in (2) is defined over the entire space shown in Fig. 1, the most straightforward sampling method is to generate as many samples as possible spreading over the "entire" sampling space that is infeasible due to unaffordable computational efforts.

Instead, the proposed method chooses a few "representative" samples $\vec{x}_j = [x_{1,j}, x_{2,j}, \ldots]$ shown as marked stems in Fig. 1, which should satisfy the following conditions.

1) The samples of each element of $\vec{x}$ (e.g., $x_i$) should follow its known marginal distribution (e.g., pdf$(x_i)$).
2) Various correlations and other relationships between the elements of $\vec{x}$ should remain intact.
3) These chosen samples should fully cover the entire sampling space to provide closer approximation.
4) These samples should be incoherent so that the number of required samples can be kept to be the minimum.

To meet above requirements, we propose to leverage well-established LHS method [22] along with correlation control technique [23] as discussed in Section III-C.

### C. LHS and Correlation Control

1) *Latin Hypercube Sampling:* The LHS method [22] is a widely used variant of MC method that can "efficiently" generate samples. LHS method first divides the cumulative

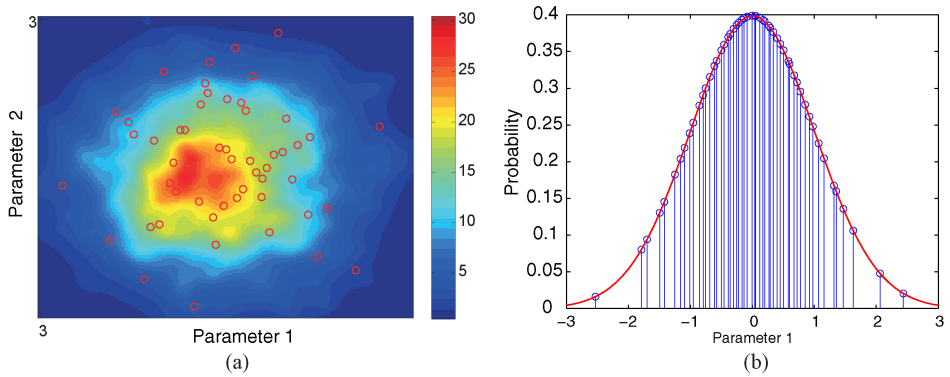



Fig. 2. Samples from LHS method. (a) Probability density map and LHS samples. (b) LHS samples in 1-D.

distribution of each random variable into several intervals with equal probability and picks one sample from each interval randomly. Then, LHS transforms these samples into the desired probabilistic distribution using inverse CDF. As such, the samples for each variable can be paired randomly to generate LHS samples. Note that LHS is "efficient" because each random variable will be sampled only once from each of its intervals. Thereby, LHS method can use a small number of samples to ensure a full coverage of the sampling space.

For example, we plot some LHS samples in the probability density map consisting of two standard normally distributed variables in Fig. 2(a), where all samples are dispersed over the entire parameter space and there are no duplicate/overlapped samples. In addition, all samples have been projected into a 1-D space in Fig. 2(b), which clearly demonstrates that all samples follow the known marginal distribution $N(0, 1)$ and fully cover the entire sampling space. Moreover, it can be observed from Fig. 2(b) that there are no duplicate samples, which implies that any two different LHS samples have different values for the same random variable so that these two LHS samples are incoherent.

Therefore, LHS samples can meet all requirements mentioned in Section III-B except for the condition of correlation, and thus correlation control technique [23] is needed.

2) *Correlation Control Technique:* The distributions of individual elements in $\vec{x}$ that can be characterized by a correlation matrix $C$ can be correlated (e.g., entry $C_{ij} \in [-1, 1]$ is the correlation coefficient between $x_i$ and $x_j$ variables). As an illustrative example, we consider random variables in $\vec{x}$ to be independent where the target correlation matrix becomes an identity matrix.

The conventional sampling scheme is to generate samples for individual variable $x_i$ independently and then pair them randomly (as combinations) to produce samples of $\vec{x}$. In particular, we are more interested in the case when sample size $N$ is small (e.g., $N$ is linear with the number of variables), as there is an important observation in this case.

*Property 2:* The conventional sampling scheme can provide samples with desired correlation only if the sample size is large enough. When sample size is small, the generated samples have WRONG correlation relationship.

*Proof:* The reason behind this observation is when sample size is small, arbitrary correlations between individual elements of $\vec{x}$ can be introduced during the pairing/combination stage. However, this phenomena disappears as the sample size increases, since large sample size can ensure a close approximation to purely random combination.

