Structure-Aware High-Dimensional Performance Modeling for Analog and Mixed-Signal Circuits

Shupeng Sun[1], Xin Li[1] and Chenjie Gu[2]

[1]Electrical & Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA, USA, 15213
[2]Intel Strategic CAD Labs, Hillsboro, OR, USA, 97124

Abstract—**Efficient high-dimensional performance modeling of nanoscale analog and mixed signal (AMS) circuits is extremely challenging. In this paper, we propose a novel structure-aware modeling (SAM) technique. The key idea of SAM is to accurately solve the model coefficients by applying an efficient statistical algorithm to exploit the underlying structure of AMS circuits. As a result, SAM dramatically reduces the required number of sampling points and, hence, the computational cost for performance modeling. Several circuit examples designed in a commercial 32nm CMOS process demonstrate that SAM achieves more than 2× runtime speedup over the traditional sparse regression technique without surrendering any accuracy.**

I. INTRODUCTION

The aggressive technology scaling of integrated circuits (IC) leads to large-scale process variations [1], significantly impacting the parametric yield of analog and mixed-signal (AMS) circuits. As an important tool for variability analysis, performance modeling aims to approximate the circuit-level performance (e.g., frequency of ring oscillator) as an analytical (e.g., linear, quadratic, etc) function of device-level variations (e.g., ΔV_{th}, ΔL, etc) [2]-[3]. Once such performance models are available, they can be applied to several important applications such as yield estimation [4], worst-case corner extraction [5], design optimization [6], etc.

Although performance modeling was extensively studied in the past, several new technical challenges arise due to the recent evolution of nanoscale IC technologies. Today, a large number of random variables must be employed to accurately capture device-level variations. For example, more than 40 random variables are used to model the random mismatches for a single transistor in a commercial 32nm CMOS process. Even if we consider a small AMS circuit with 100 transistors only, there exist more than 4000 random variables for device-level mismatch modeling. In addition, it is almost impossible to pre-select a subset of these random variables for variability analysis, since the impact of device mismatches is circuit- and performance-dependent. It, in turn, requires us to fit a high-dimensional performance model based on a large number of sampling points that must be generated by expensive transistor-level simulations [7]. A major challenge here is how to minimize the required number of sampling points so that the modeling cost can be substantially reduced.

Recently, it has been found that although there are a large number of unknown coefficients associated with a high-dimensional performance model for a given performance of interest (PoI), most of these coefficients are close to zero [7]. By exploiting this unique property, sparse regression technique has been developed for high-dimensional performance modeling [7]. The key idea is to automatically identify the important (i.e., non-zero) model coefficients by applying a statistical algorithm. As such, a high-dimensional performance model can be accurately fitted with a small number of sampling points.

In this paper, we further improve the traditional sparse regression method and propose a novel *structure-aware modeling* (SAM) technique for AMS circuits. Our proposed method is motivated by the observation that many coefficients of a performance model often share similar magnitude. For instance, consider a ring oscillator that is composed of identically-sized inverters. In this example, the threshold mismatches (i.e., ΔV_{th}'s) of all NMOS (or PMOS) transistors have similar contribution to the performance variability due to its symmetric circuit topology. Therefore, the model coefficients associated with these ΔV_{th}'s should be similar.

Based on this observation, SAM attempts to group the "similar" model coefficients and solve them together. Towards this goal, a new algorithm, referred to as *simultaneous orthogonal matching pursuit* (S-OMP), is adopted from the statistics community [8] and applied to our high-dimensional performance modeling problem. In addition, a number of important heuristic methodologies are proposed in order to make S-OMP of great efficiency, as will be discussed in detail in Section II. Several circuit examples in Section III demonstrate that SAM achieves more than 2× runtime speedup over the traditional sparse regression technique without surrendering any accuracy.

The remainder of this paper is organized as follows. In Section II, we describe SAM and the S-OMP algorithm integrated with our proposed heuristics. The efficacy of SAM is demonstrated by several circuit examples in Section III. Finally, we conclude in Section IV.

