CMCal: An Accurate Analytical Approach for the Analysis of Process Variations with Non-Gaussian Parameters and Nonlinear Functions

M. Zhang$^1$, M. Olbrich$^1$, D. Seider$^2$, M. Frerichs$^2$, H. Kinzelbach$^3$, E. Barke$^1$

$^1$Institute of Microelectronic Systems
Leibniz University of Hannover, Germany
{zhang, olbrich, barke}@ims.uni-hannover.de

$^2$Qimonda AG
{David.Seider, Martin.Frerichs}@qimonda.com

$^3$Infineon Technologies AG
{Harald.Kinzelbach}@infineon.com

Abstract — As technology rapidly scales, performance variations (delay, power etc.) arising from process variation are becoming a significant problem. The use of linear models has been proven to be very critical in many today’s applications. Even for well-behaved performance functions, linearising approaches as well as quadratic model provide serious errors in calculating expected value, variance and higher central moments. In this paper, we present a novel approach to analyse the impacts of process variations with low efforts and minimum assumption. We formulate circuit performance as a function of the random parameters and approximate it by Taylor Expansion up to 4th order. Taking advantage of the knowledge about higher moments, we convert the Taylor series to characteristics of performance distribution. Our experiments show that this approach provides extremely exact results even in strongly non-linear problems with large process variations. Its simpleness, efficiency and accuracy make this approach a promising alternative to the Monte Carlo Method in most practical applications.

I. INTRODUCTION

With rapidly shrinking technology process variations are becoming increasingly important in digital, analog and mixed-signal designs. While circuit performances mostly vary within a range of 10% in 180 nm technology, the fluctuations at the 90 nm node and below increase up to 50% [1]-[4]. The existing approaches exploring impacts of manufacturing variations can be classified into two main categories: stochastical and analytical.

The basic Monte Carlo Method (MCM) - the most known stochastical method - is generally accurate, easy to apply, but very time consuming. Employing different sampling techniques, the basic MCM can be significantly accelerated [5]. Recently, several methods using Importance Sampling (IS) technique have been proposed to estimate the stochastical characteristics [6] as well as to analyse rare failure events [7]. However, the stochastical nature of these methods always requires a minimum sampling size, which might be critical in low dimensional applications (small number of varying parameters).

Results from Corner Case Analysis are often too pessimistic since the probability of the occurrence of the worst case is quite small. Moreover, the corners of parameters do not always imply the corners of performance [8], [9].

The linearising method is probably the most common analytical method. It assumes a linear dependency of performance on process variations for the sake of simpleness of calculation [10]-[13]. From the stochastical point of view it is merely a linear transformation of distributions and holds predefined limitations: e.g. in case of Gaussian distributed parameters the result of this approach is a Gaussian distribution about the nominal value, which obviously does not mirror the true performance distribution. Furthermore, such approaches are logically not suitable for strongly non-linear cases (more in Section IV).

Recent study suggests the use of a quadratic model in Statistical Static Timing Analysis (SSTA), which appears to achieve 10x accuracy improvement [14] [15]. However, our experiments show that the results are still not sufficiently precise in many applications (see Section IV).

Further approaches combining corner case analysis and linearising method have been intensively investigated. The geometrical approaches, e.g., are applied for performance analysis and design centering in analog circuits [16][17].

In this paper, we propose a novel analytical approach based on Taylor Expansion and Central Moment Calculation. This approach aims to explore impacts of process variations with very low effort and to calculate stochastical characteristics of the circuit performance under a minimum number of assumptions. Our experiments show that this approach provides extremely exact results even in strongly non-linear cases and can be generally used as an alternative to MCM. The main features of this approach are (see Fig. 1):

- Similar to MCM, existing simulators will be used as black box in this approach. The performance T(x), which is actually determined by various complex numerical procedures (which we call “simulator” in this paper), will
be formulated as a function of the varying parameters.

- Our approach works with Central Moments (CM) instead of Probability Density Functions (PDF). Since the CMs of most distributions are known, there are little assumption about the distribution on the input side. Using CMs on the output side, different simulators are easily to be cascaded.
- The distribution of outputs will also be described by the CMs. The analysis of CMs is extremely useful for design optimization, the discussion of which is however not the subject of this paper.