As an example, we can consider two "independent" random variables that follow standard normal distribution and, ideally, the correlation matrix should be an identity. We compare the correlation matrices with different numbers of samples as follows:

$$\underbrace{\begin{pmatrix} 1 & -0.0082 \\ -0.0082 & 1 \end{pmatrix}}_{100 \text{ samples}} \qquad \underbrace{\begin{pmatrix} 1 & -0.73 \\ -0.73 & 1 \end{pmatrix}}_{4 \text{ samples}}.$$

Clearly, the correlation matrix extracted from 100 samples is closer to an identity matrix so that samples of two independent variables are obtained, while four samples introduce a strong correlation between these two independent random variables. ■

Is there any way to fix the correlation issue of cases with small sample size? The short answer is positive, and we will now discuss the reasoning for this answer.

Since the incorrect correlation is introduced in the pairing/combination stage, it is a natural choice to "re-pair or recombine" samples for different random variables to achieve the correct correlation, and thus the correlation control technique developed by Iman and Conover [23] can be used. The theoretical basis can be briefly described as follows.

Let us consider $m$ random variables and the desired correlation matrix $C$ that is positive definite and symmetric. First, $n$ random samples can be generated for each variable and we can build a $n \times m$ matrix denoted as $X$, where $X_{ij}$ is the $i$th sample for the $j$th variable. We assume $R$ is the correlation matrix extracted from these $n$ random samples, which would be different from $C$. Note that any positive-definite and symmetric matrix has Cholesky decompositions such as $C = PP^T$ and $R = QQ^T$, where $P$ and $Q$ are lower triangular matrices.

In principle, the procedure in [23] re-pairs the samples in $X$ in order to obtain $\hat{X}$ that has the closest correlation matrix to $C$. To do so, it builds another $n \times m$ matrix $K$, where each column has a random permutation of $m$ van der Waerden scores (the inverse of the standard normal distribution [24]). In this way, $\hat{K} = K(Q^{-1})^T P^T$ has the closest correlation matrix

to $C$. Then, the desired matrix $\hat{X}$ can be obtained by simply re-pairing the samples in $X$ in the same order as the samples in $\hat{K}$. Therefore, $\hat{X}$ has the same correlation matrix as $\hat{K}$ that is a close approximation to $C$.

For example, we apply the correlation control technique to the four samples for two independent variables in the above example and the correlation matrix becomes quite close to an identity matrix

$$\underbrace{\begin{pmatrix} 1 & -0.73 \\ -0.73 & 1 \end{pmatrix}}_{\text{Before correlation control}} \Rightarrow \underbrace{\begin{pmatrix} 1 & -0.098 \\ -0.098 & 1 \end{pmatrix}}_{\text{After correlation control}}.$$

Therefore, the condition of correlation in Section III-B can also be satisfied.

Note that we assume independent variables in this paper for illustration purpose; however, the random process variables are spatially correlated in practice. Therefore, a correlation matrix extracted from the measurements is typically needed to generate the correlated samples.

In addition, it is worthwhile to point out that the correlation control technique [23] can only be applied to joint normal distributions, since only second-order statistics is needed as the input (i.e., the covariance matrix).

### D. Moments Via Sampling Methods

Next, we need to approximately evaluate the integral in (2) with a small number of "representative" LHS samples. In particular, the $k$th-order probabilistic moment $m_y^k$ can be estimated as

$$m_y^k = E(y^k) = \int y^k \cdot \text{pdf}(y)dy \approx \frac{1}{N} \cdot \sum_{j=1}^{N} f(\vec{x}_j)^k \qquad (9)$$

where $\vec{x}_j (j = 1, \ldots, N)$ are the $j$th samples of $\vec{x}$ using LHS method and correlation control technique. $f(\vec{x}_j)$ is the performance merit of the circuit with input $\vec{x}_j$. This approach is actually the sampled form of the expectation value $E(y^k)$ and only utilizes these representative samples $f(\vec{x}_j)$.

### E. Discussion of Proposed Methods

The proposed method has the following positive features.

1) The proposed method needs no response-surface model, therefore, it avoids the exponential complexity and bias-variance tradeoff in the existing RSM models.
2) LHS method is used to generate samples which is a variant of MC method and intrinsically capable of handling high-dimension problems efficiently.

In the meantime, the proposed method has a major drawback. These methods pick a few samples as "representatives" of the entire sampling space, which implicitly implies that the neighborhood around each sampling point $\vec{x}_j$ has the similar circuit behavior $f(\vec{x}_j)$. This is a linearized assumption; therefore, more samples could be needed to accurately describe the strongly nonlinear circuit behavior in strongly nonlinear problems.