II. METHODOLOGY

Without loss of generality, we assume that there are N device-level random variables $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \varepsilon_2 \ ... \ \varepsilon_N]^T$ modeling process variations. Given a PoI denoted as y, performance modeling aims to construct an analytical function to capture the relation between y and $\boldsymbol{\varepsilon}$. In this paper, we use the following linear performance model as an example to illustrate our proposed SAM methodology:

$$y = \alpha_0 + \sum_{n=1}^{N} \alpha_n \cdot \varepsilon_n = \boldsymbol{\alpha}^T \cdot \mathbf{x}, \qquad (1)$$

where $\boldsymbol{\alpha} = [\alpha_0 \ \alpha_1 \ ... \ \alpha_N]^T$ represents the model coefficients and $\mathbf{x} = [1 \ \varepsilon_1 \ \varepsilon_2 \ ... \ \varepsilon_N]^T$ contains the constant term 1 and the random variables $\boldsymbol{\varepsilon}$. It should be noted that SAM can be further

extended to other nonlinear performance models, even though the details of nonlinear performance modeling are not discussed in this paper due to the page limit.

The unknown model coefficients $\boldsymbol{\alpha}$ in (1) are often solved from a set of sampling points. In particular, if M sampling points are collected, the following linear equation is formulated to solve $\boldsymbol{\alpha}$:

$$\mathbf{X} \cdot \boldsymbol{\alpha} = \mathbf{y}, \qquad (2)$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & \varepsilon_1^{(1)} & \cdots & \varepsilon_N^{(1)} \\ 1 & \varepsilon_1^{(2)} & \cdots & \varepsilon_N^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \varepsilon_1^{(M)} & \cdots & \varepsilon_N^{(M)} \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(M)} \end{bmatrix}. \qquad (3)$$

In (3), $\varepsilon_n^{(m)}$ and $y^{(m)}$ denote the values of ε_n and y for the m-th sampling point, respectively.

For today's nanoscale AMS circuits, the total number of unknown model coefficients can be extremely large (e.g., $10^3 \sim 10^4$). However, most of these coefficients are close to zero, as is demonstrated in the literature [7]. In addition, a number of the model coefficients often share similar magnitude. For instance, consider the latch-based comparator in Fig. 1 as an example. In this compactor, each transistor contains 10 multipliers. If we consider the input offset as PoI, the model coefficients associated with the threshold mismatches (i.e., ΔV_{th}'s) of all multipliers of the two input transistors M_1 and M_2 should have similar magnitude, since these threshold mismatches have similar contribution to the variability of PoI. While such structure-based information defined by circuit topology is completely ignored by the traditional performance modeling techniques [7], it will be exploited in this paper to improve the modeling accuracy and/or reduce the modeling cost.
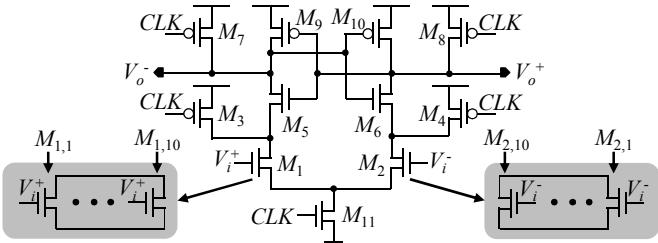
Fig. 1. Simplified circuit schematic is shown for a latch-based comparator where each transistor has 10 multipliers. For illustration purpose, the multipliers of the two input transistors M_1 and M_2 are highlighted in grey color.

In particular, we propose to "group" the model coefficients that are expected to have similar magnitude. These groups are formed purely based on circuit topology and design knowledge without running any transistor-level simulation. The aforementioned grouping strategy would help us to accurately identify the important (i.e., non-zero) model coefficients based on very few sampling points. Our proposed structure-aware performance modeling technique consists of two major steps: (i) model coefficient grouping, and (ii) model coefficient fitting. In what follows, we will describe the technical details of these two steps and highlight the novelty.

