Quadratic Statistical MAX Approximation for Parametric Yield Estimation of Analog/RF Integrated Circuits

Xin Li, *Member, IEEE*, Yaping Zhan, *Member, IEEE*, and Lawrence T. Pileggi, *Fellow, IEEE*

***Abstract*—In this paper, we propose an efficient numerical algorithm for estimating the parametric yield of analog/RF circuits, considering large-scale process variations. Unlike many traditional approaches that assume normal performance distributions, the proposed approach is particularly developed to handle multiple correlated nonnormal performance distributions, thereby providing better accuracy than the traditional techniques. Starting from a set of quadratic performance models, the proposed parametric yield estimation conceptually maps multiple correlated performance constraints to a single auxiliary constraint by using a MAX operator. As such, the parametric yield is uniquely determined by the probability distribution of the auxiliary constraint and, therefore, can easily be computed. In addition, two novel numerical algorithms are derived from moment matching and statistical Taylor expansion, respectively, to facilitate efficient quadratic statistical MAX approximation. We prove that these two algorithms are mathematically equivalent if the performance distributions are normal. Our numerical examples demonstrate that the proposed algorithm provides an error reduction of 6.5 times compared to a normal-distribution-based method while achieving a runtime speedup of 10–20 times over the Monte Carlo analysis with 10^3 samples.**

***Index Terms*—Analog/RF circuits, MAX operator, parametric yield.**

I. Introduction

THE CONTINUOUS scaling of IC feature size has resulted in relatively large process variations [2]. For analog/RF circuits designed in sub-90-nm-technology nodes, parametric yield loss due to random variations becomes a significant or even dominant portion of the total yield loss. Therefore, accurately predicting parametric yield, considering large-scale process variations, becomes an important task for today's robust analog/RF design.

During the last decade, various techniques have been proposed for statistical analysis of analog/RF circuits [3]–[14]. Most of these works [3]–[8] apply response surface modeling to approximate the performances of interest (e.g., gain and bandwidth) as polynomial (e.g., linear or quadratic) functions of process parameters (e.g., V_{TH} and T_{OX}). These response surface models are then utilized to estimate the parametric yield of a given circuit design.

Many existing response surface modeling techniques (e.g., [4] and [5]) rely on linear approximation, which is efficient and accurate when process variations are sufficiently small. Given the increasingly larger variations in nanoscale technologies, however, such a linear approximation can yield inaccurate results, particularly because a great number of analog/RF performances are strongly nonlinear in the presence of large-scale process variations. As will be demonstrated by the numerical examples in Section IV, a 9% absolute error is observed for parametric yield estimation by using linear performance models for several analog/RF circuits designed in commercial manufacturing processes.

To achieve better accuracy, quadratic response surface modeling [6]–[8] can be used, but it makes parametric yield estimation much more difficult. Although linear response surface modeling can easily map a set of performance constraints to a well-defined polyhedron [9]–[13] (called the *feasible space*) in the process parameter space, such a mapping becomes nonlinear for quadratic modeling. In general, when quadratic response surface modeling is applied, the feasible space can be nonconvex or even discontinuous. Therefore, the parametric yield, which is equal to the integral of the probability density function over the feasible space, becomes much more difficult to compute [14].

The Monte Carlo analysis is applied in [6] and [7] to estimate the parametric yield defined by a set of quadratic performance constraints. Traditionally, the quadratic response surface modeling cost dominates the overall computational cost, since it requires running expensive transistor-level simulations to generate a number of sampling points. In such cases, the Monte Carlo analysis cost is negligible. However, recent advances in statistical circuit optimization make it possible to extract the response surface models only once in a local design space and then use them to estimate the parametric

Manuscript received July 3, 2007; revised October 9, 2007. This work was supported in part by the Center for Circuit and System Solutions, (C2S2), Focus Center Research Program, under Contract 2003-CT-888, by the Semiconductor Research Cooperation under Contract 2005-TJ-1298, and by Mentor Graphics Inc. This paper was presented in part at the 45th IEEE/ACM Design Automation Conference (DAC 2007), San Diego, CA. This paper was recommended by Associate Editor H. E. Graeb.

X. Li and L. T. Pileggi are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: xinli@ece.cmu.edu; pileggi@ece.cmu.edu).

Y. Zhan was with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA. He is now with Advanced Micro Devices, Austin, TX 78741 USA (e-mail: yaping.zhan@amd.com).

Digital Object Identifier 10.1109/TCAD.2008.917582

0278-0070/$25.00 © 2008 IEEE

yield values at different design points [5], [8]. In this case, the yield estimation takes a large portion of the total computational cost, since it is repeatedly performed inside the optimization loop.

Li *et al.* [15], [16] propose an asymptotic probability extraction (APEX) algorithm to estimate the performance distribution of a given quadratic response surface model. The APEX approach, however, can only be applied to a single performance metric, whereas the parametric yield of most analog/RF circuits is defined by multiple performance constraints. The challenging problem here is how one can simultaneously consider *multiple* and *correlated* performance metrics whose probability distributions are nonnormal.

In this paper, we propose a novel algorithm for efficient parametric yield estimation of multiple correlated nonnormal performance distributions. Our proposed algorithm conceptually maps multiple performance constraints to a single *auxiliary constraint* by using a MAX operator. The auxiliary constraint is analytically approximated as a quadratic function of process parameters. As such, the parametric yield is uniquely determined by a single quadratic constraint (i.e., the auxiliary constraint) and, therefore, can easily be estimated using the APEX algorithm [15], [16], which is based on moment matching.

Another important contribution of this paper is to propose two novel numerical algorithms for efficient statistical MAX approximation. Unlike many statistical timing analysis techniques that linearly approximate the MAX operator [18]–[22], we approximate MAX as a quadratic function to capture the nonlinearities that are observed for most analog/RF performance variations. The proposed MAX approximation algorithms are derived from *moment matching* and *statistical Taylor expansion*, respectively. We prove that the moment matching algorithm yields an optimal approximation, for which the weighted squared error is minimized.

Although both moment matching and statistical Taylor expansion have widely been applied to linear MAX approximation for statistical timing analysis [18]–[22], the theoretical connection between these two approaches is not well studied and remains unclear. In this paper, we prove that the statistical Taylor expansion is exactly equivalent to the moment matching method if the performance distributions are normal. This observation provides a theoretical foundation for mathematically connecting these two algorithms. For many practical circuit examples, both moment matching and statistical Taylor expansion are equally efficient in approximation accuracy and computational cost. As will be demonstrated by the numerical examples in Section IV, our proposed quadratic MAX approximation provides an error reduction of 6.5 times compared to the traditional linear approximation while achieving a runtime speedup of 10–20 times over the Monte Carlo analysis with 10^3 samples.

The remainder of this paper is organized as follows. In Section II, we review the background materials, and then, in Section III, we propose our parametric yield estimation algorithm. The efficacy of the proposed algorithm is demonstrated by numerical examples in Section IV. Finally, we conclude in Section V.

II. Background

A. Process Variation Modeling

Process variations are the deviations from the intended or designed values for the structural or electrical parameters of concern. According to the geometrical scale of their occurrence, process variations can be classified into two broad categories: 1) *interdie variations* and 2) *intradie variations*. Interdie variations model the common/average variations across the die, whereas intradie variations model the individual but spatially correlated local variations within the same die.

In most practical applications, both interdie and intradie variations are modeled as the random variables that are jointly normal. In such cases, *principal component analysis* (PCA) can be applied to find a set of independent factors to represent the original correlated random variables [26].

Given N process parameters $\eta = [\eta_1, \eta_2, \ldots, \eta_N]^T$, the process variations $\Delta\eta = \eta - \eta_0$, where η_0 contains the mean values of η, are modeled as zero-mean normal distributions. The correlation of $\Delta\eta$ can be represented by a symmetric positive semidefinite covariance matrix R [26]. PCA decomposes R as follows:

$$R = V \cdot \Theta \cdot V^T \tag{1}$$

where $\Theta = \mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_N)$ contains the eigenvalues of R, and $V = [V_1 V_2, \ldots, V_N]$ contains the corresponding eigenvectors that are orthonormal, i.e., $V^T V = I$ (I is an identity matrix). Based on Θ and V, PCA defines a set of new random variables, i.e.,

$$\varepsilon = \Theta^{-0.5} \cdot V^T \cdot \Delta\eta. \tag{2}$$

These new random variables $\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]^T$ are called the principal components or factors. It is easy to verify that $\{\varepsilon_i; i = 1, 2, \ldots, N\}$ are independent and standard normal (i.e., have zero mean and unit variance) [26].

The essence of PCA can be interpreted as a coordinate rotation of the space defined by the original random variables. In addition, if the magnitude of the eigenvalues θ_i dramatically deceases, it is possible to use a small number of random variables, i.e., a small subset of principal components, to approximate the original N-dimensional space. More details on PCA can be found in [26].