## IV. PDF/CDF Calculation With Moments

### A. Normalized PDF for Error Prevention

The next step is to compute the residues $a_r$ and the poles $b_r$ in (6) with high-order moments so that the impulse response $h(t)$ in (7) can be evaluated to approximate pdf$(y)$. Note that probabilistic moments $m_y^k$ should be multiplied by a scaling factor to compute time moments in (4). As such, (6) results in a nonlinear equation system as

$$-\begin{bmatrix} \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots \frac{a_M}{b_M} \\ \frac{a_1}{b_1^2} + \frac{a_2}{b_2^2} + \cdots \frac{a_M}{b_M^2} \\ \frac{a_1}{b_1^3} + \frac{a_2}{b_2^3} + \cdots \frac{a_M}{b_M^3} \\ \vdots \\ \frac{a_1}{b_1^{2M}} + \frac{a_2}{b_2^{2M}} + \cdots \frac{a_M}{b_M^{2M}} \end{bmatrix} = \begin{bmatrix} \widehat{m}_y^0 \\ \widehat{m}_y^1 \\ \widehat{m}_y^2 \\ \vdots \\ \widehat{m}_y^M \end{bmatrix}. \qquad (10)$$

This nonlinear system can be efficiently solved with the numerical solution presented in [12]. However, the calculated residues $a_r$ and the poles $b_r$ may suffer from numerical noises such as roundoff error. Therefore, we propose to normalize the PDF calculated from (7) to cancel out the roundoff error.

Let us denote $\widehat{m}_y^k$ as the exact value of the $k$th-order time moment in (4), and $\tilde{m}_y^k$ as the estimated value of the $k$th-order time moment. Also, we assume $\tilde{m}_y^k = \text{const} \cdot \widehat{m}_y^k$ due to roundoff error, where const is a scaling constant. As such, the right-hand side of (10) should be substituted by $\tilde{m}_y^k$, which leads to $\tilde{a}_j = \text{const} \cdot a_j$ and $a_j$ are exact values of the residues.

In order to eliminate the scaling constant in $\tilde{a}_j$, we propose to normalize pdf$(y)$ as follows. First, $y$ can be discretized into several discrete points $y_p$, $(p = 1, \ldots, K)$. Then, the PDF value on $p$th discrete point can be divided with the sum of PDF values for all discrete points as

$$\text{pdf}_{\text{norm}}(y_p) = \frac{\sum_{r=1}^{M} \text{const} \cdot a_r \cdot e^{\widehat{b}_r^{k+1} \cdot y_p}}{\sum_{p=1}^{K} \sum_{r=1}^{M} \text{const} \cdot a_r \cdot e^{\widehat{b}_r^{k+1} \cdot y_p}}. \qquad (11)$$

In this way, the scaling constant can be canceled out and thus the normalization procedure improves the numerical stability of the proposed algorithm.

### B. Error Estimation

It is significant to estimate the approximation error of pdf$(y)$ using AWE method [12] but exact PDF is usually not available. Instead, we consider the approximation with first $q+1$ order moments as the exact value and use the relative error of $\Phi(\omega)$ [i.e., the Fourier transform of pdf$(y)$ in (3)] to measure the accuracy of PDF approximation using first $q$-order moments

$$\text{Error} = \left| \frac{\Phi^{q+1}(\omega) - \Phi^q(\omega)}{\Phi^{q+1}(\omega)} \right| \qquad (12)$$

$$= \left| \frac{(-j\omega)^{q+1}}{(q+1)!} \cdot \left( \sum_{p=0}^{q+1} \frac{(-j\omega)^p}{p!} \cdot \frac{m_y^p}{m_y^{q+1}} \right)^{-1} \right|.$$

**Algorithm 2** Overall proposed algorithm

**Input:** random variables $\vec{x} = (x_1, \cdots, x_M)$ with known probabilistic distributions and correlation matrix $C$.
**Output:** the estimated PDF/CDF of circuit behavior $y$.

**/* Step 1: High-Order Moments Calculation */**
Use LHS method to generate $N$ samples $\vec{x}_j$ $(j = 1, \cdots, N)$.
Re-pair these samples with correlation control technique to achieve the correlation matrix $C$.
Run SPICE simulations on these samples for corresponding circuit behavior $y_j$ $(j = 1, \cdots, N)$.
Compute moments $m_y^k$ with $y_j$ as (9).

**/* Step 2: Moment Matching */**
Calculate the time moments $\widehat{m}_y^k$ with $m_y^k$ as (4).
Make $\widehat{m}_y^k$ equal to $\widehat{m}_t^k$ by matching first several moments.
Solve the resulting nonlinear equation system in (10) for residues $a_r$ and poles $b_r$.

**/* Step 3: PDF Calculation */**
Compute normalized PDF/CDF of $y$ with $a_r$ and $b_r$.

When $|m_y^p| \geq |m_y^{q+1}|$ $(p \leq q+1)$, above estimation error has upper bound

$$\text{Error} \leq \left| \frac{(-j\omega)^{q+1}}{(q+1)!} \cdot \left( \sum_{p=0}^{q+1} \frac{(-j\omega)^p}{p!} \right)^{-1} \right|. \tag{13}$$

The above error analysis results in an important observation as follows.

