Efficient Statistical Capacitance Variability Modeling with Orthogonal Principle Factor Analysis

Rong Jiang†, Wenyin Fu†, Janet Meiling Wang‡, Vince Lin♮ and Charlie Chung-Ping Chen†
Department of Electrical & Computer Engineering, University of Wisconsin, Madison, WI 53705 †
Department of Electrical & Computer Engineering, University of Arizona, Tucson, AZ 85721 ‡
Springsoft, No. 25 Industry East Road IV, Science-Based Industrial Park, Hsinchu, Taiwan 300 ♮
{jiang,wenyinf}@cae.wisc.edu, wml@ece.arizona.edu, vince_lin@springsoft.com.tw, chen@engr.wisc.edu

Abstract— **Due to the ever-increasing complexity of VLSI designs and IC process technologies, the mismatch between a circuit fabricated on the wafer and the one designed in the layout tool grows ever larger. Therefore, characterizing and modeling process variations of interconnect geometry has become an integral part of analysis and optimization of modern VLSI designs. In this paper, we present a systematic methodology to develop a closed form capacitance model, which accurately captures the nonlinear relationship between parasitic capacitances and dominant global/local process variation parameters. The explicit capacitance representation applies the orthogonal principle factor analysis to greatly reduce the number of random variables associated with modeling conductor surface fluctuations while preserving the dominant sources of variations, and consequently the variational capacitance model can be efficiently utilized by statistical model order reduction and timing analysis tools. Experimental results demonstrate that the proposed method exhibits over $100\times$ speedup compared with Monte Carlo simulation while having the advantage of generating explicit variational parasitic capacitance models of high order accuracy.**

Categories and Subject Descriptors: B.7.2 [Integrated Circuits]: Design Aids – simulation, verification.

General Terms: Design, Algorithms.

Keywords: Process variations, capacitance, parasitic extraction, random variable reduction, principle factor analysis.

I. Introduction

As VLSI circuits have entered deep submicron dimensions, increasing complexity of VLSI designs and IC process technologies increases the mismatch between design and manufacturing. Process induced variations in the device and interconnect structures are posing a significant challenge to parasitic modeling and signal integrity analysis. To determine the extent of such effects, the distribution of various electrical parameters, such as interconnect resistances and capacitances due to variations in the manufacturing process must be determined. Once this distribution is known, which is also called the design envelope, the design corners can then be identified.

During the modern Damascene process, the dielectric is usually patterned by reactive ion etching (RIE), followed by the linear and metal (Cu) deposition. Then chemical-mechanical planarization (CMP) is applied to remove excessive metal and provide a global planarization. During RIE, the ideal eroded rectangular trenches in dielectric, and hence later deposited metals and liners, may become trapezoidal due to the aspect dependent etch rate (ARDE) effect. During the CMP overpolishing process, regions of high metal pattern density tend to erode faster and hence show higher metal and dielectric removal rates than regions of low metal pattern density [1]. The non-uniform metal removal rates across the wafer can lead to varying metal line thickness for interconnects sited in the same metal layer. Also during the pattern transferring in lithography process, photomask geometries may be distorted due to nonlinear distortions caused by optical diffraction and resist process effects, so that the tips and corners of interconnect will become round shape.

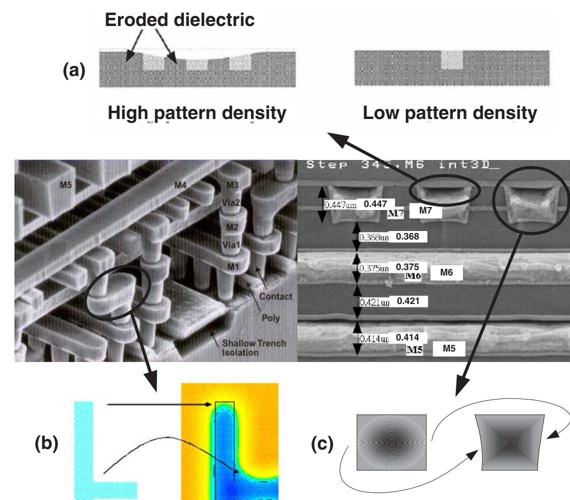


Fig. 1. Process variations due to (a) chemical-mechanical planarization, (b) optical diffraction, and (c) chemical etching. (Picture courtesy of TSMC, Hsin-Chu, Taiwan.)

Therefore, for deep submicron technologies, a combination of device physics, die location dependence, optical proximity effects, micro-loading in etching and deposition may lead to heterogeneous and non-monotonic relationships among the process random variables. Also parasitic capacitance does not change monotonically or linearly according to those random parameters, which have varying effects on interconnect geometries depending on local characteristics of the layout and uncertainties in fabrication. Since all these process variations are random in nature, statistical parasitic capacitance models

having the ability to capture those complicated nonlinear relationships become indispensable.