A. Model Coefficient Grouping

Our objective is to partition all model coefficients into several groups where the coefficients within the same group share similar magnitude. If a model coefficient has different magnitude from all other coefficients, it should form a separate group containing a single coefficient only. In this paper, the grouping step is performed by taking the design knowledge from a user as the input. The following shows two examples how coefficient groups can be formed.

- **Example I**: The same mismatch variables (e.g., ΔV_{th}'s) from all multipliers of the same transistor can be grouped together. Taking the comparator in Fig. 1 as an example, the threshold variations of all multipliers for each device (e.g., $\{\Delta V_{th1,k}; k = 1, 2, ..., 10\}$ for the transistor M_1) can be grouped together. Here, we assume that all multipliers of the same transistor are placed close to each other and there is no significant systematic difference between these multipliers. Hence, the mismatch variables of all multipliers will have a similar impact on the variability of PoI. Note that once the layout of a circuit is given, setting up the coefficient groups in this example is straightforward and it is almost independent of the circuit functionality.

- **Example II**: The same mismatch variables (e.g., ΔV_{th}'s) from different transistors can be grouped together, if our prior knowledge reveals that these mismatch variables have similar contribution to the variability of PoI. Again, taking the comparator in Fig. 1 as an example, it is easy to note that the mismatch variables of the two input transistors M_1 and M_2 have a similar impact on the variability of input offset. Hence, the mismatch variables $\{\Delta V_{th1,k}; k = 1, 2, ..., 10\}$ for M_1 and $\{\Delta V_{th2,k}; k = 1, 2, ..., 10\}$ for M_2 can be grouped together, if the input offset is considered as PoI. Compared to Example I, the grouping strategy in this example is circuit- and performance-dependent. It is based upon a deep understanding of the circuit operation.

Once the coefficient groups are identified, a statistical algorithm should be further applied to select the important groups containing non-zero model coefficients and then solve these coefficient values based on a small number of sampling points. In the next sub-section, a modified S-OMP algorithm will be introduced to address this coefficient fitting problem.

B. Model Coefficient Fitting

The key idea of S-OMP [8] is to identify the non-zero model coefficients based on the "correlation" between each group and the PoI. To illustrate the mathematical formulation of S-OMP, we first re-order all columns of the matrix $\mathbf{X}$ in (3) based on the coefficient groups. Namely, the matrix $\mathbf{X}$ is re-arranged as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{1,1} & \cdots & \mathbf{x}_{1,N_1} & \cdots & \mathbf{x}_{K,1} & \cdots & \mathbf{x}_{K,N_K} \end{bmatrix}, \qquad (4)$$

where the vector $\mathbf{x}_{k,n}$ represents one column of the matrix $\mathbf{X}$ and all vectors $\{\mathbf{x}_{k,n}; n = 1, 2, ..., N_k\}$ belonging to the k-th group are adjacent to each other in the re-ordered matrix $\mathbf{X}$. After re-ordering, the matrix $\mathbf{X}$ in (4) shows a unique structure

that contains K groups in total and N_k vectors associated with the k-th group.

Next, S-OMP selects the most important group based on the following criterion:

$$\arg\max_{k} \quad g_k = \frac{1}{N_k} \cdot \sum_{n=1}^{N_k} \left| \langle \mathbf{x}_{k,n}, \mathbf{y} \rangle \right|, \qquad (5)$$

where the operator $\langle \bullet, \bullet \rangle$ calculates the inner product between two vectors, and g_k stands for the "correlation score" for the k-th group. Intuitively, if the score g_k is large, it implies that all terms (either the constant term or the device-level random variables) within the k-th group are strongly correlated with the PoI y. Hence, the corresponding model coefficients are likely to be non-zero, and the k-th group should be identified as an important group.