*Property 3:* Assume high-order moments $m_y^k (k = 1, \ldots, N)$ are used to predict pdf$(y)$. When the moments $m_y^k$ decay as the moment order increases, the approximation of pdf$(y)$ has upper error bound.

In this paper, we consider a much stronger condition $y \in [0, 1]$ with the help of linear transformations of $y$ (e.g., scaling, shifting, and so on). As such, the approximation of PDF/CDF can provide high accuracy with high-order moments.

## V. Overall Algorithm

### A. Algorithm Flow

The overall algorithm flow for PDF/CDF approximation has been summarized in Algorithm 2.

### B. Implementation Details

We briefly discuss several implementation issues as follows.

1) *Linear transformation*: To ensure that an upper error bound exists, we propose to transform $y$ into the interval [0, 1] so that moments can decay as the moment order increases. In general, the transformations include scaling, shifting, flipping, and other linear operations. Note that the extracted PDF/CDF should be converted back to be the real results.

TABLE I
Process Parameters of MOSFETs

| Variable Name | $\sigma/\mu$ | Unit |
|---|---|---|
| Flat-band voltage ($V_{\text{fb}}$) | 0.1 | V |
| Gate oxide thickness ($t_{\text{ox}}$) | 0.05 | m |
| Mobility ($\mu_0$) | 0.1 | $\text{m}^2/Vs$ |
| Doping concentration at depletion ($N_{\text{dep}}$) | 0.1 | $\text{cm}^{-3}$ |
| Channel-length offset ($\Delta L$) | 0.05 | m |
| Channel-width offset ($\Delta W$) | 0.05 | m |
| Source/drain sheet resistance ($R_{\text{sh}}$) | 0.1 | $\text{Ohm/mm}^2$ |
| Source-gate overlap unit capacitance ($C_{\text{gso}}$) | 0.1 | F/m |
| Drain-gate overlap unit capacitance ($C_{\text{gdo}}$) | 0.1 | F/m |

2) *Numerical instability*: In principle, more high-order moments can improve the approximation accuracy of PDF/CDF. Unfortunately, it is not true because the moments can decay dramatically and be close to zero when the moment order increases. Therefore, the inevitable numerical noise (e.g., ill-conditioned moment matrix) prevents further accuracy improvement. This is an intrinsic drawback of the moment-matching method [12].
3) *PDF/CDF shifting*: The approximated PDF/CDF of $y$ can be far from the real location and display a large delay in the time domain [13]. Therefore, the PDF/CDF shifting technique using the modified Chebyshev inequality in [13] can be used to ensure the PDF/CDF approximations match the MC results.

## VI. Experimental Results

The proposed algorithm has been implemented in MATLAB environment with HSPICE and BSIM4 transistor model. We use a two-stage operational amplifier (OPAMP) and a static random-access memory (SRAM) bit-cell to demonstrate the accuracy and efficiency of proposed algorithm on both ac and transient performance merits by comparing against MC method. In order to introduce process variations to these circuits, we consider nine process parameters for each transistor shown in Table I, which are physically independent parameters [25] and modeled as independent Gaussian random variables.

For comparison purpose, all experiments involve three different approaches as follows.

1) *MC method*: Calculate the probabilistic distributions (PDF and CDF) from a huge number of MC samples. This is the "gold standard" for the comparison in this section.
2) *MC + Moment-matching method*: The probabilistic distributions are computed with moment-matching method [12], [13], where the moments are evaluated using MC samples.
3) *Proposed method*: The moments are calculated with (9) using a few samples from the LHS method coupled with correlation control technique. Also, the probabilistic distributions of circuit performance are approximated using the moment-matching method in [12] and [13].

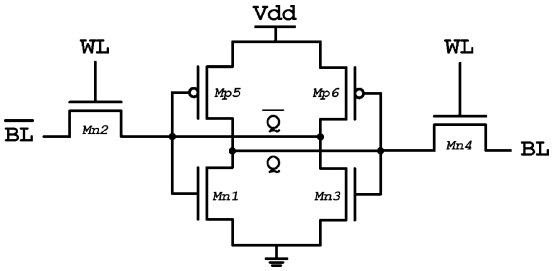



Fig. 3. Schematic of a 6-T SRAM bit cell.

### A. 6-T SRAM Bit Cell

Let us first study a typical design of six-transistor SRAM bit cell as shown in Fig. 3 [26], which stores one memory bit and consists of six transistors: the four transistors $Mn1$, $Mn3$, $Mp5$, and $Mp6$ forms two inverters to keep either a logic "0" or "1." Two additional access transistors $Mn2$ and $Mn4$ control the access to the bit cell during read and write operations. The word line (WL) is used to determine whether the bit cell should be accessed [connected to bit lines (BLs)] and the BL and $\bar{\text{BL}}$ are used to read/write the actual data from/to the cell.

