

# Efficient Parametric Yield Extraction for Multiple Correlated Non-Normal Performance Distributions of Analog/RF Circuits

Xin Li and Lawrence T. Pileggi  
 Department of ECE, Carnegie Mellon University  
 5000 Forbes Avenue, Pittsburgh, PA 15213, USA  
 {xinli, pileggi}@ece.cmu.edu

## ABSTRACT

In this paper we propose an efficient numerical algorithm to estimate the parametric yield of analog/RF circuits with consideration of large-scale process variations. Unlike many traditional approaches that assume Normal performance distributions, the proposed approach is especially developed to handle multiple correlated non-Normal performance distributions, thereby providing better accuracy than other traditional techniques. Starting from a set of quadratic performance models, the proposed parametric yield extraction conceptually maps multiple correlated performance constraints to a single auxiliary constraint using a  $MAX(\bullet)$  operator. As such, the parametric yield is uniquely determined by the probability distribution of the auxiliary constraint and, therefore, can be easily computed. In addition, a novel second-order statistical Taylor expansion is proposed for an analytical  $MAX(\bullet)$  approximation, facilitating fast yield estimation. Our numerical examples in a commercial BiCMOS process demonstrate that the proposed algorithm provides 2~3x error reduction compared with a Normal-distribution-based method, while achieving orders of magnitude more efficiency than the Monte Carlo analysis with  $10^4$  samples.

## Categories and Subject Descriptors

B.7.2 [Integrated Circuits]: Design Aids – Verification

## General Terms

Algorithms

## Keywords

Parametric Yield, Analog/RF Circuits

## 1. INTRODUCTION

The continuous scaling of IC feature size has resulted in relatively large process variations [1]. For analog/RF circuits designed in sub-90nm 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 with consideration of large-scale process variations becomes an important task in today's robust analog/RF design flow.

During the past decade, various techniques have been proposed for statistical analysis of analog/RF circuits [2]-[8]. Most of these works [2]-[7] apply response surface modeling to approximate the performances of interest (e.g., gain, bandwidth, etc.) as polynomial (e.g., linear or quadratic) functions of process

parameters (e.g.,  $V_{TH}$ ,  $T_{OX}$ , etc.). These response surface models are then utilized to estimate the parametric yield of a given circuit design.

Many previous response surface modeling techniques (e.g., [3]-[4]) 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, especially because many analog/RF performances can be strongly nonlinear in the presence of large-scale variations. As will be demonstrated in Section 4, a 10% absolute error is observed for parametric yield estimation by using the linear approximation for a commercial 0.25 $\mu$ m BiCMOS process. It should be noted that such an error is expected to dramatically increase as the feature size further scales to 65nm and below.

To achieve higher accuracy, quadratic response surface modeling [5]-[7] can be used, but it makes parametric yield estimation much more difficult. While linear response surface models easily map a set of performance constraints to a well-defined polyhedron (called *feasible space*) in the process parameter space, such a mapping becomes nonlinear for quadratic models. In general, when quadratic response surface modeling is applied, the feasible space can be non-convex or even non-continuous. Therefore, the parametric yield, which equals the integral of the probability density function over the feasible space, becomes much more difficult to compute [8].

The authors of [5]-[6] apply Monte Carlo analysis to estimate the parametric yield defined by a set of quadratic 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, the recent advances in analog performance macromodeling and statistical circuit optimization suggest a need to revisit this assumption.

Firstly, today's analog modeling techniques make it possible to accurately extract unified performance macromodels over a large design space [9]-[10]. These performance macromodels are extracted only once and then used repeatedly for a given circuit topology. As such, the sampling points required by response surface modeling can be created using the pre-extracted performance macromodels rather than expensive transistor-level simulations, resulting in significant reduction of the response surface modeling cost.

Most importantly, for statistical circuit optimization, response surface models are extracted only once in a local design space and then used to estimate the parametric yield values at different design points [4], [7]. In this case, the yield estimation takes a large portion of the total computational cost, since it is repeatedly performed inside the optimization loop. For example, today's commercial analog sizing tool requires visiting up to  $10^5$  design points for one synthesis run [9]. In this case, even if one Monte Carlo analysis only takes 5 seconds, it will take 5.8 *days* to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2007, June 4-8, 2007, San Diego, California, USA.

Copyright 2007 ACM 978-1-59593-627-1/07/0006 ...\$5.00.

evaluate the parametric yield for all  $10^5$  design points!