Furthermore, capacitance extraction with process variations can never be the final goal. Capacitance variation analysis needs to provide a model fully compatible with statistical model order reduction and statistical timing analysis tools, most of which require representing parasitic capacitances as functions of some common random variables [2]–[6]. Also recent study shows that the first order canonical model is not sufficient enough to represent the nonlinear dependency of parasitic capacitances on many variation sources [7]. To our best knowledge, although many efficient 3D capacitance extraction algorithms [8]–[15] have been proposed in the literature and there have been some pioneer works [7], [16] on capacitance extraction with the consideration of process variations, no algorithm has the functionality to efficiently supply an explicit statistical capacitance model with high order accuracy.

Unsatisfied with those limitations, in this paper, we propose a systematic way to develop an explicit quadratic form representation for parasitic capacitance and the quadratic model can be easily extended to achieve even higher order accuracy. First, we adopt the boundary element method with hierarchical panel refinement to generate the sparse linear system after conductor surface discretization. The sparse system can be efficiently solved by preconditioned iterative matrix solvers, such as preconditioned conjugate gradient (PCG) or GMRES. The surface fluctuation of a conductor is then modeled as statistical position perturbations of those most delicate panels normal to their smooth surfaces. Each position perturbation can be modeled as a random variable and the spatial correlations among those perturbations can be described by the Gaussian correlation function with an appropriate correlation length. Then we apply orthogonal principle factor analysis (OPFA) to find several dominant global/local factors that cause the position variations. The quadratic approximation in terms of those dominant factors can be obtained by solving the sparse linear system several times.

The rest of paper will be organized as follows: Section II introduces the boundary element method and the hierarchical capacitance algorithm. Section III presents the main idea of our variational capacitance extraction. Extensive simulation results are shown in Section IV and the paper is concluded in Section V.

II. PRELIMINARY

The capacitances among m conductors can be summarized by an $m \times m$ capacitance matrix C,

$$C\tilde{\boldsymbol{v}} \quad = \quad \tilde{\boldsymbol{q}}, \tag{1}$$

where $\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{v}} \in \Re^{m\times 1}$ are conductor charge distribution and potential vectors, respectively. The diagonal entries C_{ii} of C are positive, representing the self-capacitance of conductor i. The non-diagonal entries C_{ij} are negative, representing the coupling capacitance between conductors i and j. The jth column of C can be calculated by solving for the total charges on each of the conductors when the jth conductor is at unit potential and all the other conductors are at zero potential. Then the charge on conductor i, $\tilde{\boldsymbol{q}}_i$, is equal to C_{ij}. This procedure is repeated m times to compute all columns of C.

A. BEM Capacitance Extraction

Boundary element methods (BEM), also referred to as panel methods or the method of moments, have been adopted as the main approach for 3D capacitance calculation. Due to the fact that the charge is restricted to the surface of the conductors, the surfaces of m conductors with non-uniform charge distribution need to be discretized into a total of n two-dimensional panels and the charge distribution on each panel is assumed to be even. Then for each panel k, an equation is written that relates the potential at the center of the kth panel to the sum of contributions to that potential from the charge distribution on all n panels and the contribution from the lth panel is determined by the potential coefficient,

$$P_{kl} = \frac{1}{a_l}\int_{panel\ l} \frac{1}{\parallel x_k - x' \parallel} da' \approx \frac{1}{\parallel x_k - x_l \parallel}, \tag{2}$$

where x_l and x_k are the centers of the lth and k^{th} panels.

Then a system of equations can be constructed to solve for the discretized conductor surface charges

$$P\boldsymbol{q} \quad = \quad \boldsymbol{v}, \tag{3}$$

where $P \in \Re^{n\times n}$ is the potential coefficient matrix and $\boldsymbol{q}, \boldsymbol{v} \in \Re^{n\times 1}$ are panel charge distribution and potential vectors. To compute the jth column of the capacitance matrix, Eq. 3 must be solved for $\boldsymbol{q}$, given a $\boldsymbol{v}$ vector where $\boldsymbol{v}_k = 1$ if panel k is on the jth conductor, and $\boldsymbol{v}_k = 0$ otherwise [8]. Then C_{ij} of the capacitance matrix is computed by summing all the panel charges on the ith conductor,

$$C_{ij} \quad = \sum_{k\in conductor\ i} \boldsymbol{q}_k. \tag{4}$$

B. Hierarchical Capacitance Algorithms

The main obstacle of solving $\boldsymbol{q}$ is that the coefficient matrix in Eq. 3 is very dense and direct linear system solvers, such as Gaussian elimination or Cholesky decomposition, become computationally intractable if the number of panels exceeds several hundred. Therefore, multipole accelerated [8], [14] and hierarchical algorithms [9], [10] have been proposed to address this problem. Hierarchical algorithms will be introduced in this section.

Conductor surfaces can be hierarchically, instead of uniformly, divided into smaller panels [9], [10]. The hierarchical panel refinement can be fully described by a multiple-tree shown in Fig. 2, in which the root panel of each tree corresponds to a conductor surface. If the estimated potential coefficient between two panels is larger than a threshold value P_ϵ, they are further divided into smaller panels. Otherwise, a link recording the potential coefficient is created between these two panels.

All recorded potential coefficients compose a link matrix $H \in \Re^{N\times N}$ [10], where N is the number of all panels.