To model the process variations, we introduce random variations to nine process parameters shown in Table I of each transistor, which implies totally 54 independent random variables in this example.

As an illustration, we investigate the discharge behavior on $\bar{\text{BL}}$ during the reading operation when $Q$ node stores 1. In details, both $\bar{\text{BL}}$ and BL are first precharged to $Vdd$ and then $\bar{\text{BL}}$ starts to discharge when the WL becomes high. When the voltage difference $\Delta v$ between $\bar{\text{BL}}$ and BL becomes large, $\Delta v$ can be sensed by the sense amplifier connected to both $\bar{\text{BL}}$ and BL.

However, the process variations can significantly change the discharge behavior and, particularly, a reading failure can happen when $\Delta v$ is too small to be sensed by the sense amplifier at the end of reading operation. Therefore, it is of great interest to study the probabilistic distribution of node voltage $\bar{\text{BL}}$ at the end time step of reading operation considering the process variations. Note that the reading failure of SRAM bit cell is a "rare event" with extremely small probability [27] that is *not* in the scope of this paper, while the overall stochastic discharge behavior in this SRAM cell will be studied.

1) *Comparison of PDF/CDF Approximation:* We have applied all different methods (i.e., MC, MC + moment matching, proposed method) to this SRAM bit-cell example and plotted their approximations of PDFs and CDFs with first 20 moments in Figs. 4(a) and 4(b), respectively. Note that we use kernel density estimation method [28] to estimate the PDF using 1E+5 MC samples and then analytically integrate the PDF to get CDF.

For comparison purpose, we plotted PDF and CDF from proposed method using first 10 moments in the same figures, which have significant accuracy loss and only provide rough approximations. When the approximation order increases to 20, the PDF and CDF estimations are clearly improved and closely match with MC results except for the tail regions.

In addition, the PDF from the proposed method (the curve with circle marks) contains numerical oscillations in the tails.

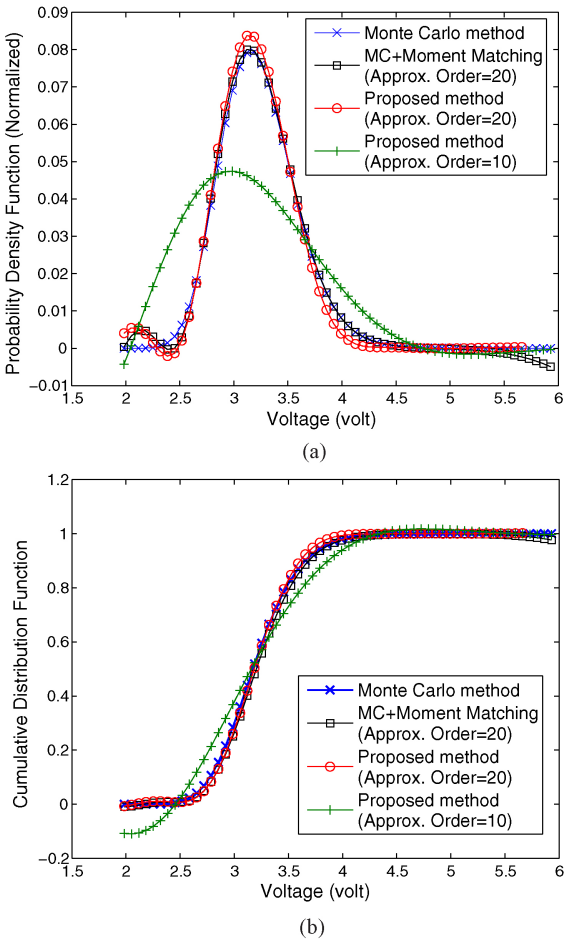



Fig. 4. PDF/CDF from proposed method for SRAM bit-cell example. (a) PDF approximation. (b) CDF approximation.

In fact, the similar oscillations are demonstrated in the PDF from the MC+moment matching method (the curve with square marks) where the "exact" moment values from MC samples are used. In principle, the approximated PDF should eliminate these oscillations and asymptotically approach the exact PDF when the approximation order increases. However, the moment matrix used to calculate the residues $a_r$ and poles $b_r$ in (6) becomes severely ill conditioned and leads to inaccurate and unstable results of $a_r$ and $b_r$. The numerical instability issue of moment-matching method [12] prevents further accuracy improvement with more probabilistic moments and remains a challenging issue.

Moreover, the approximation accuracy can be quantitatively characterized by the average error over several specific points of the CDF. The proposed method achieves 5.03% relative error on average when compared with the CDF from MC samples.

2) *Comparison of Efficiency:* We study the efficiency of different methods in Table II, where the CDF from MC method with 1E+5 samples serves as the exact CDF. In addition, the efficiency is measured by the number of required samples which equals to the number of SPICE simulations, since the transistor-level simulations are the most time-consuming calculations.