The authors of [11] propose an asymptotic probability extraction algorithm (APEX) 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, while the parametric yield values of most analog/RF circuits are defined by multiple performance constraints. The challenging problem here is how to simultaneously consider *multiple* and *correlated* performance metrics whose probability distributions are all non-Normal.

In this paper we propose a novel algorithm for efficient parametric yield estimation of multiple correlated non-Normal performance distributions. The proposed algorithm conceptually maps multiple performance constraints to a single *auxiliary constraint* using a  $MAX(\bullet)$  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 and, therefore, can be easily estimated using the APEX algorithm [11] that is based on high-order moment matching.

An important contribution of this paper is to propose a *second-order statistical Taylor expansion* for the  $MAX(\bullet)$  approximation. The proposed technique is derived from the statistical sensitivity concept proposed for timing analysis [12] and is especially tuned for analog/RF applications. Unlike most statistical timing analysis algorithms where the  $MAX(\bullet)$  operator is approximated linearly [13]-[16], we approximate  $MAX(\bullet)$  as a quadratic function to capture the nonlinearities that are observed in most analog/RF performance variations. In addition, the proposed  $MAX(\bullet)$  approximation has an analytical formulation and, therefore, is much more efficient than Monte Carlo analysis. As will be demonstrated by numerical examples in Section 4, the proposed yield estimation algorithm can achieve up to 200x speed-up over the Monte Carlo analysis with  $10^4$  random samples.

The remainder of this paper is organized as follows. In Section 2 we review the background materials and then propose our parametric yield estimation algorithm in Section 3. The efficacy of the proposed algorithm is demonstrated by numerical examples in Section 4. Finally, we conclude in Section 5.

## 2. BACKGROUND

### 2.1 Response Surface Modeling

Given a circuit topology, the circuit performance (e.g. gain, bandwidth, etc.) is a function of both design parameters (e.g. bias current, transistor sizes, etc.) and process parameters (e.g.,  $V_{TH}$ ,  $T_{OX}$ , etc.). The design parameters are optimized and fixed during the design phase; however, the process parameters must be modeled as random variables to account for any uncertain manufacturing fluctuations. Given a set of fixed design parameters, the circuit performance  $f$  can be approximated by a linear response surface model [3]-[4]:

$$f(\boldsymbol{\varepsilon}) = B^T \cdot \boldsymbol{\varepsilon} + C \quad (1)$$

where  $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N]^T$  denotes  $N$  random variables to model process variations, and  $B \in R^N$  and  $C \in R$  are the model coefficients. The process variations in (1), i.e.  $\boldsymbol{\varepsilon}$ , are often modeled as zero-mean Normal distributions.

The linear response surface model in (1) is accurate when process variations are small. However, the large-scale variations that are expected for nanoscale technologies suggest that applying quadratic response surface models is required to provide sufficient accuracy [5]-[7]:

$$f(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}^T \cdot A \cdot \boldsymbol{\varepsilon} + B^T \cdot \boldsymbol{\varepsilon} + C \quad (2)$$

where  $A \in R^{N \times N}$ ,  $B \in R^N$  and  $C \in R$  are the model coefficients.

### 2.2 Linear $MAX(\bullet)$ Approximation

The linear  $MAX(\bullet)$  approximation has been extensively studied for statistical timing analysis [13]-[16]. Given two random variables  $x$  and  $y$ , the linear  $MAX(\bullet)$  approximation attempts to approximate  $MAX(x, y)$  as a linear function of  $x$  and  $y$ , i.e.:

$$MAX(x, y) = \alpha \cdot x + \beta \cdot y + \gamma \quad (3)$$

where  $\alpha$  and  $\beta$  are determined by the following probabilities:

$$\begin{aligned} \alpha &= PROB(x \geq y) \\ \beta &= PROB(y \geq x) \end{aligned} \quad (4)$$

and the constant term  $\gamma$  is determined by matching the mean value:

$$\gamma = E[ MAX(x, y) ] - \alpha \cdot E[x] - \beta \cdot E[y]. \quad (5)$$

In (4)-(5),  $PROB(\bullet)$  represents the probability and  $E(\bullet)$  stands for the expected value.