B. Response Surface Modeling

Given a circuit topology, the circuit performance (e.g., gain or bandwidth) is a function of both design parameters (e.g., bias current and transistor sizes) and process parameters (e.g., V_{TH} and T_{OX}). The design parameters are determined during the design phase; however, the process parameters must be modeled as random variables to account for uncertain manufacturing fluctuations. Given a set of fixed design parameters, the circuit performance f can be approximated as a linear response surface model [4], [5], [28], i.e.,

$$f(\varepsilon) = B^T \cdot \varepsilon + C \tag{3}$$

where $\varepsilon = [\varepsilon_1 \varepsilon_2, \ldots, \varepsilon_N]^T$ denotes the independent random variables defined in (2), and $B \in R^N$ and $C \in R$ are the model coefficients.

The linear response surface model in (3) is accurate if the process variations are sufficiently small. However, the large-scale variations in nanoscale technologies suggest that quadratic response surface models are required to improve the modeling accuracy [6]–[8], [28], i.e.,

$$f(\varepsilon) = \varepsilon^T \cdot A \cdot \varepsilon + B^T \cdot \varepsilon + C \tag{4}$$

where $A \in R^{N \times N}$, $B \in R^N$, and $C \in R$ are the model coefficients. The unknown model coefficients in (3) and (4) can be determined by solving the following overdetermined linear equations at a number of sampling points [28]:

$$B^T \varepsilon_{(m)} + C = \tilde{f}_{(m)} \quad (m = 1, 2, \ldots, M) \tag{5}$$

$$\varepsilon_{(m)}^T A \varepsilon_{(m)} + B^T \varepsilon_{(m)} + C = \tilde{f}_{(m)} \quad (m = 1, 2, \ldots, M) \tag{6}$$

where $\varepsilon_{(m)}$ and $\tilde{f}_{(m)}$ denote the values of ε and f at the mth sampling point, respectively, and M is the total number of sampling points.

It should be noted that the aforementioned response surface models are substantially different from the macromodels proposed in [24]. Since macromodels are utilized for circuit synthesis [24], they must cover an extremely large analog design space (e.g., 100 times of changes in design variables) and, therefore, are strongly nonlinear. In contrast, our response surface models are extracted to capture process variations where most process parameters only vary by $\pm 20 \sim 40\%$. Hence, the response surface models in this paper will likely be weakly nonlinear. In practice, the response surface modeling accuracy depends on multiple factors, e.g., the circuit topology, the performance of interest, and the probability distribution of process variations. We notice that quadratic response surface models are sufficiently accurate for many practical circuits designed in today's commercial manufacturing processes. However, as IC technologies are scaled to finer feature sizes, process variations will increasingly become larger. The quadratic model assumption may become invalid, and a strongly nonlinear model may be required in the future.

C. Parametric Yield Estimation

If all analog/RF performance metrics are approximated as response surface models, the performance constraints can be expressed as the following standard form:

$$f_k(\varepsilon) \leq 0 \quad (k = 1, 2, \cdots, K) \tag{7}$$

where $f_k(\varepsilon)$ represents the response surface model of the kth performance metric, and K is the total number of performance constraints. The standard form in (7) is capable of handling several extensions. For example, $f_k(\varepsilon) \leq C$ and $f_k(\varepsilon) \geq C$ can be expressed as $f_k(\varepsilon) - C \leq 0$ and $-f_k(\varepsilon) + C \leq 0$, respectively.

Given the performance constraints in (7), the *parametric yield* is equal to the probability that all performance constraints

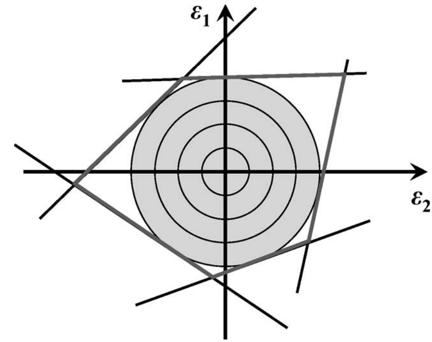


Fig. 1. Approximate the feasible space by a maximal inscribed ellipsoid for parametric yield estimation. In this example, the ellipsoid becomes a ball, since the random variables ε_1 and ε_2 defined in (2) are independent and standard normal after principal component analysis.

are satisfied, i.e.,

$$\text{Yield} = P(f_1 \leq 0 \quad \& \quad f_2 \leq 0 \quad \& \cdots \& \quad f_K \leq 0) \tag{8}$$

where $P(\bullet)$ denotes the probability. The probability in (8) depends on all performance distributions as well as their *correlations*.

If the response surface models in (7) are linear, then

$$f_k(\varepsilon) = B_k^T \varepsilon + C_k \quad (k = 1, 2, \ldots, K) \tag{9}$$

where $B_k \in R^N$ and $C_k \in R$ are the linear model coefficients of the kth performance function, and the *feasible space*, i.e.,

$$F = \left\{ \varepsilon | B_k^T \varepsilon + C_k \leq 0 \quad (k = 1, 2, \ldots, K) \right\} \tag{10}$$

is a *polyhedron*. The parametric yield in (8) is equal to the integral of the probability density function $PDF_\varepsilon(\varepsilon)$ over the feasible space, i.e.,

$$\text{Yield} = \int_F PDF_\varepsilon(\varepsilon) \cdot d\varepsilon \tag{11}$$

where $PDF_\varepsilon(\bullet)$ is the joint probability density function of the N-dimensional random variable ε.

To estimate the parametric yield, one traditional approach is to approximate the feasible space in (10) as a maximal inscribed ellipsoid, and then, the integral in (11) can easily be evaluated, as shown in Fig. 1. Such an ellipsoid approximation has widely been used in both analog and digital applications [9]–[13].

The aforementioned ellipsoid approximation, however, suffers from several major limitations. First, as shown in Fig. 1, the maximal inscribed ellipsoid does not cover the entire feasible space. Therefore, the estimated yield value is a lower bound of the actual parametric yield. In practice, the estimation error can be large, particularly if the feasible space is high dimensional. Second, the ellipsoid approximation cannot easily handle quadratic response surface models for which the feasible space can be nonconvex or even discontinuous. These observations, therefore, motivate us to develop a more accurate parametric yield estimation algorithm in this paper.

D. Linear Statistical MAX Approximation

Our proposed parametric yield estimation is based on the statistical approximation of the MAX operator. Linear statistical MAX approximation is a traditional technique that has widely been studied for statistical timing analysis [17]–[22]. Given two random variables x and y, the linear MAX approximation attempts to approximate $MAX(x, y)$ as a linear function of x and y, i.e.,

$$MAX(x, y) = \alpha \cdot x + \beta \cdot y + \gamma. \quad (12)$$

The coefficients α and β can be determined by the following probabilities [19]:

$$\alpha = P(x \geq y)$$
$$\beta = P(y \geq x). \quad (13)$$

$P(x \geq y)$ and $P(y \geq x)$ in (13) are referred to as the *tightness probabilities* of x and y [19], respectively. Intuitively, (13) implies that the weight for x (or y) is large if x (or y) will likely be greater than y (or x). The constant term γ in (12) can be determined by matching the mean value [19] as follows:

$$\gamma = E\left[MAX(x, y)\right] - \alpha \cdot E[x] - \beta \cdot E[y] \quad (14)$$

where $E(\bullet)$ stands for the expected value.

Li *et al.* [22] further prove that the tightness probabilities in (13) are equal to the following first-order statistical sensitivities:

$$\alpha = P(x \geq y) = \frac{\partial\left\{E[MAX(x, y)]\right\}}{\partial\left\{E[x]\right\}}$$
$$\beta = P(y \geq x) = \frac{\partial\left\{E\left[MAX(x, y)\right]\right\}}{\partial\left\{E[y]\right\}}. \quad (15)$$

Although the MAX operator is not analytical (i.e., does not have continuous derivatives), it can statistically be approximated as in the form of (12)–(15), which is similar to the traditional Taylor expansion. We refer to (12) as the *linear statistical Taylor expansion* in this paper.

The linear statistical Taylor expansion has widely been applied to statistical timing analysis, where digital delay variations can accurately be captured by such a linear approximation. Most analog/RF performance variations, however, are strongly nonlinear in the presence of large-scale process variations and cannot accurately be approximated by the aforementioned linear models. To intuitively illustrate the limitation of the linear MAX approximation, we consider a simple example where two random variables $x \sim N(0, 1/9)$ and $y \sim N(0, 1)$ are independent and normal. Fig. 2 shows the probability density functions of x, y, and $MAX(x, y)$. In this example, $MAX(x, y)$ is strongly nonlinear, because the probability density functions of x and y are significantly overlapped. As a result, the simple linear MAX approximation yields a large error compared to the Monte Carlo analysis result (10^6 samples) in Fig. 2. This observation, therefore, motivates us to propose a novel quadratic statistical MAX approximation in this paper to improve the approximation accuracy.