TABLE II
EFFICIENCY COMPARISON OF CDF APPROXIMATIONS

| | MC Method[a] | MC Method | Proposed Method |
|---|---|---|---|
| Accuracy[b] | 100% | 95% | 95% |
| #Samples | 1E+5 | 8.6E+3 (159X) | 54 (1X) |

[a]CDF from MC with 1E+5 samples serves as the exact CDF.
[b]Accuracy of CDF approximation is measured by the average accuracy over several specific points of the CDF.

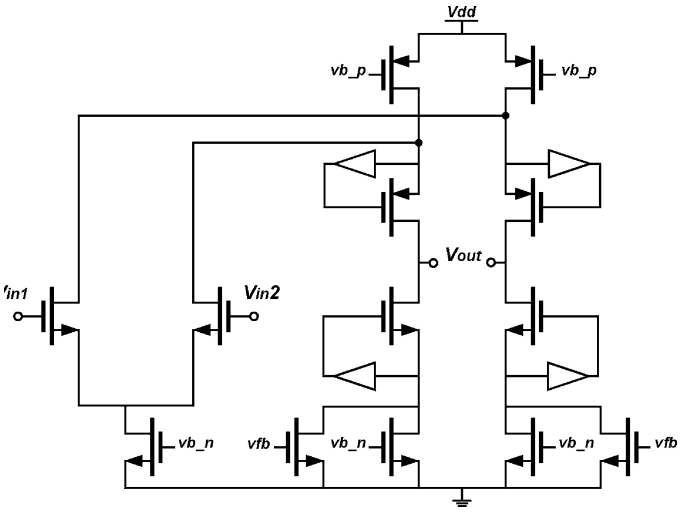



Fig. 5. Simplified Schematic of OPAMP.

In order to provide a fair comparison, we incrementally add MC samples to provide the same accuracy as the proposed method. In fact, the MC method with 8.6E+3 samples offers 95% accuracy on average; therefore, the proposed method is 159X faster than MC method for the same accuracy.

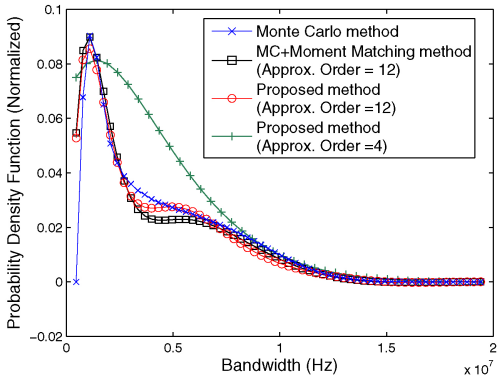
## B. Operational Amplifier

We have validated the proposed method on SRAM bit-cell example that involves transient performance merits (e.g., voltage discharge in time domain) and Gaussian-like probabilistic distribution. Next, an OPAMP, shown in Fig. 5, is used to study the proposed method on ac performance of merits (e.g., bandwidth) and non-Gaussian behavioral distributions.

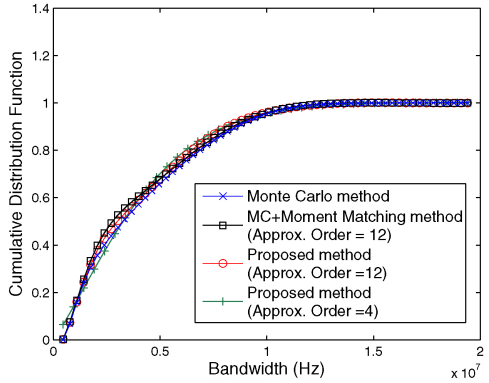
In the OPAMP, $vb\_n$, $vb\_p$ are bias voltages for nMOS and pMOS devices, respectively. *vfb* is the feedback voltage set separately by common mode feedback block. The small triangular devices denote the gain boosting components. We introduce the variations to process parameters in the Table I for all transistors except transistors associated with $vb\_n$, $vb\_p$, and *vfb*. As such, there exist totally 90 random variables to model the local random variations, which is a typical large-scale problem in practice.

The circuit behavior of the OPAMP is described by its *bandwidth*, therefore, we aim to extract the "arbitrary" unknown distributions (PDF and CDF) of bandwidth under process variations in this case. We applied all different methods on this example and compared their performance in the following.

1) *Comparison of PDF/CDF Approximation:* We plot the approximated PDF and CDF from all different methods in Figs. 6(a) and 6(b), respectively. The PDF from MC method




(a)




(b)

Fig. 6. Extracted PDF/CDF from proposed method for OPAMP example. (a) PDF approximation. (b) CDF approximation.