The authors in [12] further prove that the probabilities in (4) are equal to the first-order statistical sensitivities:

$$\begin{aligned} \alpha &= PROB(x \geq y) = \frac{\partial \{E[ MAX(x, y) ]\}}{\partial \{E[x]\}} \\ \beta &= PROB(y \geq x) = \frac{\partial \{E[ MAX(x, y) ]\}}{\partial \{E[y]\}} \end{aligned} \quad (6)$$

Although the  $MAX(\bullet)$  operator is not analytical (i.e., does not have continuous derivatives), it can be statistically approximated as the form of Eqn. (3)-(6) that is similar to the traditional Taylor expansion. We refer to (3) as the first-order *statistical Taylor expansion* in this paper.

The first-order statistical Taylor expansion has been widely applied to statistical timing analysis where digital delay variations can be accurately captured by linear approximation. Most analog/RF performance variations, however, are strongly nonlinear in the presence of large-scale process variations and cannot be accurately approximated by such linear models. This, in turn, motivates us to propose a novel second-order (i.e., quadratic) statistical Taylor expansion in this paper for analog/RF applications.

## 3. PARAMETRIC YIELD EXTRACTION

The proposed parametric yield extraction is facilitated by two key techniques, including (1) an auxiliary constraint formulation using a  $MAX(\bullet)$  operator; and (2) a second-order statistical Taylor expansion for the  $MAX(\bullet)$  approximation. In this section we describe the details of these algorithms and highlight their novelty.

### 3.1 Parametric Yield Formulation

The first step of our parametric yield extraction is to approximate all analog/RF performance metrics as quadratic response surface models. Such response surface modeling requires a number of sampling points and then solves the unknown model coefficients from a list of linear equations [5]-[7]. Based on these extracted response surface models, all analog/RF performance constraints can be expressed as the following standard form:

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

where  $f_k(\boldsymbol{\varepsilon})$  represents the response surface model of the  $k$ -th performance metric, and  $K$  is the total number of performance constraints. The standard form in (7) is ready to handle 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.

The *parametric yield* is equal to the probability that all performance constraints are satisfied, i.e.:

$$Yield = PROB[f_1(\varepsilon) \leq 0 \ \& \ f_2(\varepsilon) \leq 0 \ \& \ \dots \ f_k(\varepsilon) \leq 0]. \quad (8)$$

We define an additional auxiliary performance constraint:

$$f_{aux}(\varepsilon) = MAX[f_1(\varepsilon), f_2(\varepsilon), \dots, f_k(\varepsilon)]. \quad (9)$$

Eqn. (9) maps multiple performance metrics to a single auxiliary performance metric using a  $MAX(\bullet)$  operator. It is straightforward to verify that the parametric yield defined in (8) can be uniquely determined by the auxiliary constraint, i.e.:

$$Yield = PROB[f_{aux}(\varepsilon) \leq 0]. \quad (10)$$

If the auxiliary performance  $f_{aux}$  is approximated as a quadratic function of  $\varepsilon$ , the probability distribution of  $f_{aux}(\varepsilon)$  and, consequently, the parametric yield can be easily estimated by using the APEX algorithm proposed in [11]. Next, we propose a novel second-order statistical Taylor expansion to efficiently approximate the  $MAX(\bullet)$  operator in (9).

### 3.2 Quadratic $MAX(\bullet)$ Approximation

In this sub-section we focus on the  $MAX(\bullet)$  approximation of two random variables, i.e.,  $MAX(x, y)$ , since a multi-variable operation can be easily broken down into multiple two-variable cases. Such a two-variable  $MAX(\bullet)$  operator can be further transformed to a single-variable one:

$$MAX(x, y) = x + MAX(0, z) \quad (11)$$

where:

$$z = y - x. \quad (12)$$

The remainder of this sub-section will derive the second-order statistical Taylor expansion for the single-variable operator  $MAX(0, z)$ .

Extending the statistical Taylor expansion in (3)-(6) to second order and expanding  $MAX(0, z)$  at the expansion point  $E[z]$  yield:

$$MAX(0, z) = 0.5 \cdot \lambda_2 \cdot \{z - E[z]\}^2 + \lambda_1 \cdot \{z - E[z]\} + \lambda_0 \quad (13)$$

where the linear and quadratic coefficients  $\lambda_1$  and  $\lambda_2$  are determined by the statistical derivatives:

$$\lambda_1 = \frac{d\{E[ MAX(0, z) ]\}}{d\{E[z]\}} \quad (14)$$

$$\lambda_2 = \frac{d^2\{E[ MAX(0, z) ]\}}{d\{E[z]\}^2} = \frac{d\lambda_1}{d\{E[z]\}} \quad (15)$$

and the constant term  $\lambda_0$  is determined by matching the mean value:

$$\lambda_0 = E[ MAX(0, z) ] - 0.5 \cdot \lambda_2 \cdot E[\{z - E[z]\}^2]. \quad (16)$$

Next, we show how to compute the coefficients  $\lambda_0$ ,  $\lambda_1$  and  $\lambda_2$  in (14)-(16) efficiently.