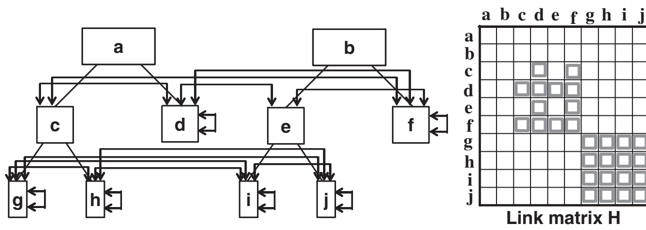
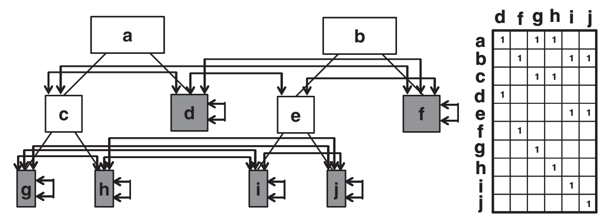


Fig. 2. Hierarchical capacitance extraction algorithm represented by a multiple tree data structure. The dimension of the link matrix H is equal to the total number of panels.



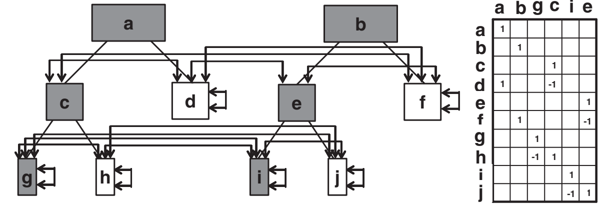


Fig. 3. Sparsifying the potential coefficient matrix based on basis transformation. The basis including all leaf panels in (a) will lead to a dense potential coefficient matrix while the potential coefficient matrix P related to the basis containing all root panels and left hand side panels (b) is sparse and contains $O(n)$ non-zero entries.

For any two panels i and j with no links in between, the corresponding entry in H is zero, otherwise, the recorded potential coefficient evaluated by using Eq. 2 is filled into H_{ij}.

For each panel in the hierarchical data structure, its charge is equal to the summation of charges on its two child panels. Therefore, one can choose a specific set of panels, called basis panels, such that all panel charges can be uniquely represented as linear combinations of charges on those panels [17]. The coefficient matrix of those linear combinations is called the structure matrix $J \in \mathcal{R}^{N \times n}$, where n is the number of leaf panels.

For a particular multiple-tree structure, there are many possible bases and each of them has its own structure matrix J and potential coefficient matrix P, which can be expressed in terms of H and J as

$$P = J^T H J. \tag{5}$$

Then in the new linear system, the $\boldsymbol{q}$ vector in Eq. 3 will represent charges on those basis panels instead of leaf panels. It has been discovered that the potential coefficient matrix in Eq. 5 is dense when leaf panels are chosen as the basis. On the contrary, if all root panels and left hand side panels are chosen as the basis, it is provable that the P matrix related to this basis is sparse and contains $O(n)$ non-zeros (Fig. 3). Therefore, equations in Eq. 3 can be efficiently solved by preconditioned Krylov subspace solvers in linear time.

III. Statistical Capacitance Extraction

The following four issues will be discussed in this section for modeling parasitic capacitance variations: (1) how to efficiently solve the system equations associated with the variational capacitance model; (2) how to mathematically model the surface fluctuation due to process variations; (3) how to reduce the large number of random variables used to model the surface fluctuation; (4) how to obtain the probability density function without using time consuming Monte Carlo simulation.

A. Variational Capacitance Approximation

Assume for now that process variations induce some perturbations in the nominal potential coefficient P_{kl} between panel k and panel l in Eq. 2, and the variational potential coefficient $\bar{P}_{kl}$ can be represented in terms of the nominal value P_{kl} and k normal random variables $\boldsymbol{\delta} = [\delta_1 \ \delta_2 \ \cdots \ \delta_k]^T$ as

$$\bar{P}_{kl} = P_{kl} + \sum_i \Delta P^i_{kl} \delta_i + \sum_{i,j} \Delta P^{ij}_{kl} \delta_i \delta_j + h.o.t. \tag{6}$$

How to represent $\bar{P}_{kl}$ in the such a form will be presented in the following sections.

The expression of $\bar{P}_{kl}$ in terms of $\boldsymbol{\delta}$ can be extended to higher orders. If the first three terms is used, Eq. 6 is the quadratic form of the potential coefficient $\bar{P}_{kl}$. The second term represents the canonical linear model while the third term captures the nonlinear relationship between $\bar{P}_{kl}$ and $\boldsymbol{\delta}$. In the rest of this paper, our discussion will be based on the quadratic form, since higher order approximations can be easily extended using the presented derivation.

Since each entry of the variational link matrix $\bar{H}$ has the form shown in Eq. 6, the entire $\bar{H}$ can also be expressed in a quadratic form as follows:

$$\bar{H} = H + \sum_i \Delta H^i \delta_i + \sum_{i,j} \Delta H^{ij} \delta_i \delta_j, \tag{7}$$

where $H, \Delta H^i, \Delta H^{ij} \in \Re^{N \times N}$ are constant coefficient matrices.