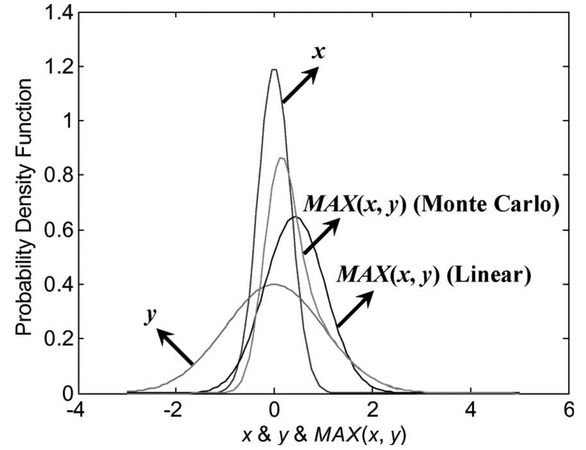


Fig. 2. Probability density functions of x, y, and $MAX(x, y)$.

III. Parametric Yield Estimation

Our proposed parametric yield estimation conceptually maps multiple correlated performance constraints to a single auxiliary constraint by using a MAX operator. The auxiliary constraint is then analytically approximated as a quadratic function of process parameters. As such, the parametric yield can easily be estimated from the probability distribution of the auxiliary constraint. In this section, we first mathematically define the auxiliary constraint and then derive two novel algorithms from moment matching and statistical Taylor expansion, respectively, for quadratic MAX approximation.

A. Parametric Yield Formulation

Given the parametric yield formulation in (8), we conceptually define the auxiliary performance constraint as follows:

$$f_{\text{aux}}(\varepsilon) = MAX\left[f_1(\varepsilon), f_2(\varepsilon), \ldots, f_K(\varepsilon)\right]. \quad (16)$$

Equation (16) maps multiple performance metrics to a single auxiliary performance metric by using a MAX operator. It is straightforward to verify that the parametric yield defined in (8) can uniquely be determined by the following auxiliary constraint:

$$\text{Yield} = P\left[f_{\text{aux}}(\varepsilon) \leq 0\right]. \quad (17)$$

If the auxiliary performance f_{aux} is approximated as a quadratic function of ε, the probability distribution of $f_{\text{aux}}(\varepsilon)$ and, consequently, the parametric yield can easily be estimated by the APEX algorithm proposed in [15] and [16].

Next, we propose two novel algorithms for efficiently approximating the MAX operator in (16). We will focus on the MAX approximation of two random variables, i.e., $MAX(x, y)$, since any multivariable operation can be broken down into multiple two-variable cases. Such a two-variable operator $MAX(x, y)$ can further be transformed to a single-variable operator as follows:

$$MAX(x, y) = x + MAX(0, z) \quad (18)$$

where

$$z = y - x. \tag{19}$$

Therefore, the remainder of this section will derive the quadratic statistical MAX approximation for the single-variable operator $MAX(0, z)$.

B. Quadratic MAX Approximation by Moment Matching

We approximate the MAX operator as the following quadratic function:

$$MAX(0, z) \approx 0.5 \cdot \sigma_2 \cdot z^2 + \sigma_1 \cdot z + \sigma_0 \tag{20}$$

where σ_2, σ_1, and σ_0 are the unknown model coefficients. These coefficients can be determined by matching the following moments:

$$E[MAX(0, z)] = 0.5 \cdot \sigma_2 \cdot E[z^2] + \sigma_1 \cdot E[z] + \sigma_0 \tag{21}$$

$$E[z \cdot MAX(0, z)] = 0.5 \cdot \sigma_2 \cdot E[z^3] + \sigma_1 \cdot E[z^2] + \sigma_0 \cdot E[z] \tag{22}$$

$$E\left[z^2 \cdot MAX(0, z)\right] = 0.5 \cdot \sigma_2 \cdot E[z^4] + \sigma_1 \cdot E[z^3] + \sigma_0 \cdot E[z^2]. \tag{23}$$

Since z is equal to $y - x$ in (19), where x and y are two performance metrics approximated as quadratic models, z is also a quadratic function of the random variable ε, i.e.,

$$z(\varepsilon) = \varepsilon^T \cdot A_z \cdot \varepsilon + B_z^T \cdot \varepsilon + C_z \tag{24}$$

where A_z, B_z, and C_z are the model coefficients. Given (24), the high-order moments $E[z]$, $E[z^2]$, $E[z^3]$, and $E[z^4]$ can be computed by the binominal moment evaluation algorithm proposed in [15] and [16]. In addition, the probability distribution of z can be extracted using APEX [15], [16], and consequently, $E[MAX(0, z)]$, $E[z(MAX(0, z)]$, and $E[z^2(MAX(0, z)]$ can be calculated using the following 1-D numerical integrations:

$$E[MAX(0, z)] = \int_0^{+\infty} z \cdot PDF_z(z) \cdot dz \tag{25}$$

$$E[z \cdot MAX(0, z)] = \int_0^{+\infty} z^2 \cdot PDF_z(z) \cdot dz \tag{26}$$

$$E\left[z^2 \cdot MAX(0, z)\right] = \int_0^{+\infty} z^3 \cdot PDF_z(z) \cdot dz \tag{27}$$

where $PDF_z(\bullet)$ denotes the probability density function of z. Given these high-order moments, the linear equation (21)–(23) can be solved to determine σ_2, σ_1, and σ_0. The $MAX(0, z)$ in (20) can then be approximated as a quadratic function of the random variable ε by substituting (24) into (20) and ignoring all high-order terms. Once the quadratic function $MAX(0, z)$ is determined, it can be substituted into (18) to calculate the quadratic model for $MAX(x, y)$.

It should be noted that our moment matching in (21)–(23) is substantially different from the algorithm proposed in [23]. Zhan *et al.* [23] attempt to match the moments for all random variables $\{\varepsilon_i; i = 1, 2, \ldots, N\}$ and their cross-product terms, whereas we match the moments only for a single random variable z, thereby significantly reducing the computational cost. In addition, the following theorem proves that the proposed moment matching results in an optimal approximation that minimizes the weighted squared error. The detailed proof of Theorem 1 is given in the Appendix.

Theorem 1: Given a random variable z and the quadratic statistical MAX approximation in (20), the coefficients σ_2, σ_1, and σ_0 determined by (21)–(23) minimize the weighted squared error, i.e.,

$$\Delta(\sigma_2, \sigma_1, \sigma_0) = \int_{-\infty}^{+\infty} \left[MAX(0, z) - (0.5 \cdot \sigma_2 \cdot z^2 + \sigma_1 \cdot z + \sigma_0)\right]^2 \cdot PDF_z(z) \cdot dz. \tag{28}$$

C. Quadratic MAX Approximation by Taylor Expansion

In addition to moment matching, an alternative method for quadratic MAX approximation can be derived from statistical Taylor expansion. In this section, we develop the algorithm for quadratic statistical Taylor expansion and then show the theoretical connection between the proposed two MAX approximation schemes.

Extending the linear statistical Taylor expansion in (12)–(15) to second-order and expanding $MAX(0, z)$ at the expansion point $E[z]$ yield

$$MAX(0, z) \approx 0.5 \cdot \lambda_2 \cdot \{z - E[z]\}^2 + \lambda_1 \cdot \{z - E[z]\} + \lambda_0 \tag{29}$$

where the linear and quadratic coefficients λ_1 and λ_2 are determined by the following statistical derivatives:

$$\lambda_1 = \frac{d\{E[MAX(0, z)]\}}{d\{E[z]\}} \tag{30}$$

$$\lambda_2 = \frac{d^2\{E[MAX(0, z)]\}}{d\{E[z]\}^2} = \frac{d\lambda_1}{d\{E[z]\}}. \tag{31}$$

The constant term λ_0 is determined by matching the mean value as follows:

$$\lambda_0 = E[MAX(0, z)] - 0.5 \cdot \lambda_2 \cdot E\left[\{z - E[z]\}^2\right]. \tag{32}$$

Next, we show how we can efficiently compute the coefficients λ_0, λ_1, and λ_2 in (30)–(32).

1) Linear Coefficient λ_1: As described in Section II-D, the first-order derivative in (30) is equal to the following

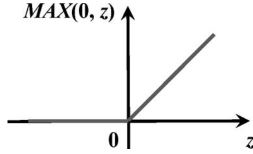


Fig. 3. Single variable function $MAX(0, z)$ is convex.

probability [22]:

$$\lambda_1 = \frac{d\{E[MAX(0,z)]\}}{d\{E[z]\}} = P(z \geq 0) = 1 - CDF_z(0) \tag{33}$$

where $CDF_z(\bullet)$ stands for the cumulative distribution function of z. Given the quadratic model $z(\varepsilon)$ in (24), $CDF_z(\bullet)$ can be extracted using APEX [15], [16].