#### A. Linear Coefficient $\lambda_1$

As summarized in Section 2.2, the first-order derivative in (14) is equal to the probability:

$$\lambda_1 = \frac{d\{E[ MAX(0, z) ]\}}{d\{E[z]\}} = PROB(z \geq 0) = 1 - CDF_z(0) \quad (17)$$

where  $CDF_z(\bullet)$  stands for the cumulative distribution function of the random variable  $z$ . Since  $z$  is equal to  $y - x$  in (12), where  $x$  and  $y$  are two performance metrics approximated as quadratic models,  $z$  is also a quadratic function of the random variations  $\varepsilon$ , i.e.:

$$z(\varepsilon) = \varepsilon^T \cdot A_z \cdot \varepsilon + B_z^T \cdot \varepsilon + C_z \quad (18)$$

where  $A_z$ ,  $B_z$  and  $C_z$  are the model coefficients. Given (18), the cumulative distribution function of  $z$  can be extracted using the APEX algorithm [11].

#### B. Quadratic Coefficient $\lambda_2$

Substituting (17) into (15) yields:

$$\lambda_2 = \frac{d\{1 - CDF_z(0)\}}{d\{E[z]\}}. \quad (19)$$

To calculate the derivative value in (19), we re-write  $z$  as:

$$z = \mu + \delta \quad (20)$$

where  $\mu$  is the mean value of  $z$  and  $\delta = z - \mu$  is a random variable with zero mean. Substituting (20) into (19) yields:

$$\begin{aligned} \lambda_2 &= \frac{d\{1 - CDF_{\mu+\delta}(0)\}}{d\mu} = \frac{d\{1 - CDF_{\delta}(-\mu)\}}{d\mu} \\ &= PDF_{\delta}(-\mu) = PDF_{\mu+\delta}(0) = PDF_z(0) \end{aligned} \quad (21)$$

where  $PDF_z(\bullet)$  stands for the probability density function of the random variable  $z$ . Again, the probability distribution of  $z$  can be extracted by APEX [11] to calculate  $\lambda_2$  in (21).

Importantly, the quadratic coefficient  $\lambda_2$  in (21) has two interesting properties:

- $\lambda_2 = PDF_z(0)$  is *non-negative*. Intuitively, as shown in Fig. 1, the function  $MAX(0, z)$  is convex and, therefore, the quadratic model coefficient should be non-negative [17].

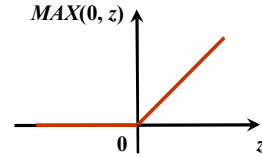


Fig. 1. The single variable function  $MAX(0, z)$ .

- $\lambda_2$  indicates the *nonlinearity*. Considering the first two cases in Fig. 2,  $MAX(0, z)$  can be accurately approximated as 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 of Fig. 2, however,  $MAX(0, z)$  is strongly nonlinear, corresponding to a non-zero  $PDF_z(0)$ .

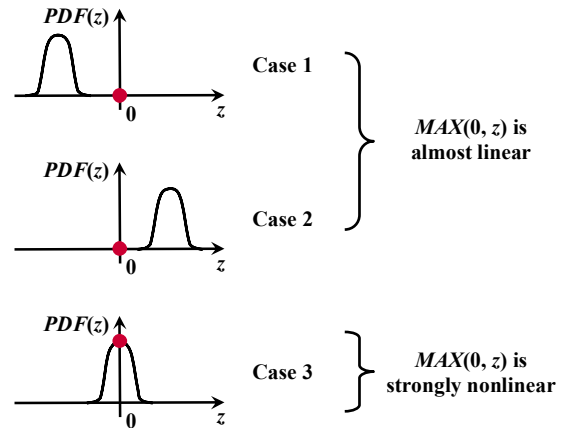


Fig. 2. Three different cases for the  $MAX(0, z)$  approximation.