TABLE III
EFFICIENCY COMPARISON OF CDF APPROXIMATIONS

| | MC Method[a] | MC Method | Proposed Method |
|---|---|---|---|
| Accuracy[b] | 100% | 98% | 98% |
| #Samples | 5E+5 | 1.5E+5 (1666X) | 90 (1X) |

[a]CDF from MC with 5E+5 samples is treated as the exact CDF.
[b]Accuracy of CDF approximation is measured by the average accuracy over several specific points of the CDF.

is estimated using kernel density estimation method [28] with 5E+5 samples and is the "exact" PDF of bandwidth. For comparison purpose, the PDF from proposed method with first four moments is plotted in the same figure, which clearly deviates from the exact PDF from MC method (the curve with cross marks) by a large amount.

We further increase the approximation order to 12 and the extracted PDF (the curve with circle marks) becomes much closer to the exact PDF from MC. However, there are some numerical oscillations in the PDFs from moment matching based methods (i.e., proposed method and the MC+moment matching method), which result from the numerical noise during moment matching.

The approximation accuracy is measured by the average error over several specific points of the CDF. In this OPAMP case, the proposed method achieves 1.65% relative error on average when compared with the exact CDF from MC method.

2) *Comparison of Efficiency:* We use the number of required samples to measure the efficiency between different methods in Table III. The CDF from MC method with 5E+5

samples is treated as the exact CDF. The proposed method uses 90 samples to provide 98% accuracy while MC method needs 1.5E+5 samples for the same accuracy. It implies the proposed method offers nearly linear complexity and is 1666X faster than MC method.

## VII. Conclusion

In this paper, we proposed an efficient moment-matching-based algorithm to extract the probabilistic distributions of stochastic circuit behavior. Our approach performed an efficient evaluation of high-order moments of circuit behavior and thus circumvented the use of RSMs. Moreover, the proposed method has been successfully extended to deal with high-dimension problems with nearly linear complexity. Experiments showed that the proposed method can provide up to 1666X runtime speedup with the same accuracy when compared to the MC method.

The work presented in this paper has three limitations. First, the deployed circuit examples are relatively small when compared with industrial designs, where 1000+ random variables can be easily involved. Second, most process variation sources are spatially correlated in practice but we only assume the independent random variables for illustration purpose. Third, the conventional moment-matching method suffers from numerical noise and instability issues that remain significant challenges. Our future study will investigate the correlated process variation sources, study large-scale industrial problems, and deal with the numerical noise issues.

## References


[1] S. Nassif, "Modeling and analysis of manufacturing variations," in *Proc. IEEE Custom Integr. Circuits Conf.*, May 2001, pp. 223–228.

[2] X. Li, Y. Zhan, and L. Pileggi, "Quadratic statistical MAX approximation for parametric yield estimation of analog/RF integrated circuits," *IEEE Trans. Comput.-Aided Design*, vol. 27, no. 5, pp. 831–843, May 2008.

[3] K. Bowman, S. Duvall, and J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, Feb. 2002.

[4] R. Chang, Y. Cao, and C. Spanos, "Modeling the electrical effects of metal dishing due to CMP for on-chip interconnect optimization," *IEEE Trans. Semicond. Manuf.*, vol. 51, no. 10, pp. 1577–1583, Oct. 2004.

[5] M. Eisele, J. Berthold, D. Schmitt-Landsiedel, and R. Mahnkopf, "The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 5, no. 12, pp. 360–368, Dec. 1997.

[6] H. Yu, X. Liu, H. Wang, and S. Tan, "A fast analog mismatch analysis by an incremental and stochastic trajectory piecewise linear macromodel," in *Proc. IEEE/ACM ASPDAC*, Jan. 2010, pp. 211–216.

[7] F. Gong, H. Yu, and L. He, "PiCAP: A parallel and incremental full-chip capacitance extraction considering random process variation," in *Proc. ACM/IEEE DAC*, Jul. 2009, pp. 764–769.

[8] F. Gong, H. Yu, Y. Shi, D. Kim, J. Ren, and L. He, "Quickyield: An efficient global-search based parametric yield estimation with performance constraints," in *Proc. ACM/IEEE DAC*, June 2010, pp. 392–397.

[9] S. Vrudhula, J. M. Wang, and P. Ghanta, "Hermite polynomial based interconnect analysis in the presence of process variations," *IEEE Trans. Comput.-Aided Design*, vol. 25, no. 10, pp. 2001–2011, Oct. 2006.

[10] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. New York: Springer-Verlag, 1996.

[11] D. Xiu and G. E. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic differential equations," *SIAM J. Sci. Comput.*, vol. 24, no. 2, pp. 619–644, Feb. 2002.

[12] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. Comput.-Aided Design*, vol. 9, no. 4, pp. 352–366, Apr. 1990.

[13] X. Li, J. Le, P. Gopalakrishnan, and L. T. Pileggi, "Asymptotic probability extraction for non-normal distributions of circuit performance," in *Proc. IEEE/ACM ICCAD*, Nov. 2004, pp. 2–9.