By using Eq. 5, the variational potential coefficient matrix $\bar{P}$ can also be represented in terms of P and $\boldsymbol{\delta}$

$$\begin{aligned} \bar{P} &= J^T H J + \sum_i J^T \Delta H^i J \delta_i + \sum_{i,j} J^T \Delta H^{ij} J \delta_i \delta_j, \\ &= P + \underbrace{\sum_i \Delta P^i \delta_i + \sum_{i,j} \Delta P^{ij} \delta_i \delta_j}_{\Delta P}, \end{aligned} \tag{8}$$

where $\Delta P^i = J^T \Delta H^i J$ and $\Delta P^{ij} = J^T \Delta H^{ij} J$. P is the potential coefficient matrix without considering the process

variations, and ΔP, which is the summation of the second and third terms in Eq. 8, represents the variational part of $\bar{P}$.

Let $\bar{\boldsymbol{q}}$ denote the variational charge distribution vector, our goal is then to express $\bar{\boldsymbol{q}}$ in a quadratic form, such that

$$\bar{\boldsymbol{q}} = \boldsymbol{q} + \underbrace{\sum_i \Delta \boldsymbol{q}^i \delta_i + \sum_{i,j} \Delta \boldsymbol{q}^{ij} \delta_i \delta_j}_{\Delta \boldsymbol{q}}, \quad (9)$$

where $\boldsymbol{q}, \Delta \boldsymbol{q}^i, \Delta \boldsymbol{q}^{ij} \in \Re^{n \times 1}$. From Eq. 9, it is clear that the quadratic expressions of self and coupling capacitances can be easily obtained by using Eq. 4.

From Eq. 8 and Eq. 9, the variational linear system can be then represented as

$$(P + \Delta P)(\boldsymbol{q} + \Delta \boldsymbol{q}) = \boldsymbol{v}. \quad (10)$$

Substituting the normal equation in Eq. 3 into Eq. 10 and applying the Taylor expansion, $\Delta \boldsymbol{q}$ can be expressed as

$$\begin{aligned} \Delta \boldsymbol{q} &= -(I + P^{-1}\Delta P)P^{-1}\Delta P \boldsymbol{q} \\ &= -\underbrace{P^{-1}\Delta P \boldsymbol{q}}_{\Delta \boldsymbol{q}_1} + \underbrace{P^{-1}\Delta P P^{-1} \Delta P \boldsymbol{q}}_{\Delta \boldsymbol{q}_2} + \cdots \\ &= A\boldsymbol{q} + A^2 \boldsymbol{q} + \cdots = \sum_{i=1}^{\infty} A^i \boldsymbol{q}, \end{aligned} \quad (11)$$

where $A = -P^{-1}\Delta P$.

Theorem 1: The variational charge distribution vector $\Delta \boldsymbol{q}$ can be represented as $\Delta \boldsymbol{q} = \sum_{i=1}^{\infty} A^i \boldsymbol{q}$, where $A = -P^{-1}\Delta P$. The Taylor expansion series of $\Delta \boldsymbol{q}$ converges under the condition $\| P^{-1}\Delta P \|_p < 1$. So high order terms can be iteratively calculated by using the following equation

$$P\Delta \boldsymbol{q}_{i+1} = -\Delta P \Delta \boldsymbol{q}_i. \quad (12)$$

Since in practice, the perturbation matrix ΔP is normally smaller than the normal potential coefficient matrix P, the convergence condition can be almost always satisfied. Let the quadratic form representation of the first term on the right hand side of Eq. 11, $\Delta \boldsymbol{q}_1$, to be

$$\Delta \boldsymbol{q}_1 = \sum_i \Delta \boldsymbol{q}_1^i \delta_i + \sum_{i,j} \Delta \boldsymbol{q}_1^{ij} \delta_i \delta_j. \quad (13)$$

By using Eq. 8 and $P\Delta q_1 = -\Delta P q$, we can get

$$\begin{aligned} P\Delta \boldsymbol{q}_1^i &= -\Delta P^i \boldsymbol{q}, \\ P\Delta \boldsymbol{q}_1^{ij} &= -\Delta P^{ij} \boldsymbol{q}, \end{aligned} \quad (14)$$

Therefore, the quadratic expression of $\Delta \boldsymbol{q}_1$ can be calculated by solving $(k + k^2)$ linear systems. Since P is sparse, each linear system in Eq. 14 can be efficiently solved by preconditioned iterative methods with $O(n)$ complexity. So the total complexity of solving $\Delta \boldsymbol{q}_1$ is $O((k^2 + k)n)$. Usually, the number of random variables, k, is much smaller than the total number of leaf panels n.