2) Quadratic Coefficient λ_2*:* Substituting (33) into (31) yields

$$\lambda_2 = \frac{d[1 - CDF_z(0)]}{d\{E[z]\}}. \tag{34}$$

To calculate the derivative in (34), we rewrite z as follows:

$$z = \mu + \delta \tag{35}$$

where μ is the mean value of z, and $\delta = z - \mu$ is a random variable with zero mean. Substituting (35) into (34) yields

$$\begin{aligned}\lambda_2 &= \frac{d[1 - CDF_{\mu+\delta}(0)]}{d\mu} = \frac{d[1 - CDF_\delta(-\mu)]}{d\mu} \\ &= PDF_\delta(-\mu) = PDF_{\mu+\delta}(0) = PDF_z(0)\end{aligned} \tag{36}$$

where the probability density function $PDF_z(\bullet)$ can be extracted using APEX [15], [16].

It is worth noting that the quadratic coefficient λ_2 in (36) has two interesting properties.

1) $\lambda_2 = PDF_z(0)$ *is nonnegative*. Intuitively, as shown in Fig. 3, the function $MAX(0, z)$ is convex, and therefore, the quadratic model coefficient should be nonnegative [29].
2) $\lambda_2 = PDF_z(0)$ *indicates the nonlinearity*. For the first two cases in Fig. 4, $MAX(0, z)$ can accurately be approximated as the linear models, i.e., $MAX(0, z) \approx 0$ and $MAX(0, z) \approx z$, respectively. This is consistent with the fact that $PDF_z(0) \approx 0$ in both cases. In the third case in Fig. 4, however, $MAX(0, z)$ is strongly nonlinear, corresponding to a nonzero $PDF_z(0)$.

3) Constant Term λ_0*:* After λ_1 and λ_2 are known, computing the constant term λ_0 in (32) requires further knowing $E[MAX(0, z)]$ and $E[\{z - E[z]\}^2]$. $E[MAX(0, z)]$ can be calculated using the 1-D numerical integration in (25). Since z is a quadratic function of ε, as shown in (24), its second-order central moment is determined by the following analytical equation [21]:

$$\begin{aligned}E\left[\{z - [z]\}^2\right] = B_z^T \cdot \Sigma \cdot B_z + 2 \\ \cdot TRACE(\Sigma \cdot A_z \cdot \Sigma \cdot A_z)\end{aligned} \tag{37}$$

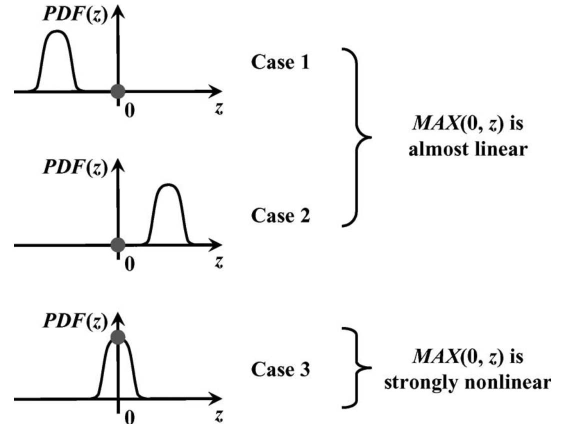


Fig. 4. Three different cases for quadratic MAX approximation.

where $TRACE(\bullet)$ represents the trace of a matrix (the sum of all diagonal elements), and Σ stands for the covariance matrix of the N-dimensional random variable ε. Substituting (25), (36), and (37) into (32) yields the constant term λ_0. After the coefficients λ_0, λ_1, and λ_2 are known, the $MAX(0, z)$ in (29) can be approximated as a quadratic function of ε by substituting (24) into (29) and ignoring all high-order terms. Once the quadratic function $MAX(0, z)$ is determined, it can be substituted into (18) to calculate the quadratic model for $MAX(x, y)$.

It can be proven that the proposed quadratic statistical Taylor expansion is mathematically equivalent to the moment matching approach if the random variable z is normal. This conclusion can formally be stated as the following theorem. The detailed proof of Theorem 2 is given in the Appendix.

Theorem 2: If the random variable z is normal, the quadratic MAX approximations in (20) (based on moment matching) and (29) (based on statistical Taylor expansion) are equivalent.

If the random variable z is nonnormal, the theoretical connection between moment matching and statistical Taylor expansion is difficult to find. However, for many practical applications, these two approaches are equally efficient in both approximation accuracy and computational cost, as will be demonstrated by the numerical examples in Section IV.

D. Summary

A simplified flow of the proposed parametric yield estimation is summarized in Algorithm 1. Our proposed algorithm only involves APEX [15], [16], the 1-D numerical integration in (25)–(27), and a couple of other simple analytical operations and, therefore, is computationally efficient. As will be demonstrated by the numerical examples in Section IV, the proposed parametric yield estimation can achieve a runtime speedup of 10–20 times over the Monte Carlo analysis with 10^3 samples.

Algorithm 1: Parametric Yield Estimation.

(1) Start from a set of quadratic response surface models $\{f_k(\varepsilon); k = 1, 2, \ldots, K\}$.
(2) Represent all performance constraints as in the standard form in (7).
(3) Set $f_{\text{aux}}(\varepsilon) = f_1(\varepsilon)$.

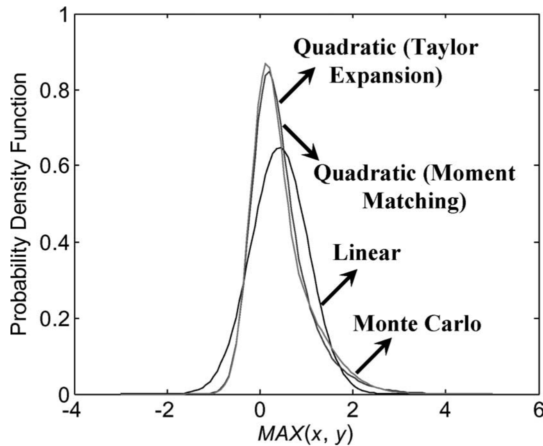


Fig. 5. Estimated probability density functions of $MAX(x, y)$.

(4) For $k = 2, 3, \ldots, K$
(5) Represent $z(\varepsilon) = f_{\mathrm{aux}}(\varepsilon) - f_k(\varepsilon)$ as in the form of (24).
(6) Compute the quadratic approximation of $MAX(0, z)$ by either moment matching (Section III-B) or statistical Taylor expansion (Section III-C).
(7) Compute $f_{\mathrm{aux}}(\varepsilon) = f_k(\varepsilon) + MAX(0, z)$.
(8) End for
(9) Compute the probability distribution of $f_{\mathrm{aux}}(\varepsilon)$ by using APEX [15], [16].
(10) Compute the parametric yield defined in (17).

IV. Numerical Examples

In this section, we demonstrate the efficacy of the proposed parametric yield estimation by several analog/RF circuit examples. In particular, we compare the proposed quadratic MAX approximation with several linear approximation techniques to demonstrate the improvement in accuracy. All numerical experiments are performed on a LINUX 2.8-GHz server.

A. Simple Example for MAX Approximation

We first consider the example in Fig. 2, where $x \sim N(0, 1/9)$ and $y \sim N(0, 1)$ are independent normal random variables. This simple example allows us to test the efficacy of the proposed MAX approximation and compare it with other traditional techniques.

Four different approaches, namely, the traditional linear MAX approximation (Section II-D), the quadratic MAX approximation that uses moment matching, the quadratic MAX approximation that uses statistical Taylor expansion, and the Monte Carlo analysis with 10^6 samples, are applied to estimate the probability distribution of $MAX(x, y)$. Fig. 5 shows the probability density functions estimated by these techniques. By studying Fig. 5, one would notice that the moment matching and the statistical Taylor expansion yield the same results (except for the numerical error). This observation is consistent with our theoretical result in Theorem 2.

In this example, the distribution of $MAX(x, y)$ is not symmetric due to the nonlinearity of the MAX operator. The traditional linear approximation cannot capture such a nonzero skewness and, therefore, results in a large approximation error.

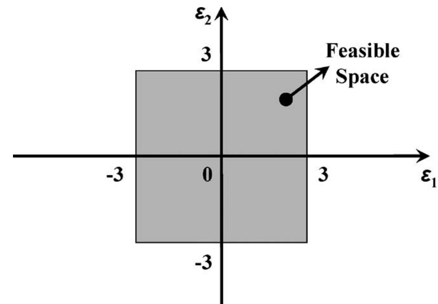


Fig. 6. Constraints in (38) define a 2-D square if N is equal to 2.