[14] X. Li and H. Liu, "Statistical regression for efficient high-dimensional modeling of analog and mixed-signal performance variations," in *Proc. DAC*, Jun. 2008, pp. 38–43.

[15] X. Li, "Finding deterministic solution from underdetermined equation: Large-scale performance variability modeling of analog/RF circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 29, no. 11, pp. 1661–1668, Nov. 2010.

[16] T. R. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York: Springer-Verlag, 2008.

[17] A. Papoulis and S. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2001.

[18] M. H. DeGroot and M. J. Schervish, *Probability and Statistics*. Reading, MA: Addison-Wesley, 2011.

[19] F. Gong, H. Yu, and L. He, "Stochastic analog circuit behavior modeling by point estimation method," in *Proc. Int. Symp. Phys. Design*, Mar. 2011, pp. 175–182.

[20] E. Rosenblueth, "Point estimation for probability moments," *Proc. Nat. Acad. Sci. USA*, vol. 72, no. 10, pp. 3812–3814, 1975.

[21] Y.-G. Zhao and T. Ono, "New point estimation for probability moments," *J. Eng. Mech.*, vol. 126, no. 4, pp. 433–436, 2000.

[22] M. Stein, "Large sample properties of simulations using latin hypercube sampling," *Technometrics*, vol. 29, pp. 143–151, May 1987.

[23] R. Iman and W. Conover, "A distribution-free approach to inducing rank correlation among input variables," *Commun Stat: Simul. Comput.*, vol. B11, no. 3, pp. 311–334, 1982.

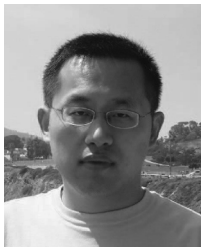
[24] W. J. Conover, *Practical Nonparametric Statistics*. New York: Wiley, 1980.

[25] P. Drennan and C. McAndrew, "Understanding MOSFET mismatch for analog design," *IEEE J. Solid State Circuits*, vol. 38, no. 3, pp. 450–456, Mar. 2003.

[26] J. Wang, S. Yaldiz, X. Li, and L. T. Pileggi, "SRAM parametric failure analysis," in *Proc. ACM/IEEE DAC*, July 2009, pp. 496–501.

[27] M. Denny, "Introduction to importance sampling in rare-event simulations," *Eur. J. Phys.*, vol. 22, no. 4, pp. 403–411, 2001.

[28] A. W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. New York: Oxford Univ. Press, 1997.

**Fang Gong** (S'08) received the B.S. degree from the Department of Computer Science, Beijing University of Aeronautics and Astronautics, Beijing, China, in 2005, and the M.S. degree from the Department of Computer Science, Tsinghua University, Beijing, in 2008. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, University of California, Los Angeles.

His current research interests include numerical computing and stochastic techniques for computer-aided design, including fast circuit simulation, yield estimation, and optimization.

**Sina Basir-Kazeruni** (S'09) received the B.A.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2010, and the M.S. degree in electrical engineering from the University of California, Los Angeles, where he is currently pursuing the Ph.D. degree.

During his undergraduate studies, he interned at various companies including NVIDIA, Santa Clara, CA, Synopsys, Mountain View, CA, and Honeywell Aerospace, Phoenix, AZ. His current research interests include design and optimization of digital integrated circuits and power/area-efficient very large-scale integrated systems.

**Lei He** (M'99–SM'08) received the Ph.D. degree in computer science from the University of California, Los Angeles (UCLA), in 1999.

He is currently a Professor with the Department of Electrical Engineering, UCLA. From 1999 to 2002, he was a Faculty Member with the University of Wisconsin, Madison. He is the author or co-author of one book and over 200 technical papers. His current research interests include modeling and simulation, very large-scale integrated circuits and systems, and cyber-physical systems.

Dr. He was nominated 12 times for the Best Paper Award and five times for the Best Paper or Best Contribution Award in numerous conferences, including the Design Automation Conference, the International Conference on Computer-Aided Design, and the *ACM Transactions on Electronic System Design Automation*.

**Hao Yu** (S'02–M'06) received the B.S. degree from Fudan University, Shanghai, China, and the M.S. and Ph.D. degrees from the Department of Electrical Engineering, University of California, Los Angeles.

He was a Senior Research Staff Member with Berkeley Design Automation, Santa Clara, CA. Since October 2009, he has been an Assistant Professor with the Circuits and Systems Division, School of Electrical and Electronic Engineering School, Nanyang Technological University, Singapore. He is the author or co-author of 43 refereed international publications and five book chapters. He holds four pending patent applications.

Dr. Yu was a recipient of one Best Paper Award from the *ACM Transactions on Design Automation of Electronic Systems*, two Best Paper Award nominations in the Design Automation Conference and the International Conference of Computer-Aided Design, and one Inventor Award from the Semiconductor Research Corporation.