The second term $\Delta \boldsymbol{q}_2 = \sum_{i,j} \Delta \boldsymbol{q}_2^{ij} \delta_i \delta_j$ in Eq. 11 can be obtained by using $\Delta \boldsymbol{q}_1$

$$P\Delta \boldsymbol{q}_2 = -\Delta P \Delta \boldsymbol{q}_1. \quad (15)$$

Let the right hand side vector in Eq. 15 to be $\tilde{\boldsymbol{q}}_1 = \Delta P \Delta \boldsymbol{q}_1$, then the quadratic approximation of $\tilde{\boldsymbol{q}}_1$ can be expressed as

$$\tilde{\boldsymbol{q}}_1 = \sum_{i,j} \Delta P^i \Delta \boldsymbol{q}_1^j \delta_i \delta_j + h.o.t. \quad (16)$$

Therefore, the coefficient vectors of $\Delta \boldsymbol{q}_2$ can be obtained by

$$P\Delta \boldsymbol{q}_2^{ij} = -\Delta P^i \Delta \boldsymbol{q}_1^j, \quad (17)$$

So the quadratic expression of $\Delta \boldsymbol{q}_2$ requires the solving of k^2 linear systems and hence the complexity is $O(k^2 n)$.

Therefore, by using the quadratic expressions of $\Delta \boldsymbol{q}_1$ and $\Delta \boldsymbol{q}_2$, the the quadratic expression of $\Delta \boldsymbol{q}$ is then obtained by

$$\begin{aligned} \Delta \boldsymbol{q}^i &= \Delta \boldsymbol{q}_1^i, \\ \Delta \boldsymbol{q}^{ij} &= \Delta \boldsymbol{q}_1^{ij} + \Delta \boldsymbol{q}_2^{ij}. \end{aligned} \quad (18)$$

So the total computational complexity of calculating the quadratic form of $\Delta \boldsymbol{q}$ is $O(k^2 n)$.

Also, one may notice that the first order terms are only generated by $\Delta \boldsymbol{q}_1$ while the second order terms are generated by $\Delta \boldsymbol{q}_1$ and $\Delta \boldsymbol{q}_2$. Therefore, for the quadratic form approximation, when $i > 2$, $\Delta \boldsymbol{q}_i$ does not contain the first and second order terms, and hence can be safely truncated. In the follow subsections, we will present how to express the variational potential coefficients in a form given in Eq. 6 in terms of k normal random variables.

B. Process Variation Modeling

After the hierarchical panel discretization process, the positions of those most delicate panels, leaf panels, may be varying due to process variations. The surface fluctuation of a conductor can be described as a statistical perturbation on each nominal leaf panel smooth surface along its normal direction as shown in Fig. 4.

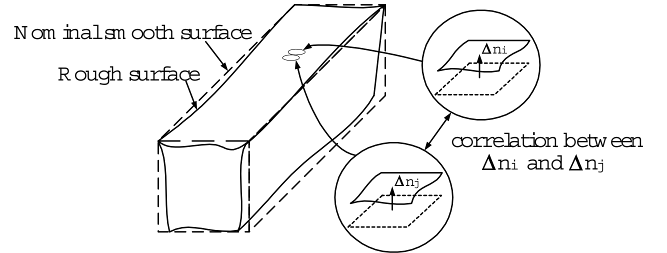


Fig. 4. Process variation modeling with correlated statistical position perturbations on leaf panels.

Although leaf panel position variations may not be truly random, they can often be accurately modeled by assuming an appropriate spatial correlation [16]. We denote leaf panel position variations as a random variable vector $\Delta \tilde{\boldsymbol{n}}$, where the ith element in $\Delta \tilde{\boldsymbol{n}}$, $\Delta \tilde{\boldsymbol{n}}_i$, is the random perturbation on the leaf panel i. For simplicity, one can assume that the expectation of $\Delta \tilde{\boldsymbol{n}}$ is $\mu(\Delta \tilde{\boldsymbol{n}}) = \boldsymbol{0}$.

Obviously, the larger the distance between two leaf panels, the weaker the correlation will be. This spatial relationship can be accurately modeled by using the Gaussian correlation

function [16]. For two leaf panels i and j, the correlation between them is determined by

$$\Gamma_{ij} = e^{-\|x_i - x_j\|^2/\eta^2}, \tag{19}$$

where e is Euler constant and η is user-specified correlation length. x_i and x_j are the centers of leaf panels i and j, respectively. Then the correlation matrix can be written as

$$\Gamma(\Delta\tilde{\boldsymbol{n}}) = (\Gamma_{ij})_{n\times n}. \tag{20}$$

Many small terms in $\Gamma(\Delta\tilde{\boldsymbol{n}})$ can be truncated to make it sparse if the corresponding two leaf panels are separated faraway enough. Also if the variance on leaf panel i is assumed to be σ_i, then the variance-covariance matrix Σ of $\Delta\tilde{\boldsymbol{n}}$ can be obtained as

$$\Sigma(\Delta\tilde{\boldsymbol{n}}) = (\Gamma_{ij}\sigma_i\sigma_j)_{n\times n}. \tag{21}$$

Therefore, the surface fluctuation can be modeled by the random vector $\Delta\tilde{\boldsymbol{n}}$ with mean $\mu(\Delta\tilde{\boldsymbol{n}}) = \mathbf{0}$ and the variance-covariance matrix $\Sigma(\Delta\tilde{\boldsymbol{n}})$ given in Eq. 21.

C. Random Variable Reduction

Although the process variations can be modeled as position perturbations on leaf panels, the number of random variables can easily exceeds several thousand and this may greatly limit the size of the problem that can be analyzed.