The proposed quadratic approximation, however, accurately models the nonzero skewness by including second-order terms.

B. Simple Example for Parametric Yield Estimation

To compare the proposed parametric yield estimation with the traditional ellipsoid approximation (Section II-C), we consider the following constraints:

$$\begin{cases} \varepsilon_1 \geq -3 & \& \quad \varepsilon_1 \leq 3 \\ \varepsilon_2 \geq -3 & \& \quad \varepsilon_2 \leq 3 \\ \quad\vdots & \vdots \qquad \vdots \\ \varepsilon_N \geq -3 & \& \quad \varepsilon_N \leq 3 \end{cases} \tag{38}$$

where $\{\varepsilon_i; i = 1, 2, \ldots, N\}$ are independent and standard normal, and N is the total number of random variables (i.e., the problem dimension). The constraints in (38) define a hypercube in the N-dimensional space. For illustration purposes, Fig. 6 shows a 2-D square, where $N = 2$.

Given the constraints in (38), it is easy to verify that the parametric yield can be analytically computed as follows [27]:

$$\text{Yield} = [\Phi(3) - \Phi(-3)]^N \tag{39}$$

where $\Phi(\bullet)$ stands for the cumulative distribution function of standard normal distribution. For testing and comparison purposes, three different approaches, namely, the traditional ellipsoid approximation (Section II-C), the quadratic MAX approximation that uses moment matching, and the quadratic MAX approximation that uses statistical Taylor expansion, are applied to estimate the parametric yield. The absolute difference between the estimated yield and the actual yield in (39) is used as a measure of the estimation error for accuracy comparison. Fig. 7 shows the estimation error as a function of the problem dimension N. Two important observations can be made from the data in Fig. 7.

First, in this example, although all performance constraints in (38) are linear and, therefore, their probability distributions are normal, the moment matching and the statistical Taylor expansion do not yield exactly identical results. This is because for any $N \geq 2$, multiple two-variable MAX operations are involved. Taking $N = 2$ as an example, the auxiliary performance constraint in (16) is defined as follows:

$$f_{\mathrm{aux}}(\varepsilon) = MAX[-\varepsilon_1 - 3, \varepsilon_1 - 3, -\varepsilon_2 - 3, \varepsilon_2 - 3]. \tag{40}$$

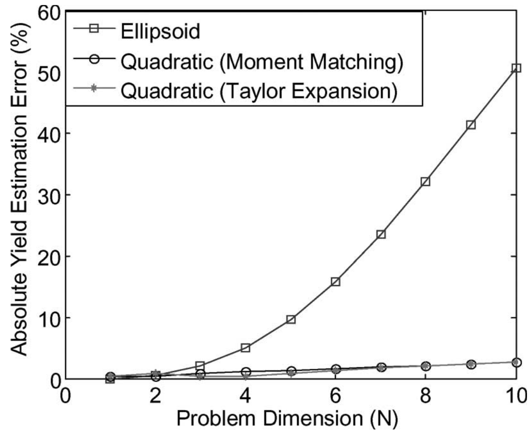


Fig. 7. Absolute yield estimation error varies as a function of the problem dimension N.

This four-variable MAX operator can be broken down into the following three two-variable MAX operators:

$$g_1(\varepsilon) = MAX[-\varepsilon_1 - 3, \varepsilon_1 - 3] \quad (41)$$

$$g_2(\varepsilon) = MAX[g_1(\varepsilon), -\varepsilon_2 - 3] \quad (42)$$

$$f_{\text{aux}}(\varepsilon) = MAX[g_2(\varepsilon), \varepsilon_2 - 3]. \quad (43)$$

It is easy to verify that only the first MAX operator in (41) computes the maximum of two normal distributions: 1) $-\varepsilon_1 - 3$ and 2) $\varepsilon_1 - 3$. The other two MAX operators in (42) and (43) take the nonnormal distributions $g_1(\varepsilon)$ and $g_2(\varepsilon)$ as the input, respectively. In this case, Theorem 2 is not applicable. However, as shown in Fig. 7, the estimation error of the moment matching and the statistical Taylor expansion remains close, although the theoretical connection between these two approaches is not clear.

Second, the error of the ellipsoid approximation continuously increases with N. As shown in Fig. 1, the maximal inscribed ellipsoid does not cover the entire feasible space. If the feasible space is high dimensional, the ellipsoid approximation error can extremely be large. For $N = 10$ in Fig. 7, the absolute yield estimation error of the ellipsoid approximation reaches 50.54%. It is 18.9 times larger than the error of our proposed method, which is based on quadratic MAX approximation.

C. Low-Noise Amplifier

Fig. 8 shows a low-noise amplifier designed in a commercial 90-nm CMOS process. In this example, the variations of both MOS transistors and passive components (i.e., resistors, capacitors, and inductors) are considered. The probability distribution and the correlation information of these variations are specified in the process design kit from the foundry. After PCA, nine principal factors are identified to model the process variations.

1) Response Surface Modeling: The performance of the low-noise amplifier is characterized by seven specifications. Given a fixed circuit design, each circuit performance is a function of process variations. We approximate these performance functions by linear and quadratic response surface models, respectively. The linear models are extracted by the direct

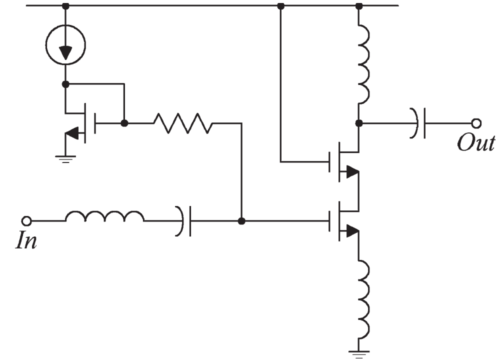


Fig. 8. Circuit schematic of a low-noise amplifier designed in a commercial 90-nm CMOS process.

TABLE I
RESPONSE SURFACE MODELING ERROR FOR THE LOW-NOISE AMPLIFIER

Performance	Linear Model Error	Quadratic Model Error
F0	2.64%	2.33%
Gain	4.08%	0.45%
S11	6.55%	1.69%
S22	9.84%	4.99%
NF	9.05%	2.73%
IIP3	6.21%	1.04%
Power	2.45%	0.04%

TABLE II
MINIMAL/MAXIMAL NORMALIZED EIGENVALUES OF THE QUADRATIC MODEL COEFFICIENT MATRIX FOR THE LOW-NOISE AMPLIFIER

Performance	Minimal Eigenvalue	Maximal Eigenvalue
F0	-0.37	0.67
Gain	-0.79	0.02
S11	-0.72	0.58
S22	-0.73	0.65
NF	-0.07	0.99
IIP3	-0.99	0.07
Power	-0.01	1.00

fitting method [28], and the quadratic models are extracted by the PROBE algorithm proposed in [25]. In this example, the linear and quadratic modeling takes 37 and 75 s, respectively, including the computational time for transistor-level simulation to collect all sampling points. Table I shows the response surface modeling error for all performance metrics. Note that the quadratic modeling error is four times smaller than the linear modeling error on the average.

Table II shows the minimal and maximal normalized eigenvalues of the quadratic coefficient matrix, i.e., the matrix A in (4). For a given performance function $f(\varepsilon)$, A is related to the Hessian matrix $\nabla^2 f(\varepsilon)$ (i.e., the second-order derivative). Therefore, it can be used to test the convexity of the performance function $f(\varepsilon)$. For instance, A should be positive semidefinite (or negative semidefinite) if $f(\varepsilon)$ is convex (or concave) [29]. For most performance metrics in this example, the matrix A contains both positive and negative eigenvalues with comparable magnitude. It, in turn, implies that these performance functions are neither convex nor concave. Hence, the feasible space defined by the performance constraints is not a convex set.

2) Parametric Yield Estimation: For testing and comparison purposes, we randomly select 100 different design

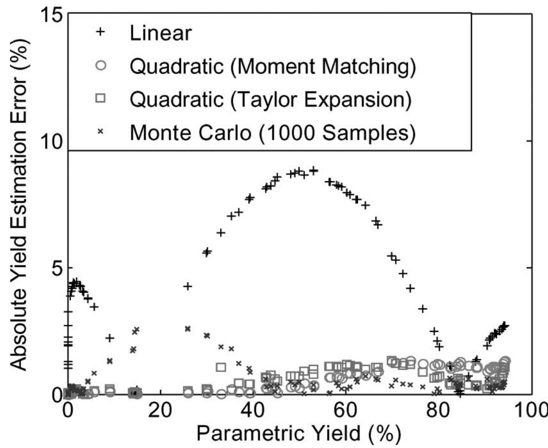


Fig. 9. Absolute yield estimation error for the low-noise amplifier.