The correlation score g_k in (5) is derived based upon the assumption that we do not know the sign of the correlation (i.e., either positive or negative) between $\mathbf{x}_{k,n}$ and $\mathbf{y}$. In this case, the "average" correlation is calculated for the absolute value of the individual correlation. In practice, if we know the "relative sign" of the correlation (i.e., whether the correlation between $\mathbf{x}_{k,n}$ and $\mathbf{y}$ has the same sign for any $n \in \{1, 2, ..., N_k\}$), such sign information can be explicitly incorporated into the formulation in (5), resulting in the following criterion to calculate the correlation score:

$$\arg\max_{k} \quad g_k = \frac{1}{N_k} \cdot \left| \sum_{n=1}^{N_k} s_{k,n} \cdot \langle \mathbf{x}_{k,n}, \mathbf{y} \rangle \right|, \qquad (6)$$

where $s_{k,n} \in \{-1, 1\}$ represents the relative sign of the correlation between $\mathbf{x}_{k,n}$ and $\mathbf{y}$. For instance, if the correlation between $\mathbf{x}_{k,1}$ and $\mathbf{y}$ and the correlation between $\mathbf{x}_{k,2}$ and $\mathbf{y}$ share the same sign, $s_{k,1}$ and $s_{k,2}$ should be set to the same value. Otherwise, they should be set to two opposite values.

Compared to (5), Eq. (6) first calculates the average of the individual correlation, and then takes the absolute value of the average correlation. As such, the accuracy of the estimated correlation score g_k can be substantially improved. To understand the reason, we consider a simple example where the correlation between $\mathbf{x}_{k,n}$ and $\mathbf{y}$ is zero for any $n \in \{1, 2, ..., N_k\}$. In this case, the inner product $\langle \mathbf{x}_{k,n}, \mathbf{y} \rangle$ may not be exactly zero, since it is calculated from a small number of sampling points. Hence, if the correlation score g_k is estimated by (5), the average of the absolute value of the individual correlation is not equal to zero. On the other hand, the correlation score g_k calculated by (6) can be much closer to zero (i.e., the true correlation value), since it attempts to average out the error of the individual correlation. The difference between (5) and (6) will be further discussed in Section III, along with our numerical examples.

Once an important group is identified by S-OMP, all model coefficients associated with this group are considered to be non-zero. The performance model in (1) is fitted with all other model coefficients set to zero. Next, the modeling error is calculated and an additional group is identified by S-OMP to maximally reduce the modeling error. The aforementioned iteration steps are repeatedly performed until the modeling error is sufficiently small. Algorithm 1 summarizes the major steps of the proposed SAM technique based on S-OMP. Due to the page limit, more details about S-OMP are not included in this paper, but they can be found in [8].

Algorithm 1: Structure-Aware Modeling (SAM)

1. Start from a set of sampling points and formulate the linear equation in (2).
2. Group the model coefficients based on circuit structure and design knowledge.
3. Re-order the matrix $\mathbf{X}$ to the form of (4) where the column vectors within the same group are adjacent to each other.
4. Set the residual $\mathbf{r} = \mathbf{y}$ and the set $\Omega = \{\}$.
5. If the sign information is unknown for the correlation between $\mathbf{x}_{k,n}$ and $\mathbf{y}$, select the most important group (say, the k-th group) by using:

$$\arg\max_{k} \quad g_k = \frac{1}{N_k} \cdot \sum_{n=1}^{N_k} \left| \langle \mathbf{x}_{k,n}, \mathbf{r} \rangle \right|. \qquad (7)$$

Otherwise, if the sign information is known, select the most important group (say, the k-th group) by using:

$$\arg\max_{k} \quad g_k = \frac{1}{N_k} \cdot \left| \sum_{n=1}^{N_k} s_{k,n} \cdot \langle \mathbf{x}_{k,n}, \mathbf{r} \rangle \right|. \qquad (8)$$