#### C. Constant Term $\lambda_0$

After  $\lambda_1$  and  $\lambda_2$  are extracted, computing the constant term  $\lambda_0$

in (16) requires further knowing  $E[\text{MAX}(0, z)]$  and  $E[\{z - E[z]\}^2]$ .  $E[\text{MAX}(0, z)]$  can be calculated using the following one-dimensional numerical integration:

$$E[\text{MAX}(0, z)] = \int_0^{+\infty} \tau \cdot \text{PDF}_z(\tau) \cdot d\tau \quad (22)$$

Since  $z$  is a quadratic function of  $\varepsilon$  shown in (18), its second-order central moment can be determined by the following analytical equation [16]:

$$E[\{z - [z]\}^2] = B_z^T \cdot \Sigma \cdot B_z + 2 \cdot \text{TRACE}(\Sigma \cdot A_z \cdot \Sigma \cdot A_z) \quad (23)$$

where  $\text{TRACE}(\bullet)$  represents the trace of a matrix (the sum of all diagonal elements), and  $\Sigma$  stands for the correlation matrix of the random process variations  $\varepsilon$ . Substituting (21) and (22)-(23) into (16) yields the constant term  $\lambda_0$ .

After the coefficients  $\lambda_0$ ,  $\lambda_1$  and  $\lambda_2$  are known,  $\text{MAX}(0, z)$  in (13) can be approximated as a quadratic function of the random variations  $\varepsilon$  by substituting (18) into (13) and ignoring all high-order terms.

### 3.3 Summary

1. Start from a set of quadratic response surface models  $\{f_k(\varepsilon), k = 1, 2, \dots, K\}$ .
2. Represent all analog/RF performance constraints as the standard form in (7).
3. Set  $f_{aux}(\varepsilon) = f_1(\varepsilon)$ .  
For  $k = 2, 3, \dots, K$
4. Represent  $z(\varepsilon) = f_{aux}(\varepsilon) - f_k(\varepsilon)$  as the form of (18).
5. Compute the random distribution of  $z$  using APEX [11].
6. Compute the coefficients  $\lambda_0$ ,  $\lambda_1$  and  $\lambda_2$  using (14)-(23).
7. Compute  $\text{MAX}(0, z)$  by substituting (18) into (13).
8. Compute  $f_{aux}(\varepsilon) = f_k(\varepsilon) + \text{MAX}(0, z)$ .
- End For
9. Compute the random distribution of  $f_{aux}(\varepsilon)$  using APEX [11].
10. Compute the parametric yield defined in (10).

Fig. 3. A simplified flow of the proposed parametric yield extraction.

A simplified flow of the proposed parametric yield extraction is summarized in Fig. 3. Our proposed algorithm only involves APEX [11], one-dimensional numerical integration in (22) and other simple analytical operations and, therefore, is computationally efficient. As will be demonstrated by the numerical examples in Section 4, the proposed parametric yield extraction can achieve up to 200x speed-up over the Monte Carlo analysis with  $10^4$  random samples.

## 4. NUMERICAL EXAMPLES

In this section, we demonstrate the efficacy of the proposed algorithm using several analog/RF circuits. All circuits are implemented in a commercial  $0.25\mu\text{m}$  BiCMOS process. The numerical experiments in this section are performed on a LINUX 2.8 GHz server.

### 4.1 A Simple Example

We first consider a simple example to approximate  $\text{MAX}(x, y)$  where  $x \sim N(0, 1/3)$  and  $y \sim N(0, 1)$  are independent Normal distributions. Fig. 4 shows the probability density functions of the random variables  $x$  and  $y$ . In this example,  $\text{MAX}(x, y)$  is strongly nonlinear, because the probability density functions of  $x$  and  $y$  are

significantly overlapped. It, in turn, allows us to test the efficacy of the proposed  $\text{MAX}(\bullet)$  approximation and compare it with other traditional techniques.

Three different approaches, namely, the linear approximation, the proposed second-order statistical Taylor expansion and the Monte Carlo analysis with  $10^4$  random samples, are applied to estimate the probability distribution of  $\text{MAX}(x, y)$ . Fig. 5 shows the probability density functions estimated by these techniques. In this example, the  $\text{MAX}(x, y)$  distribution is not symmetric due to the nonlinearity. The traditional linear approximation cannot capture such a non-zero skewness and, therefore, results in large approximation error, especially at both tails of the probability density function. The proposed quadratic approximation, however, accurately models the non-zero skewness by including the second-order terms.