TABLE III
YIELD ESTIMATION COST FOR THE LOW-NOISE AMPLIFIER (ONE DESIGN POINT)

Algorithm	Computational Time (Sec.)
Linear	0.0016
Quadratic (Moment Matching)	0.0120
Quadratic (Taylor Expansion)	0.0128
Monte Carlo (10^3 Samples)	0.1382

specifications, and the parametric yield is estimated for each of these specifications by four different approaches: 1) the traditional linear MAX approximation (Section II-D); 2) the quadratic MAX approximation that uses moment matching; 3) the quadratic MAX approximation that uses statistical Taylor expansion; and 4) the Monte Carlo analysis with 10^3 samples. The estimated parametric yield is compared to the golden result extracted by the Monte Carlo analysis with 10^6 samples. Their absolute difference is used as a measure of the estimation error for accuracy comparison.

Fig. 9 shows the absolute yield estimation error of the four different approaches. The traditional linear approximation cannot accurately capture the parametric yield, and the maximal error reaches 8.82%. The proposed quadratic approximation (both moment matching and statistical Taylor expansion) achieves much better accuracy by reducing the maximal error to 1.32% (i.e., 6.68 times smaller). On the average, the proposed quadratic approximation is as accurate as the Monte Carlo analysis with 10^3 samples in this example.

Table III compares the computational cost of the four different approaches. The traditional linear approximation has the lowest computational cost; however, it cannot provide high estimation accuracy. In this example, the proposed quadratic approximation (both moment matching and statistical Taylor expansion) achieves a runtime speedup of ten times over the Monte Carlo analysis with 10^3 samples while offering the same accuracy.

Finally, it should be noted that the yield estimation cost shown in Table III is the computational time required to estimate the parametric yield of a single design point. As we know, if the parametric yield estimation is within a statistical optimization loop, it must be repeated many times. In this case, the overall computational time for yield estimation may be expensive, even if one estimation run is cheap. For example, assuming that one complete synthesis run requires visiting 10^5 design points [24] and the parametric yield estimation of one design point takes 0.13 s by using the Monte Carlo analysis with 10^3 samples, the total yield estimation cost will reach 3.8 h. However, if the proposed quadratic approximation is applied, the total yield estimation cost can be reduced to 20 min in this example.

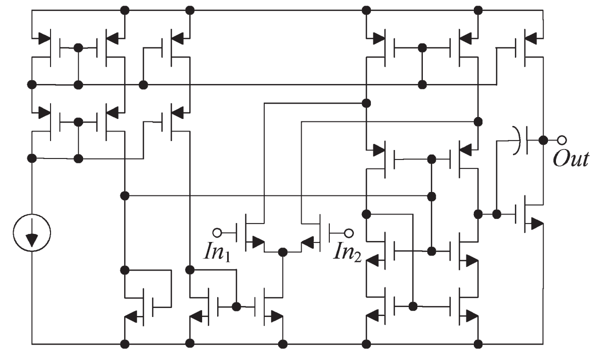


Fig. 10. Circuit schematic of an operational amplifier designed in a commercial 0.25-μm BiCMOS process.

TABLE IV
RESPONSE SURFACE MODELING ERROR FOR THE OPERATIONAL AMPLIFIER

Performance	Linear Model Error	Quadratic Model Error
Offset	0.59%	0.06%
Gain	14.65%	3.75%
Unit Gain Frequency	2.30%	0.39%
Phase Margin	2.35%	0.37%
Power	1.23%	0.40%

D. Operational Amplifier

Fig. 10 shows a two-stage folded-cascode operational amplifier designed in a commercial 0.25-μm BiCMOS process. In this example, both interdie variations and device mismatches are considered. The probability distribution and the correlation information of these variations are specified in the process design kit from the foundry. After PCA, 49 principal factors are identified to model the process variations. It is important to note that modeling mismatches are extremely important for this operational amplifier example, since they can significantly impact several circuit performances, e.g., offset voltage.

1) Response Surface Modeling: The performance of the operational amplifier is characterized by five specifications. Given a fixed circuit design, each circuit performance is a function of process variations. We approximate these performance functions by linear and quadratic response surface models, respectively. The linear models are extracted by the direct fitting method [28], and the quadratic models are extracted by the PROBE algorithm proposed in [25]. In this example, the linear and quadratic modeling takes 21 and 46 s, respectively, including the computational time for transistor-level simulation for collecting all sampling points. Table IV shows the response surface modeling error for all performance metrics. Compared with the linear response surface modeling, the quadratic response surface modeling reduces the maximal approximation

TABLE V
MINIMAL/MAXIMAL NORMALIZED EIGENVALUES OF THE QUADRATIC MODEL COEFFICIENT MATRIX FOR THE OPERATIONAL AMPLIFIER

Performance	Minimal Eigenvalue	Maximal Eigenvalue
Offset	-0.67	0.66
Gain	-0.94	0.10
Unit Gain Frequency	-0.43	0.88
Phase Margin	-0.89	0.40
Power	-0.82	0.36

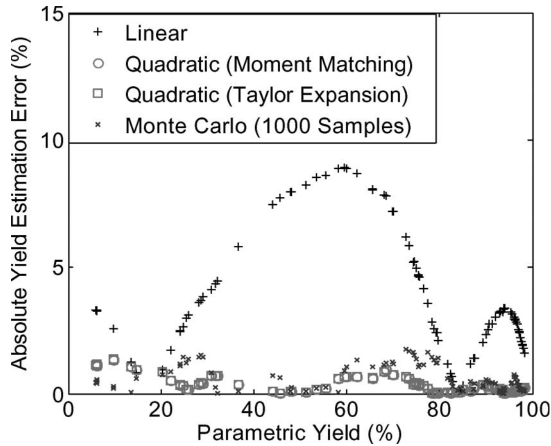


Fig. 11. Absolute yield estimation error for the operational amplifier.

error from 14.65% to 3.75% (i.e., 3.9 times smaller). Note that even though the operational amplifier is designed in a 0.25-μm process, the linear response surface modeling is not sufficiently accurate, as shown in Table IV. As IC technologies are scaled to finer feature sizes, process variations will become relatively larger, thereby making the nonlinear terms in the quadratic model even more important.

Table V shows the minimal and maximal normalized eigenvalues of the quadratic coefficient matrix, i.e., the matrix A in (4). Similar to the low-noise amplifier example in Section IV-C, most operational amplifier performance functions are neither convex nor concave. Hence, the feasible space defined by the performance constraints is not a convex set.

2) Parametric Yield Estimation: For testing and comparison purposes, we randomly select 100 different design specifications, and the parametric yield is estimated for each of these specifications by four different approaches: 1) the traditional linear MAX approximation (Section II-D); 2) the quadratic MAX approximation that uses moment matching; 3) the quadratic MAX approximation that uses statistical Taylor expansion; and 4) the Monte Carlo analysis with 10^3 samples. The estimated parametric yield is compared to the golden result extracted by the Monte Carlo analysis with 10^6 samples. Their absolute difference is used a measure of the estimation error for accuracy comparison.

Fig. 11 compares the yield estimation accuracy for the four different approaches. The traditional linear approximation results in a maximal error of 8.92%. The proposed quadratic approximation (both moment matching and statistical Taylor expansion) is much more accurate, and it reduces the maximal error to 1.37% (i.e., 6.51 times smaller). On the average, the proposed quadratic approximation is as accurate as the Monte Carlo analysis with 10^3 samples in this example.

TABLE VI
YIELD ESTIMATION COST FOR THE OPERATIONAL AMPLIFIER (ONE DESIGN POINT)

Algorithm	Computational Time (Sec.)
Linear	0.0005
Quadratic (Moment Matching)	0.0059
Quadratic (Taylor Expansion)	0.0069
Monte Carlo (10^3 Samples)	0.1300

Table VI shows the computational cost of the four different approaches. In this example, the proposed quadratic approximation (both moment matching and statistical Taylor expansion) achieves a runtime speedup of 20 times over the Monte Carlo analysis with 10^3 samples while offering the same accuracy.

V. CONCLUSION

We have proposed an efficient parametric yield estimation algorithm for multiple correlated nonnormal analog/RF performance distributions. The proposed algorithm is facilitated by two key techniques, including: 1) an auxiliary constraint formulation that uses a MAX operator and 2) two efficient algorithms (based on moment matching and statistical Taylor expansion, respectively) for quadratic statistical MAX approximation. We prove that these two quadratic MAX approximation algorithms are mathematically equivalent if the performance distributions are normal. Our numerical examples demonstrate that, compared to the traditional normal-distribution-based method, the proposed parametric yield estimation reduces the estimation error by 6.5 times while achieving a runtime speedup of 10–20 times over the Monte Carlo analysis with 10^3 samples. Our future research will further study the theoretical connection between the proposed quadratic MAX approximation algorithms for nonnormal performance distributions. In addition, we will incorporate the proposed yield estimation algorithm into a statistical optimization flow for robust analog/RF design.