![Diagram of CMCal Approach](image)

**Fig. 1.** Black Box Approach: Simulator is used as a black box in CMCal approach. The input distribution and output distribution are characterised by central moments. The remainder of this paper is organized as follows: Section II reviews the underlying mathematics; Section III describes our approach in detail; The experimental results will be presented and discussed in Section IV; Finally, conclusions with ongoing research directions will be drawn in Section V.

**II. BACKGROUND**

As mentioned in Section. I the existing simulator is embedded into our Central-Moment-Calculating (CMCal) approach (see also Fig.1). Instead of random numbers, the inputs of our approach are the CMs. Before we introduce the use of CMs, we will firstly review some definitions in probability theory.

*Expected Value (Expection)* is one of the most fundamental term in statistics:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx,$$

where $X$ is a random variable and $f(x)$ is its probability density function. $E[X]$ is also denoted as *Mean* of $X$ ($\mu = E[X]$) and indicates where the values of the random variable $X$ are “centred”.

Another important notion is *Moment (Raw Moment)*. The $r$th Moment of $X$ is defined as the expected value of $X^r$:

$$m_r = E[X^r].$$

The $r$th Moment about the *Mean* is defined as the $r$th *Central Moment*:

$$M_r = E[(X - \mu)^r].$$

Note that a measurable function of a random variable is also a random variable [18]. Therefore, these definitions are valid not only for the parameter distributions on the input side of simulator but also for the performance distribution on the output side (see Fig.1). The *Expected Value* of $T(x)$, for example, is then defined as:

$$\mu_T = E[T(X)] = \int_{-\infty}^{\infty} T(x)f(x)dx.$$  

*Expected Value* has several properties [18] which will be frequently used in our approach: Let $c_1$ and $c_2$ be constants,

- $E[c_1] = c_1$;
- $E[c_1T(X)] = c_1E[T(X)]$;
- $E[c_1T_1(X) + c_2T_2(X)] = c_1E[T_1(X)] + c_2E[T_2(X)]$.

While the entire set of moments ordinarily determine the distribution exactly, the values of *Mean*, *Variance*, *Skewness* and *Kurtosis* give most important information about a distribution in applied statistics [18].

*Variance* is defined as the second *Central Moment*. The square root of Variance, the *standard deviation* $\sigma$, is used to indicate the dispersion of distribution.

$$Var(X) = M_2 = \int_{-\infty}^{\infty} (X - \mu_x)^2 f(x)dx = E[(X - \mu_x)^2]$$  

(1)

In practice the variance is often calculated by:

$$Var(X) = E[X^2] - (E[X])^2.$$  

(2)

*Skewness* is a measure of symmetry and is defined as:

$$Skew(X) = \frac{M_3}{Var(X)^{\frac{3}{2}}}.$$  

(3)

As Fig.2 shows, a curve is skewed to the left (right) if it has a positive (negative) *Skewness*.

![pdf](image)

*Fig. 2. Skewness determines the symmetry of distribution: (a) Positive Skewness (b) Negative Skewness*

*Kurtosis* is used to indicate the degree of flatness of a density near its center:

$$Kurt(X) = \frac{M_4}{Var(X)^2} - 3.$$  

(4)

A positive kurtosis means that the density is more peaked around its center than the density of a normal distribution (see Fig.3).
Given a simulator approach is based on two considerations/facts:

For normal distribution, e.g. \( \zeta \), where the domain around nominal points

\[ \mathbf{x}_d \] indicates that the performance should be a well-behaved function of process parameters, at least within a small interval around the nominal parameters. Given a simulator \( T(\bar{x}) \) and varying inputs of \( T(\bar{x}) \), \( \bar{x} = (x_1, x_2, \ldots, x_d) \). We assume a convergent power series in a small domain around nominal points \( \mathbf{\bar{\mu}} = (\mu_1, \mu_2, \ldots, \mu_d) \), so that we have the infinite Taylor series expansion:

\[
T(\bar{x}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_d=0}^{\infty} C_{n_1 \cdots n_d} \zeta_1^{n_1} \cdots \zeta_d^{n_d},
\]

where \( \zeta_i \) denotes \( (x_i - \mu_i) \) and the Taylor series coefficients, \( C_{n_1 \cdots n_d} \), equal:

\[
C_{n_1 \cdots n_d} = \frac{1}{n_1! \cdots n_d!} \left( \frac{\partial}{\partial x_1} \right)^{n_1} \left( \frac{\partial}{\partial x_2} \right)^{n_2} \cdots \left( \frac{\partial}{\partial x_d} \right)^{n_d} T(\mathbf{\bar{\mu}}).
\]

To explain the concept conveniently, our approach will be demonstrated for a one-dimension (one varying parameter) problem in the remainder of this paper:

\[
T(x)|_{x=\mu} = \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\partial^n T(\mu)}{\partial x^n} \zeta^n \\
\approx T_0 + T_1 + T_2 + T_3 + T_4.
\]

\( T_n \) is then defined as

\[
T_n = C_n(\mu) \zeta^n,
\]

where \( C_n(\mu) \) denotes the \( n \)th derivative of function \( T(x) \) at nominal point \( \mu \). Applying the properties of the expected value (as introduced in Section II) to the both sides of Eq.(5), we get the mean of \( T(x) \):

\[
E[T(x)] = E[T_0] + E[T_1] + E[T_2] + E[T_3] + E[T_4].
\]

Each summand in Eq.(7) can be evaluated easily by the same principle, e.g.:

\[
E[T_2] = \frac{C_2(\mu)}{2!} E[\zeta^2] = \frac{C_2(\mu)}{2!} M_2,
\]

where \( M_2 \) is the 2nd CM of input variable \( x \). Note that for all distributions \( M_1 = 0 \) and for all symmetrical distributions \( M_3 = 0 \). Therefore, assuming \( x \) is symmetrically distributed, the mean of \( T(X) \) equals:

\[
E[T(x)] = T(\mu) + \frac{C_1(\mu)}{1!} M_1 + \frac{C_2(\mu)}{2!} M_2 + \frac{C_3(\mu)}{3!} M_3 + \frac{C_4(\mu)}{4!} M_4
\]

\[
= T(\mu) + \frac{C_2(\mu)}{2} Var(x) + \frac{C_4(\mu)}{8} Var^2(x).\]

Using Eq.(2), the variance of \( T(x) \) is then calculated by:

\[
Var(T(x)) = E[T(x)^2] - (E[T(x)])^2.
\]

The only unknown term is \( E[T(x)^2] \), the calculation of which should be no problem by substituting \( T(x) \) with its Taylor expansion.

\[
T^2 = (T_0 + T_1 + T_2 + T_3 + T_4)^2
\]

\[
= T_0^2 + T_1^2 + T_2^2 + T_3^2 + T_4^2 + 2T_0 T_1 + 2T_0 T_2 + 2T_0 T_3 + 2T_0 T_4 + 2T_1 T_2 + 2T_1 T_3 + 2T_1 T_4 + 2T_2 T_3 + 2T_2 T_4 \]

\[
+ 2T_3 T_4.
\]

The generalized form of the summands in Eq.(10) is:

\[
T_0^{p_0} T_1^{p_1} T_2^{p_2} T_3^{p_3} T_4^{p_4}.
\]

For each term we can calculate its TOTAL-POWER \( I_{total} \):

\[
I_{total} = \sum_{i=0}^{4} i P_i.
\]

It is easy to see that the \( I_{total} \) indicates the power of \( \zeta \) in each summand. The \( I_{total} \) of 2T2T3, for example, is 5 and this term is relevant with \( (x-\mu)^5 \).

\[
2T_2 T_3 = \frac{C_2(\mu) C_3(\mu)}{6} \zeta^5 = \frac{C_2(\mu) C_3(\mu)}{6} (x-\mu)^5
\]