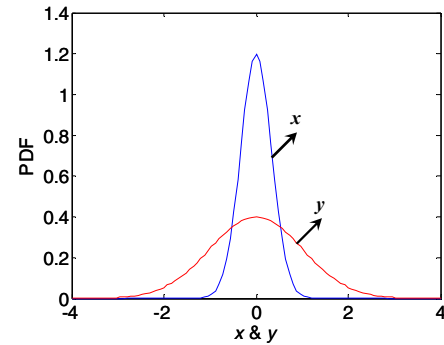


Fig. 4. The probability density functions of  $x$  and  $y$ .

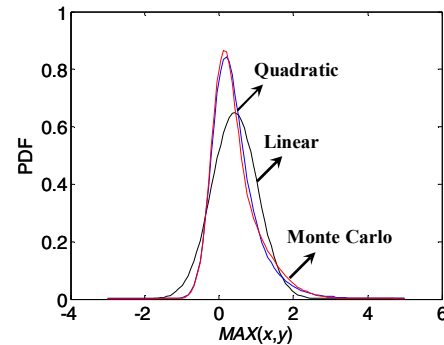


Fig. 5. The estimated probability density functions of  $\text{MAX}(x, y)$ .

## 4.2 Low Noise Amplifier

### A. Response Surface Modeling

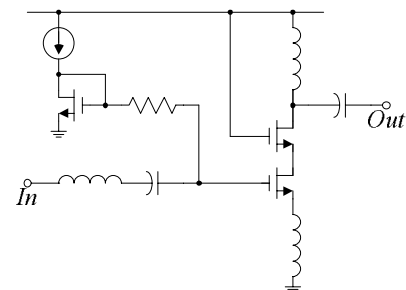


Fig. 6. The circuit schematic of a low noise amplifier.

Shown in Fig. 6 is a low noise amplifier designed in a

commercial BiCMOS 0.25 $\mu$ m process. In this example, the variations on both MOS transistors and passive components (i.e., resistors, capacitors and inductors) are considered. The probability distribution and correlation of these variations are provided in the process design kit. After PCA analysis, 8 principal factors are identified to model the critical process variations.

The performance of the LNA is characterized by 8 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. Table 1 shows the response surface modeling error for all these 8 performance metrics. In this example, the quadratic modeling error is 3.5x smaller than the linear modeling error on average.

Table 1. The response surface modeling error for LNA

Performance	Linear	Quadratic
F0	1.76%	0.32%
S11	6.40%	1.60%
S12	3.44%	0.80%
S21	2.94%	1.16%
S22	5.56%	3.67%
NF	2.38%	1.46%
IIP3	4.49%	1.64%
Power	3.79%	0.66%

It is worth mentioning that while the linear modeling error in this example is not extremely large, as IC technologies are scaled to 65nm and below, process variations will become relatively larger, thereby making the nonlinear terms in the quadratic model even more important.

### B. Parametric Yield Extraction

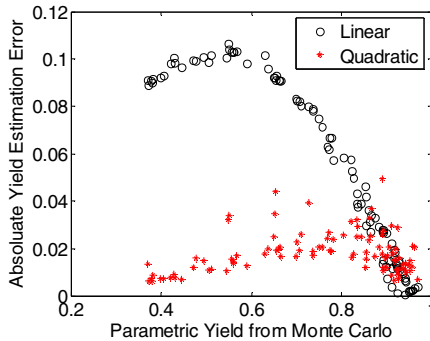


Fig. 7. The absolute yield estimation error for LNA.

Table 2. The yield estimation cost for LNA

Algorithm	Computational Time	
	One Design Point	One Synthesis Run (Visit $10^5$ Design Points)
Linear	0.002 Seconds	3.33 Minutes
Quadratic	0.021 Seconds	35.0 Minutes
Monte Carlo w/ $10^4$ Samples	4.537 Seconds	5.25 Days

For testing and comparison purpose, we randomly select 100 different design specifications and the parametric yield is estimated for each of these specifications. Fig. 7 compares the yield estimation accuracy for two different approaches: the traditional linear approximation and the proposed quadratic

approximation. The parametric yield values estimated by both techniques are compared with the Monte Carlo analysis results with  $10^4$  random samples. Their absolute difference is used as a measure of the estimation error for accuracy comparison.