The quadratic statistical MAX approximation proposed in this paper is not limited only to the parametric yield estimation of analog/RF circuits. The same idea can widely be applied to many other applications that require MAX operations, e.g., the statistical static timing analysis for digital circuits.

APPENDIX
PROOF OF THEOREM 1

Given the weighted squared error in (28), the optimal values of σ_2, σ_1, and σ_0 can be determined by the following optimization:

$$\underset{\sigma_1, \sigma_2, \sigma_3}{\text{minimize}} \quad \Delta(\sigma_2, \sigma_1, \sigma_0) \tag{44}$$

where $\Delta(\sigma_2, \sigma_1, \sigma_0)$ is given in (28). Equation (44) defines an unconstrained nonlinear programming problem. The cost function $\Delta(\sigma_2, \sigma_1, \sigma_0)$ in (44) is a quadratic function of σ_2, σ_1, and σ_0. In addition, $\Delta(\sigma_2, \sigma_1, \sigma_0)$ is always nonnegative for any σ_2, σ_1, $\sigma_0 \in R$. In other words, the quadratic cost

function $\Delta(\sigma_2, \sigma_1, \sigma_0)$ is positive semidefinite and, therefore, is convex [29].

Since the cost function in (44) is quadratic and convex, it does not contain any local minimum, and its global optimum can be determined by the following first-order gradient:

$$\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_2} = 0 \quad (45)$$

$$\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_1} = 0 \quad (46)$$

$$\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_0} = 0. \quad (47)$$

We rewrite the weighted squared error in (28) as in the form of (56). Substituting (56) into (45)–(47) yields (57)–(59). Note that the optimality conditions in (57)–(59) are exactly equivalent to the moment matching equations in (21)–(23). Therefore, the moment matching in (21)–(23) yields the optimal σ_2, σ_1, and σ_0 that minimize the weighted squared error in (28).

APPENDIX
PROOF OF THEOREM 2

Without loss of generality, we assume that the random variable z satisfies the normal distribution $N(\mu, \sigma^2)$, where the mean value is μ, and the standard deviation is σ. In this case, the high-order moments $E[z]$, $E[z^2]$, $E[z^3]$, and $E[z^4]$ in (21)–(23) can be represented as functions of μ and σ [27], i.e.,

$$E[z] = \mu \quad (48)$$

$$E[z^2] = \mu^2 + \sigma^2 \quad (49)$$

$$E[z^3] = \mu^3 + 3\mu\sigma^2 \quad (50)$$

$$E[z^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4. \quad (51)$$

In addition, the high-order moments $E[MAX(0, z)]$, $E[z(MAX(0, z)]$, and $E[z^2(MAX(0, z)]$ in (21)–(23) can be computed by the integrals in (60)–(62), where $\varphi(\bullet)$ and $\Phi(\bullet)$ denote the probability density function and the cumulative distribution function of the standard normal distribution, respectively. Substituting (48)–(51) and (60)–(62) into (20)–(23) yields the quadratic MAX approximation in (63).

On the other hand, the statistical Taylor expansion coefficients in (36), (33), and (32) are equal to

$$\lambda_2 = PDF_z(0) = \frac{1}{\sigma} \cdot \varphi\left(\frac{\mu}{\sigma}\right) \quad (52)$$

$$\lambda_1 = 1 - CDF_z(0) = 1 - \Phi\left(\frac{-\mu}{\sigma}\right) \quad (53)$$

$$\lambda_0 = E[MAX(0, z)] - 0.5 \cdot \lambda_2 \cdot \sigma^2. \quad (54)$$

Substituting (60) and (52) into (54) yields

$$\lambda_0 = \frac{\sigma}{2} \cdot \varphi\left(\frac{\mu}{\sigma}\right) + \mu \cdot \left[1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right]. \quad (55)$$

Substitute (52)–(53) and (55) into (29). After some mathematical manipulations, we obtain the approximated quadratic function in (63). Therefore, if z is normal, the moment matching and the statistical Taylor expansion are exactly equivalent

$$\begin{aligned}
&\Delta(\sigma_2, \sigma_1, \sigma_0) \\
&= \int_{-\infty}^{+\infty} \Big[MAX^2(0, z) + \frac{\sigma_2^2 \cdot z^4}{4} + \sigma_1^2 \cdot z^2 + \sigma_0^2 \\
&\qquad - \sigma_2 \cdot z^2 \cdot MAX(0, z) - 2 \cdot \sigma_1 \cdot z \cdot MAX(0, z) \\
&\qquad - 2 \cdot \sigma_0 \cdot MAX(0, z) + \sigma_2 \cdot \sigma_1 \cdot z^3 \\
&\qquad + \sigma_2 \cdot \sigma_0 \cdot z^2 + 2 \cdot \sigma_1 \cdot \sigma_0 \cdot z \Big] \cdot PDF_z(z) \cdot dz \\
&= E\left[MAX^2(0, z)\right] + \frac{\sigma_2^2}{4} \cdot E[z^4] + \sigma_1^2 \cdot E[z^2] \\
&\quad + \sigma_0^2 - \sigma_2 \cdot E\left[z^2 \cdot MAX(0, z)\right] \\
&\quad - 2 \cdot \sigma_1 \cdot E[z \cdot MAX(0, z)] - 2 \cdot \sigma_0 \cdot E[MAX(0, z)] \\
&\quad + \sigma_2 \cdot \sigma_1 \cdot E[z^3] + \sigma_2 \cdot \sigma_0 \cdot E[z^2] + 2 \cdot \sigma_1 \cdot \sigma_0 \cdot E[z]
\end{aligned} \quad (56)$$

$$\begin{aligned}
&\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_2} \\
&= 0.5 \cdot \sigma_2 \cdot E[z^4] + \sigma_1 \cdot E[z^3] + \sigma_0 \cdot E[z^2] \\
&\quad - E\left[z^2 \cdot MAX(0, z)\right] = 0
\end{aligned} \quad (57)$$

$$\begin{aligned}
&\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_1} \\
&= \sigma_2 \cdot E[z^3] + 2 \cdot \sigma_1 \cdot E[z^2] + 2 \cdot \sigma_0 \cdot E[z] \\
&\quad - 2 \cdot E[z \cdot MAX(0, z)] = 0
\end{aligned} \quad (58)$$

$$\begin{aligned}
&\frac{\partial \Delta(\sigma_2, \sigma_1, \sigma_0)}{\partial \sigma_0} \\
&= \sigma_2 \cdot E[z^2] + 2 \cdot \sigma_1 \cdot E[z] + 2 \cdot \sigma_0 \\
&\quad - 2 \cdot E[MAX(0, z)] = 0
\end{aligned} \quad (59)$$

$$\begin{aligned}
&E[MAX(0, z)] \\
&= \int_{0}^{+\infty} z \cdot PDF_z(z) \cdot dz \\
&= \sigma \cdot \varphi\left(\frac{\mu}{\sigma}\right) + \mu \cdot \left[1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right]
\end{aligned} \quad (60)$$

$$\begin{aligned}
&E[z \cdot MAX(0, z)] \\
&= \int_{0}^{+\infty} z^2 \cdot PDF_z(z) \cdot dz \\
&= \mu\sigma \cdot \varphi\left(\frac{\mu}{\sigma}\right) + (\mu^2 + \sigma^2) \cdot \left[1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right]
\end{aligned} \quad (61)$$

$$E\left[z^2 \cdot MAX(0,z)\right] = \int_{0}^{+\infty} z^3 \cdot PDF_z(z) \cdot dz = (\mu^2\sigma + 2\sigma^3) \cdot \varphi\left(\frac{\mu}{\sigma}\right) + (\mu^3 + 3\mu\sigma^2) \cdot \left[1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right] \tag{62}$$

$$MAX(0,z) \approx \frac{1}{2\sigma} \cdot \varphi\left(\frac{\mu}{\sigma}\right) \cdot z^2 + \left[1 - \Phi\left(\frac{-\mu}{\sigma}\right) - \frac{\mu}{\sigma} \cdot \varphi\left(\frac{\mu}{\sigma}\right)\right] \cdot z + \frac{\mu^2 + \sigma^2}{2\sigma} \cdot \varphi\left(\frac{\mu}{\sigma}\right). \tag{63}$$

REFERENCES

[1] X. Li and L. Pileggi, "Efficient parametric yield extraction for multiple correlated non-Normal performance distributions of analog/RF circuits," in *Proc. IEEE Des. Autom. Conf.*, 2007, pp. 928–933.

[2] Semiconductor Industry Associate, *International Technology Roadmap for Semiconductors*, 2004.

[3] M. Styblinski and J. Zhang, "Circuit performance variability reduction: Principles, problems, and practical solutions," in *Proc. IEEE Int. Conf. Comput-Aided Des.*, 1991, pp. 170–173.