Since in our approach we only approximate performance functions up to 4th order, terms whose TOTAL-POWER is greater than 4 (like “2T2T3”) will be deleted.
Similar to Eq.(8), all terms with \( l = 1,3 \) can be deleted in case of symmetrically distributed process variations. Eq.(10) is then simplified to:

\[
T^2 = T_0^2 + T_1^2 + T_2^2 + 2T_0T_2 + 2T_0T_4 + 2T_1T_3 .
\]

(13)

Applying this in Eq.(9), we get

\[
\text{Var}[T(x)] = \text{Var}(x)[C_2^2(\mu)] + \text{Var}^2(x)[\frac{C_2^2(\mu)}{2} + C_1(\mu)C_3(\mu)]) + \text{Var}^3(x)[\frac{C_2^2(\mu)C_4(\mu)}{8}] + \text{Var}^4(x)[\frac{C_2^2(\mu)}{64}].
\]

(14)

To calculate the skewness and kurtosis of \( T(x) \), \( M_3[T] \) and \( M_4[T] \) are required. Similar to Eq.(9), they can be recursively evaluated by computing \( E[T^3] \) and \( E[T^4] \):

\[
\]

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed methodology was implemented on top of different simulators to test its “robustness”, i.e., to test its accuracy for strongly non-linear applications.

The first experiment is to use CMCal to investigate the impact of process variations on the interconnect characteristics of a two-dimensional model shown in Fig. 4. GEO2D is used to evaluate the resistance and capacitance matrix [20]. Resistance \( R \) and capacitance \( C \) of the mid-interconnect have been analysed. The results are compared with those of the basic Monte Carlo Method (MCM), Importance Sampling (IS), Linear Regression Model (LM) and Quadratic Model (QM).

Fig. 5 plots the dependency of \( R, C \) on the width of the mid-interconnect. The resistance shows a more non-linear behaviour with the shrinking of technology.

Fig. 6 shows the estimation error of linearising approach for Mean calculation with respect to the parameter variation (width of interconnect) at the 50 nm node. The error is defined as the relative difference between “real” value (Mean calculated by MCM with 36,000 Samples) and nominal value (Mean calculated by linear model). It clearly indicates that estimation errors arising from linearization are not acceptable: in the presence of process variation \( W_c : N(50nm, 15nm^2) \), the use of the linearising method to calculate mean value causes an error of 12.43%.

Moreover, Fig. 6 clearly indicates that the estimation error of linearization logically depends on two factors: linearity of the performance function and variation of process parameters.

- Since the R-curve is more non-linear than the C-curve (see Fig. 5), the estimation error of the resistance is larger than that of the capacitance.
- The estimation error decreases as process becomes stable (small value of \( \sigma \)).

This observation is totally in line with Eq.(8) and suggests the use of higher order models.
TABLE I
MEAN VALUE E[\text{R}]: COMPARISON OF BASIC MCM, IS, LM, QM, AND CMCal

<table>
<thead>
<tr>
<th>Method</th>
<th>MCM</th>
<th>MCM</th>
<th>IS</th>
<th>Linear Model</th>
<th>Quadratical Model</th>
<th>CMCal</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Simulation</td>
<td>3000</td>
<td>30000</td>
<td>1000</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Result(10^7)</td>
<td>2.3205</td>
<td>2.2812</td>
<td>2.2736</td>
<td>2.0000</td>
<td>2.1465</td>
<td>2.2813</td>
</tr>
<tr>
<td>Deviation to Ref.</td>
<td>1.61%</td>
<td>-0.11%</td>
<td>-0.44%</td>
<td>-12.43%</td>
<td>-6.01%</td>
<td>-0.11%</td>
</tr>
</tbody>
</table>

TABLE II
STANDARD DEVIATION \(\sigma[R] = \sqrt{\text{Var}[R]}\): COMPARISON OF BASIC MCM, IS, LM, QM AND CMCal

<table>
<thead>
<tr>
<th>Method</th>
<th>MCM</th>
<th>MCM</th>
<th>IS</th>
<th>Linear Model</th>
<th>Quadratical Model</th>
<th>CMCal</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Simulation</td>
<td>3000</td>
<td>30000</td>
<td>1000</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Result(10^7)</td>
<td>1.6531</td>
<td>1.4251</td>
<td>1.3246</td>
<td>0.3986</td>
<td>0.6644</td>
<td>1.3979</td>
</tr>
<tr>
<td>Deviation to Ref.</td>
<td>16.7%</td>
<td>0.61%</td>
<td>-6.48%</td>
<td>-71.77%</td>
<td>-53.09%</td>
<td>-0.97%</td>
</tr>
</tbody>
</table>