As shown in Fig. 7, the traditional linear approximation cannot accurately capture the parametric yield of the LNA and the maximal absolute yield estimation error reaches 11%. The proposed quadratic approximation achieves much better accuracy and it reduces the maximal error to 5%. On average, the proposed quadratic approximation is 3x more accurate than the traditional linear approximation in this example.

Table 2 compares the computational cost of the parametric yield estimation for different approaches. The traditional linear approximation has the lowest computational cost; however, it cannot provide sufficient estimation accuracy. In this example, the proposed quadratic approximation achieves more than 200x speed-up over the Monte Carlo analysis with  $10^4$  random samples. As shown in Table 2, assuming that one complete synthesis run requires visiting  $10^5$  design points [9], the total yield estimation cost can be reduced from 5.25 days (by the Monte Carlo analysis with  $10^4$  random samples) to 35 minutes (by the proposed quadratic approximation)!

## 4.3 Operational Amplifier

### A. Response Surface Modeling

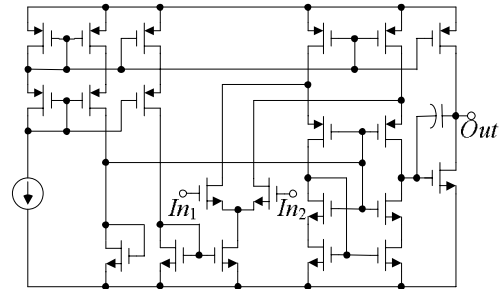


Fig. 8. The circuit schematic of an operational amplifier.

Table 3. The response surface modeling error for OpAmp

Performance	Linear	Quadratic
Gain	5.33%	0.98%
Offset	0.29%	0.07%
UGF	0.86%	0.17%
Gain Margin	0.65%	0.18%
Phase Margin	0.90%	0.15%
Slew Rate (+)	0.60%	0.44%
Slew Rate (-)	1.01%	0.25%
Power	0.48%	0.25%

Fig. 8 shows a two-stage folded-cascode operational amplifier designed in a commercial BiCMOS 0.25 $\mu$ m process. In this example, 49 principal factors are extracted by PCA to model the critical process variations, including both inter-die variations and device mismatches. It is important to note that modeling mismatches is extremely important for this operational amplifier, since they significantly impact several circuit performances, e.g., input offset voltage. The probability distribution and correlation of all random variations are provided in the process design kit.

The performance of the OpAmp is characterized by 8 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. Table 3 shows the response surface modeling error for all these 8 performance metrics. In this example, the quadratic modeling error is 4x smaller than the linear modeling error on average.

### B. Parametric Yield Extraction

Similar to the LNA example, we randomly select 100 different OpAmp design specifications and the parametric yield is estimated for each of them. Fig. 7 compares the yield estimation accuracy for two different approaches: the traditional linear approximation and the proposed quadratic approximation. The parametric yield values estimated by both techniques are compared with the Monte Carlo analysis results with  $10^4$  random samples. Their absolute difference is used as a measure of the estimation error for accuracy comparison.

As shown in Fig. 9, the traditional linear approximation results in a maximal error of 9%. The proposed quadratic approximation is much more accurate and it reduces the maximal error to 5%. On average, the proposed quadratic approximation is 2x more accurate than the traditional linear approximation in this example. In addition, the proposed quadratic approximation achieves more than 150x speed-up over the Monte Carlo analysis with  $10^4$  random samples, as shown in Table 4.

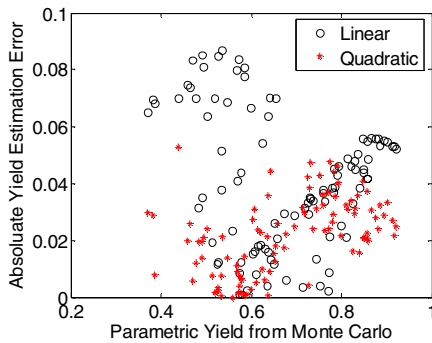


Fig. 9. The absolute yield estimation error for OpAmp.