[4] Z. Wang and S. Director, "An efficient yield optimization method using a two-step linear approximation of circuit performance," in *Proc. IEEE Eur. Des. Test Conf.*, 1994, pp. 567–571.

[5] F. Schenkel, M. Pronath, S. Zizala, R. Schwencker, H. Graeb, and K. Antreich, "Mismatch analysis and direct yield optimization by specwise linearization and feasibility-guided search," in *Proc. IEEE Des. Autom. Conf.*, 2001, pp. 858–863.

[6] A. Dharchoudhury and S. Kang, "Worse-case analysis and optimization of VLSI circuit performance," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 14, no. 4, pp. 481–492, Apr. 1995.

[7] E. Felt, S. Zanella, C. Guardiani, and A. Sangiovanni-Vincentelli, "Hierarchical statistical characterization of mixed-signal circuits using behavioral modeling," in *Proc. IEEE Int. Conf. Comput-Aided Des.*, 1996, pp. 374–380.

[8] X. Li, P. Gopalakrishnan, Y. Xu, and L. Pileggi, "Robust analog/RF circuit design with projection-based performance modeling," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 26, no. 1, pp. 2–15, Jan. 2007.

[9] H. Abdel-Malek and A. Hassan, "The ellipsoidal technique for design centering and region approximation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 10, no. 8, pp. 1006–1014, Aug. 1991.

[10] J. Wojciechowski and J. Vlach, "Ellipsoidal method for design centering and yield estimation," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 12, no. 10, pp. 1570–1579, Oct. 1993.

[11] K. Antreich, H. Graeb, and C. Wieser, "Circuit analysis and optimization driven by worst-case distances," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 13, no. 1, pp. 57–71, Jan. 1994.

[12] A. Seifi, K. Ponnambalam, and J. Vlach, "A unified approach to statistical design centering of integrated circuits with correlated parameters," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 46, no. 1, pp. 190–196, Jan. 1999.

[13] J. Jess, K. Kalafala, S. Naidu, R. Otten, and C. Visweswariah, "Statistical timing for parametric yield prediction of digital integrated circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 11, pp. 2376–2392, Nov. 2006.

[14] P. Feldmann and S. Director, "Integrated circuit quality optimization using surface integrals," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 12, no. 12, pp. 1868–1879, Dec. 1993.

[15] X. Li, J. Le, P. Gopalakrishnan, and L. Pileggi, "Asymptotic probability extraction for non-normal distributions of circuit performance," in *Proc. IEEE Int. Conf. Comput-Aided Des.*, 2004, pp. 2–9.

[16] X. Li, J. Le, P. Gopalakrishnan, and L. Pileggi, "Asymptotic probability extraction for nonnormal performance distributions," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 26, no. 1, pp. 16–37, Jan. 2007.

[17] C. Clark, "The greatest of a finite set of random variables," *Oper. Res.*, vol. 9, no. 2, pp. 145–162, Mar./Apr. 1961.

[18] H. Chang and S. Sapatnekar, "Statistical timing analysis under spatial correlations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 9, pp. 1467–1482, Sep. 2005.

[19] C. Visweswariah, K. Ravindran, K. Kalafala, S. Walker, S. Narayan, D. Beece, J. Piaget, N. Venkateswaran, and J. Hemmett, "First-order incremental block-based statistical timing analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 10, pp. 2170–2180, Oct. 2006.

[20] H. Chang, V. Zolotov, S. Narayan, and C. Visweswariah, "Parameterized block-based statistical timing analysis with non-Gaussian parameters, nonlinear delay functions," in *Proc. IEEE Des. Autom. Conf.*, 2005, pp. 71–76.

[21] L. Zhang, W. Chen, Y. Hu, J. Gubner, and C. Chen, "Correlation-preserved non-Gaussian statistical timing analysis with quadratic timing model," in *Proc. IEEE Des. Autom. Conf.*, 2005, pp. 83–88.

[22] X. Li, J. Le, M. Celik, and L. Pileggi, "Defining statistical sensitivity for timing optimization of logic circuits with large-scale process and environmental variations," in *Proc. IEEE Int. Conf. Comput-Aided Des.*, 2005, pp. 844–851.

[23] Y. Zhan, A. Strojwas, X. Li, L. Pileggi, D. Newmark, and M. Sharma, "Correlation-aware statistical timing analysis with non-Gaussian delay distributions," in *Proc. IEEE Des. Autom. Conf.*, 2005, pp. 77–82.

[24] H. Liu, A. Singhee, R. Rutenbar, and L. Carley, "Remembrance of circuit past: Macromodeling by data mining in large analog design spaces," in *Proc. IEEE Des. Autom. Conf.*, 2002, pp. 437–442.

[25] X. Li, J. Le, L. Pileggi, and A. Strojwas, "Projection-based performance modeling for inter/intra-die variations," in *Proc. IEEE Int. Conf. Comput.-Aided Des.*, 2005, pp. 721–727.

[26] G. Seber, *Multivariate Observations*. New York: Wiley, 1984.

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

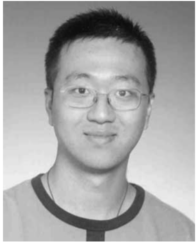
[28] R. Myers and D. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. New York: Wiley-Interscience, 2002.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Xin Li (S'01–M'06) received the B.S. and M.S. degrees in electronics engineering from Fudan University, Shanghai, China, in 1998 and 2001, respectively, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 2005.

He was an Intern with Extreme DA, Palo Alto, CA, in the summer of 2004. Since the summer of 2005, he has been a Systems Scientist with the Department of Electrical and Computer Engineering, Carnegie Mellon University. In the same year, he cofounded Xigmix Inc., a startup company in Pittsburgh, PA, for developing parametric yield analysis and optimization tools for integrated circuits, and served as the Chief Technical Officer until the company was acquired by Extreme DA in 2007. His research interests include modeling, simulation, and synthesis for analog/RF and digital systems.

Dr. Li served on the IEEE Outstanding Young Author Award Selection Committee in 2006. He received the Best Session Award from the Semiconductor Research Corporation Student Symposium in 2006, the Best Paper Nomination in the 43rd IEEE/ACM Design Automatic Conference (DAC 2006), and the IEEE/ACM William J. McCalla Best Paper Award in the 2004 IEEE/ACM International Conference on Computer Aided Design (ICCAD). He also received the Inventor Recognition Awards from the Microelectronics Advanced Research Corporation in 2006 and 2007.



Yaping Zhan (S'05–M'06) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1999 and 2002, respectively, and the Ph.D. degree in electrical engineering from Carnegie Mellon University, Pittsburgh, PA, in 2005.

He is currently a Senior CAD Design Engineer with the Advanced Micro Devices, Austin, TX. His research interests include gate- and transistor-level timing analysis, statistical timing and power optimizations, gate and interconnect modeling, noise analysis, and high-performance low-power microprocessors.

Dr. Zhan served on the Technical Program Committee of the IEEE/ACM International Conference on Computer Aided Design in 2006 and 2007, respectively.

Lawrence T. Pileggi (S'85–M'89–SM'94–F'01) received the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 1989.

He is the Tanoto Professor of Electrical and Computer Engineering and the Director of the Center for Silicon System Implementation, Carnegie Mellon University. From 1984 to 1986, he was with Westinghouse Research and Development, where he received the corporation's highest engineering achievement award in 1986. From 1989 to 1995, he was a Faculty Member with the University of Texas at Austin. In January 1996, he joined the faculty of the Department of Electrical and Computer Engineering, Carnegie Mellon University. He has consulted for several electronic design automation (EDA) and semiconductor companies. He is a coauthor of *Electronic Circuit and System Simulation Methods* (McGraw-Hill, 1995) and *IC Interconnect Analysis* (Kluwer, 2002). He has published more than 200 refereed conference proceedings and journal papers and is the holder of 14 U.S. patents, including that for the RICE simulation tool. His research interests include various aspects of digital and analog design and EDA.

Prof. Pileggi is the Technical Program Chairman of the 2001 IEEE/ACM International Conference on Computer Aided Design (ICCAD) and the Conference Chairman of ICCAD 2002. He received the Best CAD Transactions Paper Awards in 1991 and 1999, the Best Paper Award in the 40th IEEE/ACM Design Automatic Conference (DAC 2003), and the Best Paper Award in ICCAD 2004. He received the Presidential Young Investigator Award from the U.S. National Science Foundation in 1991 and the Semiconductor Research Corporation Technical Excellence Awards in 1991 and 1999, the Invention Award from the Semiconductor Research Corporation in 1993, the University of Texas Parents' Association Centennial Teaching Fellowship in 1994 for his excellence in undergraduate instruction, and the Faculty Partnership Awards from IBM in 1995 and 2005.