TABLE III
STANDARD DEVIATION \(\sigma[\text{delay}] = \sqrt{\text{Var}[\text{delay}]}\): COMPARISON OF BASIC MCM, LM, QM AND CMCal

<table>
<thead>
<tr>
<th>Method</th>
<th>MCM</th>
<th>MCM</th>
<th>Linear Model</th>
<th>Quadratical Model</th>
<th>CMCal</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Simulation</td>
<td>1000</td>
<td>5000</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Result(10^7)</td>
<td>7.2109</td>
<td>7.5691</td>
<td>2.4751</td>
<td>6.3384</td>
<td>7.4159</td>
</tr>
<tr>
<td>Deviation to Ref.</td>
<td>3.28%</td>
<td>1.52%</td>
<td>-66.80%</td>
<td>-14.98%</td>
<td>-0.53%</td>
</tr>
</tbody>
</table>

Applying the proposed CMCal method, the accuracy is significantly improved. Table I compares the results of different methods used to calculate the mean value in presence of one varying parameter. The estimation error has been reduced from 12.43% to 6.01% by using a quadratic model instead of a linear one. Applying CMCal, the result after 5 simulations is as precise as that of basic MCM after 30,000 simulations: i.e. the speedup is 6,000. Speedup is defined as \(\frac{N_{MCM}}{N_{CMCal}}\), where \(N_{MCM}\) and \(N_{CMCal}\) are the number of samples needed by the basic MCM and CMCal, respectively. It has to be noted that for multiple varying parameters more simulations will be needed. However, the large margin between 5 and 30,000 indicates that CMCal can be used as an alternative to basic MCM for low dimensional problems, even from a conservative point of view.

The most competitive advantage of our approach is to estimate the higher central moments of the circuit performance quite precisely. As Table II indicates, the results of a linear model and a quadratic model are useless for variance calculating: the estimation error for standard deviation, \(\sqrt{\text{Var}[R]}\), are \(-71.77\%\) and \(-53.09\%\). By contrast, our approach still yields an extreme high accurate level \((-0.97\%)\), for which 30,000 simulations will be needed by the basic MCM.

Fig. 7 also shows the errors of different approaches while they are applied to calculate Mean, Variance, Skewness and Kurtosis of resistance distribution. For this strongly non-linear problem, the linear and quadratic model yield error ratios over 100% for higher central moments.

The efficiency and accuracy of the CMCal approach has been re-confirmed by further applications.The objective was to monitor the change of delay variance arising from interconnect
variance in a H-Tree structure. Fig. 8 illustrates the dependency of delay on one interconnect width, while the delays are calculated by the HSPICE simulator. The analysis has been performed at the most “non-linear” position (as shown in Fig. 8) and the comparison has been shown in Table III. In this simple and common case, the estimation error arising from linearization is 66.8%! Even the second order approximation leads to 15% error, while the error of CMCal stays negligibly small (0.53%).

V. CONCLUSION

Exploiting the features of central moments as well as that of moments, we developed a novel approach to explore the impact of process variations with very low effort and minimum number of assumptions.

We approximate the performance function up to 4th order and use the higher central moments to convert knowledge of the approximating function to knowledge of the stochastic characteristics. The only assumption is that the CMs of the input parameter distribution should be known.

Our experiments show that this approach still provides extremely exact results to strongly non-linear problems with large variations. In low-dimensional problems, it yields a speedup up to 6,000. This indicates that it generally can be used as a promising alternative to Monte Carlo Method in low-dimensional applications. For high-dimensional applications it can be used after conventional dimensional reduction procedures [21].

Due to its analytical nature the CMCal approach is coupled with the number of dimensions. At present we are focusing on developing methods to reduce the number of needed simulations in high-dimensional applications.

REFERENCES