Table 4. The yield estimation cost for OpAmp

Algorithm	Computational Time	
	One Design Point	One Synthesis Run (Visit $10^5$ Design Points)
Linear	0.003 Seconds	5.00 Minutes
Quadratic	0.031 Seconds	51.7 Minutes
Monte Carlo w/ $10^4$ Samples	5.343 Seconds	6.18 Days

## 5. CONCLUSIONS

We have proposed an efficient parametric yield extraction algorithm for multiple correlated non-Normal analog/RF performance distributions that are expected in nanoscale technologies. The proposed algorithm is facilitated by two key techniques, including (1) an auxiliary constraint formulation using a  $MAX(\bullet)$  operator; and (2) a second-order statistical Taylor expansion for the  $MAX(\bullet)$  approximation. Our numerical examples in a commercial  $0.25\mu\text{m}$  BiCMOS process demonstrate that, compared with the traditional Normal-distribution-based method, the proposed parametric yield extraction reduces the estimation error by 2~3x, while simultaneously achieving up to 200x speed-up over the Monte Carlo analysis with  $10^4$  samples.

In addition, the second-order statistical Taylor expansion

proposed in this paper is not limited to the parametric yield estimation for analog/RF circuits only. The same idea can be widely applied to many other applications that require the  $MAX(\bullet)$  operation, e.g., statistical timing analysis for digital circuits.

## 6. ACKNOWLEDGEMENTS

This work has been supported by the FCRP Focus Center for Circuit & System Solutions (C2S2) under contract 2003-CT-888, the Semiconductor Research Cooperation (SRC) under contract 2005-TJ-1298 and Mentor Graphics Inc.

## 7. REFERENCES

- [1] Semiconductor Industry Associate, *International Technology Roadmap for Semiconductors*, 2004.
- [2] M. Styblinski and J. Zhang, "Circuit performance variability reduction: principles, problems, and practical solutions," *IEEE ICCAD*, pp. 170-173, 1991.
- [3] Z. Wang and S. Director, "An efficient yield optimization method using a two step linear approximation of circuit performance," *IEEE EDAC*, pp. 567-571, 1994.
- [4] F. Schenkel, M. Pronath, S. Zizala, R. Schwencker, H. Graeb and K. Antreich, "Mismatch analysis and direct yield optimization by spec-wise linearization and feasibility-guided search," *IEEE DAC*, pp. 858-863, 2001.
- [5] A. Dharchoudhury and S. Kang, "Worse-case analysis and optimization of VLSI circuit performance," *IEEE Trans. CAD*, Vol. 14, No. 4, pp. 481-492, Apr. 1995.
- [6] E. Felt, S. Zanella, C. Guardiani and A. Sangiovanni-Vincentelli, "Hierarchical statistical characterization of mixed-signal circuits using behavioral modeling," *IEEE ICCAD*, pp. 374-380, 1996.
- [7] X. Li, *Statistical Modeling, Analysis and Optimization for Analog and RF ICs*, PhD Dissertation, Carnegie Mellon University, 2005.
- [8] P. Feldmann and S. Director, "Integrated circuit quality optimization using surface integrals," *IEEE Trans. CAD*, Vol. 12, No. 12, pp. 1868-1879, 1993.
- [9] H. Liu, A. Singhee, R. Rutenbar, L. Carley, "Remembrance of circuit past: macromodeling by data mining in large analog design spaces," *IEEE DAC*, pp. 437-442, 2002.
- [10] F. Bernardinis, M. Jordan and A. Sangiovanni-Vincentelli, "Support vector machine for analog circuit performance representation," *IEEE DAC*, pp. 964-969, 2003.
- [11] X. Li, J. Le, P. Gopalakrishnan and L. Pileggi, "Asymptotic probability extraction for non-Normal distributions of circuit performance," *IEEE ICCAD*, pp. 2-9, 2004.
- [12] 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," *IEEE ICCAD*, pp. 844-851, 2005.
- [13] H. Chang and S. Sapatnekar, "Statistical timing analysis under spatial correlations," *IEEE Trans. CAD*, Vol. 24, No. 9, pp. 1467-1482, 2005.
- [14] 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. CAD*, Vol. 25, No. 10, pp. 2170-2180, 2006.
- [15] H. Chang, V. Zolotov, S. Narayan and C. Visweswariah, "Parameterized block-based statistical timing analysis with non-Gaussian parameters, nonlinear delay functions," *IEEE DAC*, pp. 71-76, 2005.
- [16] L. Zhang, W. Chen, Y. Hu, J. Gubner and C. Chen, "Correlation-preserved non-Gaussian statistical timing analysis with quadratic timing model," *IEEE DAC*, pp. 83-88, 2005.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.