6. Update the set $\Omega = \Omega \cup \{(k, 1), (k, 2), ..., (k, N_k)\}$.
7. Determine the model coefficients defined by the set Ω by solving the following least-squares fitting problem:

$$\arg\min_{\alpha_{k,n}; (k,n) \in \Omega} \quad \left\| \mathbf{y} - \sum_{(k,n) \in \Omega} \alpha_{k,n} \cdot \mathbf{x}_{k,n} \right\|_2^2. \qquad (9)$$

8. Update the residual $\mathbf{r}$:

$$\mathbf{r} = \mathbf{y} - \sum_{(k,n) \in \Omega} \alpha_{k,n} \cdot \mathbf{x}_{k,n}. \qquad (10)$$

9. If the residual is sufficiently small, stop iteration and set $\alpha_{k,n} = 0$ for any $(k, n) \notin \Omega$. Otherwise, go to Step 5.

III. EXPERIMENTAL RESULT

In this section, two circuit examples designed in a commercial 32nm CMOS process are used to demonstrate the efficacy of the proposed SAM method. Our objective is to build linear performance models to study the circuit-level performance variability with respect to device mismatches. Such a performance modeling problem is non-trivial, because device mismatches must be modeled by a large number of independent random variables, rendering a high-dimensional variation space. For testing and comparison purposes, two different techniques are implemented: (i) the traditional sparse regression method based on OMP [7], and (ii) the proposed SAM method (i.e., Algorithm 1). For each method, we measure the modeling error from 50 repeated runs where the sampling points are independently generated for each run. All numerical experiments are performed on a 2.5GHz Linux server with 64GB memory.

A. Ring Oscillator

Shown in Fig. 2 is the simplified circuit schematic of a ring oscillator consisting of 21 identically-sized inverters. There are 2835 independent random variables defined by the process design kit to model device mismatches. The oscillation frequency is considered as the PoI in this example.

Since all inverters in Fig. 2 are identically sized, the random mismatches of all NMOS (or PMOS) transistors should have a similar impact on the frequency variability. Hence, the mismatch variables associated with all NMOS (or PMOS) transistors are grouped together. Such a grouping strategy is similarly applied to the load capacitors that are connected to the output of each inverter.
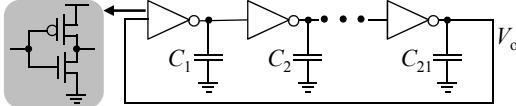

Fig. 2. Simplified circuit schematic is shown for a ring oscillator consisting of 21 identically-sized inverters.

In this example, the sign information can be easily defined. For all random variables within the same group, they play a similar role on the PoI and, hence, the corresponding model coefficients should have the same sign. Even though the sign information is available in this example, two versions of Algorithm 1, with and without the sign information respectively, are implemented and tested for comparison purposes.

Fig. 3(a) shows the modeling error as a function of the number of sampling points. Note that, given the same number of sampling points, SAM achieves superior modeling accuracy over OMP. On the other hand, to achieve the same modeling accuracy, SAM offers about $2\times$ runtime speedup over OMP, as shown in Table I. It is also observed from Fig. 3(a) that if the sign information is incorporated into Algorithm 1, SAM is able to further reduce the modeling error, especially when the number of sampling points is small. This observation is consistent with our discussion in Section II.
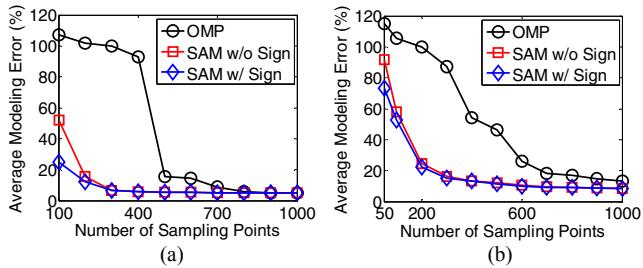

Fig. 3. The average modeling error from 50 repeated runs is plotted as a function of the number of sampling points: (a) ring oscillator, and (b) latch-based comparator.