The position perturbations of leaf panels may be caused by many unobservable variation sources, either global or local. However, some of them may have significant effects on the conductor surface fluctuation while others may not, and hence those non-significant factors can be safely neglected in our modeling process. In multivariate statistics, determining the dominant unobservable variation sources can be performed by principle factor analysis (PFA) [18] based on either the correlation matrix $\Gamma(\Delta\tilde{\boldsymbol{n}})$ in Eq. 20 or the variance-covariance matrix $\Sigma(\Delta\tilde{\boldsymbol{n}})$ in Eq. 21.

The random variable vector $\Delta\tilde{\boldsymbol{n}}$ representing the perturbations on leaf panels is observable, and has n components with the mean vector $\mu(\Delta\tilde{\boldsymbol{n}}) = \mathbf{0}$ and the variance-covariance matrix $\Sigma(\Delta\tilde{\boldsymbol{n}})$ given in Eq. 21. The principle factor analysis postulates that $\Delta\tilde{\boldsymbol{n}}$ is linearly dependent upon k ($k << n$) unobservable random variables $\boldsymbol{\delta}$, called common factors. Those k common factors are used to model the unknown and unobservable dominant process variation sources that inherently induce the perturbations on leaf panels.

Furthermore, the orthogonal principle factor analysis (OPFA), also referred to as principle component model, assumes that

$$\begin{aligned} \mu(\boldsymbol{\delta}) &= \mathbf{0}, \\ \Sigma(\boldsymbol{\delta}) &= I. \end{aligned} \tag{22}$$

The goal of orthogonal principle factor analysis is to find a loading matrix $L \in \Re^{n\times k}$, such that

$$\underset{(n\times 1)}{\Delta\tilde{\boldsymbol{n}}} = \underset{(n\times k)}{L} \times \underset{(k\times 1)}{\boldsymbol{\delta}}. \tag{23}$$

From the OPFA model in Eq. 23 and by using Eq. 22, one can easily obtain that

$$\Sigma(\Delta\tilde{\boldsymbol{n}}) = L\Sigma(\boldsymbol{\delta})L' = LL'. \tag{24}$$

Let $\Sigma(\Delta\tilde{\boldsymbol{n}})$ have eigenvalue-eigenvector pairs $(\lambda_i, \boldsymbol{e}_i)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Then the eigen-decomposition of $\Sigma(\Delta\tilde{\boldsymbol{n}})$ is given by

$$\begin{aligned} \Sigma(\Delta\tilde{\boldsymbol{n}}) &= \lambda_1\boldsymbol{e}_1\boldsymbol{e}_1' + \lambda_2\boldsymbol{e}_2\boldsymbol{e}_2' + \cdots + \lambda_n\boldsymbol{e}_n\boldsymbol{e}_n' \\ &= \left[\sqrt{\lambda_1}\boldsymbol{e}_1 \ \sqrt{\lambda_2}\boldsymbol{e}_2 \ \cdots \ \sqrt{\lambda_n}\boldsymbol{e}_n\right] \begin{bmatrix} \sqrt{\lambda_1}\boldsymbol{e}_1 \\ \sqrt{\lambda_2}\boldsymbol{e}_2 \\ \vdots \\ \sqrt{\lambda_n}\boldsymbol{e}_n \end{bmatrix} \end{aligned} \tag{25}$$

So if the loading matrix equal is equal to $L = [\sqrt{\lambda_1}\boldsymbol{e}_1 \ \cdots \sqrt{\lambda_n}\boldsymbol{e}_n]$, then we can obtain $\Sigma(\Delta\tilde{\boldsymbol{n}}) = LL'$ as in Eq. 24.

However, in this case, the principle factor analysis is not particularly useful since it employs as many common factors as there are random variables and does not lead to any approximation of $\Sigma(\Delta\tilde{\boldsymbol{n}})$, although the correlative relationships among $\Delta\tilde{\boldsymbol{n}}$ have been decoupled. We prefer models that explain the variance-covariance matrix $\Sigma(\Delta\tilde{\boldsymbol{n}})$ in terms of just a few common factors.

When the last $(n-k)$ eigenvalues are small, one can neglect the contribution of $\lambda_{k+1}\boldsymbol{e}_{k+1}\boldsymbol{e}_{k+1}' + \cdots + \lambda_n\boldsymbol{e}_n\boldsymbol{e}_n'$ to $\Sigma(\Delta\tilde{\boldsymbol{n}})$ in Eq. 25. So if one let

$$L = [\sqrt{\lambda_1}\boldsymbol{e}_1 \ \sqrt{\lambda_2}\boldsymbol{e}_2 \ \cdots \sqrt{\lambda_k}\boldsymbol{e}_k], \tag{26}$$

then neglecting this contribution leads to the approximation

$$\Sigma(\Delta\tilde{\boldsymbol{n}}) \approx \lambda_1\boldsymbol{e}_1\boldsymbol{e}_1' + \lambda_2\boldsymbol{e}_2\boldsymbol{e}_2' + \cdots + \lambda_k\boldsymbol{e}_k\boldsymbol{e}_k' = LL'. \tag{27}$$