TABLE I
COMPUTATIONAL COST FOR PERFORMANCE MODELING

	Ring Oscillator		Comparator	
	OMP	SAM	OMP	SAM
# of Sampling Points	800	400	1000	400
Modeling Error	5.72%	5.80%	13.33%	13.17%
Simulation Cost (Hour)	3.60	1.80	5.50	2.20
Fitting Cost (Sec.)	111.90	2.97	141.85	5.79
Total Cost (Hour)	3.63	1.80	5.54	2.20

B. Latch-based Comparator

As a second example, we consider the latched-based comparator shown in Fig. 1. In this example, there are 4950 independent random variables modeling device mismatches and the input offset is considered as our PoI. As previously mentioned in Section II, each transistor in Fig. 1 is composed

of 10 multipliers. The mismatch variables from all multipliers of the same transistor are grouped together. In addition, by considering the symmetry of the circuit topology, we group the mismatch variables corresponding to the following transistor pairs: $\{M_1, M_2\}$, $\{M_3, M_4\}$, $\{M_5, M_6\}$, $\{M_7, M_8\}$ and $\{M_9, M_{10}\}$. Based on the aforementioned grouping strategy, the sign information within each group can be derived according to the structure of the comparator.

Fig. 3(b) shows the modeling error as a function of the number of sampling points. Similar to the ring oscillator example, SAM requires substantially less sampling points than OMP to achieve the same modeling accuracy. Compared to OMP, SAM offers about $2.5\times$ runtime speedup, as shown in Table I. However, unlike the ring oscillator example, the sign information does not make a significant difference in modeling accuracy for the comparator example. This observation is made, because the error posed by other steps in Algorithm 1 (e.g., the regression error of the least-squares fitting in Step 7 of Algorithm 1) dominates the overall modeling error, when the number of sampling points is extremely small. For this reason, even though the sign information helps to accurately identify the important groups, it does not substantially reduce the modeling error.

IV. CONCLUSION

In this paper, a novel SAM technique is proposed for structure-aware high-dimensional performance modeling of AMS circuits with consideration of process variations. The key idea of SAM is to group the model coefficients that have similar magnitude. As such, only an extremely small number of sampling points are required to fit a high-dimensional performance model, thereby substantially reducing the computational cost of performance modeling. As is demonstrated by our circuit examples designed in a commercial 32nm CMOS process, SAM achieves more than $2\times$ runtime speedup over the traditional sparse regression technique.

V. REFERENCES

[1] Semiconductor Industry Associate, *International Technology Roadmap for Semiconductors*, 2011.
[2] X. Li, J. Le and L. Pileggi, *Statistical Performance Modeling and Optimization*, Now Publishers, 2007.
[3] T. McConaghy, "High-dimensional statistical modeling and analysis of custom integrated circuits," *IEEE CICC*, 2011.
[4] X. Li, J. Le, P. Gopalakrishnan and L. Pileggi, "Asymptotic probability extraction for nonnormal performance distributions," *IEEE Trans. on CAD*, vol. 26, no. 1, pp. 16-37, Jan. 2007.
[5] M. Sengupta, S. Saxena, L. Daldoss, G. Kramer, S. Minehane and J. Cheng, "Application-specific worst case corners using response surfaces and statistical models," *IEEE Trans. on CAD*, vol. 24, no. 9, pp. 1372-1380, 2005.
[6] A. Dharchoudhury and S. Kang, "Worse-case analysis and optimization of VLSI circuit performance," *IEEE Trans. on CAD*, vol. 14, no. 4, pp. 481-492, Apr. 1995.
[7] X. Li, "Finding deterministic solution from underdetermined equation: large-scale performance modeling of analog/RF circuits," *IEEE Trans. on CAD*, vol. 29, no. 11, pp. 1661-1668, Nov. 2010.
[8] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *J. Signal Process.*, vol. 86, pp. 572-588, Apr. 2006.