Furthermore, OPFA provides a easy way to determine how many number of common factors are necessary to achieve the user specified accuracy. Since the ith factor basically corresponds to the ith eigenvalue as shown in Eq. 25 and $\sum_{i=1}^{n} \lambda_i = tr(\Sigma(\Delta\tilde{\boldsymbol{n}}))$, the contribution of the ith factor to $\Sigma(\Delta\tilde{\boldsymbol{n}})$ can then be estimated by

$$c_i = \begin{cases} \frac{\lambda_i}{tr(\Sigma(\Delta\tilde{\boldsymbol{n}}))} & \text{factor analysis using } \Sigma(\Delta\tilde{\boldsymbol{n}}) \\ \frac{\lambda_i}{n} & \text{factor analysis using } \Gamma(\Delta\tilde{\boldsymbol{n}}) \end{cases}. \tag{28}$$

So if $\sum_{i=1}^{k} c_i$ of the first k largest eigenvalues is larger than a user specified value depending on accuracy requirement, the result k number of factors will be applied to approximate $\Delta\tilde{\boldsymbol{n}}$.

D. Potential Coefficient Approximation

For one pair of panels k and l without process variations, the potential coefficient between them is evaluated by Eq. 2. If panels k and l have variations Δn_k and Δn_l along their normal direction, then the variation potential coefficient $\bar{P}_{kl}$ is a function of Δn_k and Δn_l, $\bar{P}_{kl} = f(x_k, x_l, \Delta n_k, \Delta n_l)$. By expanding $\bar{P}_{kl}$ into Taylor series around Δn_k and Δn_l, one can obtain that

$$\bar{P}_{kl} = P_{kl} + \hat{\boldsymbol{a}}_{kl}\Delta\hat{\boldsymbol{n}} + \Delta\hat{\boldsymbol{n}}'\hat{A}_{kl}\Delta\hat{\boldsymbol{n}} + h.o.t, \tag{29}$$

where $\hat{\boldsymbol{a}}_{kl}$ is a 1×2 vector and $\hat{A}_{kl}$ is a 2×2 matrix. $\Delta\hat{\boldsymbol{n}} = [\Delta n_k \ \Delta n_l]^T$ is a random vector containing Δn_k and Δn_l.

During the hierarchical panel refinement process, the recorded links may or may not be created between two leaf panels as we have shown in Fig. 2. So $\Delta\hat{\boldsymbol{n}}$ could contain the variations on some non-leaf panels. Since our process variations and principle factor analysis are performed in terms of variations on leaf panels, it is necessary to represent $\Delta\hat{\boldsymbol{n}}$ in terms of $\Delta\tilde{\boldsymbol{n}}$.

Without loss of generality, we assume that the position variations of leaf panels are along their normal direction. Then if two panels i and j have variations Δn_i and Δn_j, the variation on their parent panel k will be $\Delta n_k = 1/2(\Delta n_i + \Delta n_j)$. So all panel variations $\Delta\boldsymbol{n}$, either leaf or non-leaf, can be expressed in terms of variations on its underlying leaf panels

$$\Delta\boldsymbol{n} \quad = \quad R\Delta\tilde{\boldsymbol{n}}, \tag{30}$$

where $R \in \Re^{N \times n}$ is a provable sparse matrix. For example, for the small tree structure shown on the right hand side in Fig. 5, panels 1, 2, and 4 are leaf panels. Panel 3 is the parent of panels 1 and 2, and hence $\Delta n_3 = 1/2(\Delta n_1 + \Delta n_2)$. Panel 5 is the parent of panels 3 and 4, and hence $\Delta n_5 = 1/2(\Delta n_3 + \Delta n_4) = 1/4(\Delta n_1 + \Delta n_2) + 1/2\Delta n_4$. The detailed algorithm for constructing the random variable transformation matrix R is presented in Fig. 6.

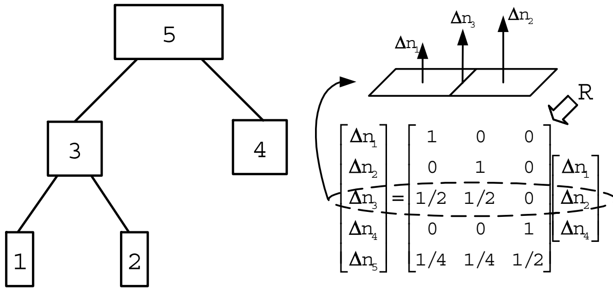


Fig. 5. Random variable transformation.

Procedure	*ConstructR*
Input:	(a) Vector $Panel$ contains the indexes of all panels;
	(b) Vector $Basis$ contains the indexes of leaf panels.
Output:	$R \in \Re^{N \times n}$, such that $\Delta\boldsymbol{n} = R \times \Delta\tilde{\boldsymbol{n}}$.

1: $n = Basis.size()$;
2: **for** $i = 1 \cdots n$ **do**
3: $\quad X = Basis[i]$;
4: $\quad InsertEntry(R, X, i, 1)$;
5: $\quad value = 1/2$;
6: $\quad$ **while** $Panel[X].parent! = NULL$ **do**
7: $\quad\quad X = Panel[X].GetParent()$;
8: $\quad\quad InsertEntry(R, X, i, value)$;
9: $\quad\quad value = 1/2 \times value$;
10: $\quad$ **end while**
11: **end for**

Fig. 6. Algorithm for constructing random variable transformation matrix. Function $InsertEntry(R, i, j, value)$ fills $value$ into the entry (i, j) of R.

Therefore, by using Eq. 30, $\Delta\hat{\boldsymbol{n}}$ can be expressed in terms of $\Delta\tilde{\boldsymbol{n}}$ as

$$\Delta\hat{\boldsymbol{n}} \quad = \quad \begin{bmatrix} R_k \\ R_l \end{bmatrix} \Delta\tilde{\boldsymbol{n}}, \tag{31}$$

where R_k and R_l are the kth and the lth rows in the transformation matrix R. And then the variational potential coefficient between panels k and l, $\bar{P}_{kl}$, can be written as according to $\Delta\tilde{\boldsymbol{n}}$

$$\bar{P}_{kl} \quad = \quad P_{kl} + \tilde{\boldsymbol{a}}_{kl}\Delta\tilde{\boldsymbol{n}} + (\Delta\tilde{\boldsymbol{n}})'\tilde{A}_{kl}\Delta\tilde{\boldsymbol{n}}, \tag{32}$$

where

$$\tilde{\boldsymbol{a}}_{kl} \quad = \quad \hat{\boldsymbol{a}}_{kl} \begin{bmatrix} R_k \\ R_l \end{bmatrix}, \tag{33}$$

and

$$\tilde{A}_{kl} \quad = \quad \begin{bmatrix} R_k \\ R_l \end{bmatrix}' \hat{A}_{kl} \begin{bmatrix} R_k \\ R_l \end{bmatrix}. \tag{34}$$

Furthermore, since the leaf panel variations can be represented using k common factors as shown in Eq. 23, the variational potential coefficient between panels k and l, $\bar{P}_{kl}$, can be further represented in terms of the k dominant common factors

$$\bar{P}_{kl} \quad = \quad P_{kl} + \boldsymbol{a}_{kl}\boldsymbol{\delta} + \boldsymbol{\delta}'A_{kl}\boldsymbol{\delta}, \tag{35}$$

where

$$\boldsymbol{a}_{kl} \quad = \quad \tilde{\boldsymbol{a}}_{kl}L, \tag{36}$$

$$A_{kl} \quad = \quad L'\tilde{A}_{kl}L. \tag{37}$$

The representation in Eq. 35 is exactly equivalent to the one in Eq. 6. The ith element in the vector $\boldsymbol{a}_{kl}$ is equal to ΔP^i_{kl} while ΔP^{ij}_{kl} is equal to $2(A_{kl})_{ij}$ if $i \neq j$ and $(A_{kl})_{ij}$ if $i = j$. So the method presented in section 3.1 can be used to solve $\Delta\boldsymbol{q}$.

E. Distribution of Parasitic Capacitance

After obtaining the quadratic expression of parasitic capacitance, Monte Carlo simulation can be applied to determine the corresponding probability density distribution (PDF). However, in this section, we will present a way to directly calculate the PDF of a parasitic capacitance given its quadratic form.

To compute the PDF of the parasitic capacitance, we first need to calculate its characteristic function. For a random variable X, its characteristic function is defined as

$$C_X(\xi) = E(e^{j\xi X}) = \int_{-\infty}^{+\infty} e^{j\xi x} f_X(x)dx, \tag{38}$$

where $f_X(x)$ is the probability density function (PDF) of X.

Since the characteristic function is actually an inverse fourier transform of the PDF, the PDF of the random variable X can easily computed if its characteristic function is known

$$f_X(x) \quad = \quad \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-j\xi x} C_X(\xi)d\xi. \tag{39}$$

The formal proof of this conclusion can be found in [19].

For a parasitic capacitance defined in the quadratic form

$$\bar{C} \quad = \quad C + \boldsymbol{a}\boldsymbol{\delta} + \boldsymbol{\delta}' A \boldsymbol{\delta}, \qquad (40)$$

where $\boldsymbol{\delta} \sim N(\mathbf{0}, \Sigma)$, its exact characteristic function can be analytically computed by [20]

$$C_{\bar{C}}(\xi) = |\Omega|^{-\frac{1}{2}} exp\{j\xi m - \frac{1}{2}\xi^2 \boldsymbol{a}' \Sigma^{\frac{1}{2}} \Omega^{-1} \Sigma^{\frac{1}{2}} \boldsymbol{a}\}, \qquad (41)$$

where $|\Omega|$ is the determinant of matrix $\Omega = I - 2j\xi\Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}}$. Once we obtain $C_{\bar{C}}(\xi)$, the PDF, and then the cumulative distribution function (CDF), can be computed from Eq. 39.

Clearly, there will be one step of eigenvalue decomposition (computing $\Sigma^{\frac{1}{2}}$) and one step of fourier transformation in order to analytically compute the distribution of a parasitic capacitance. Since our principle factor vector is $\boldsymbol{\delta} \sim N(\mathbf{0}, I)$, the Ω matrix can be simplified to $\Omega = I - 2j\xi A$. So that $C_{\bar{C}}(\xi) = |\Omega|^{-\frac{1}{2}} exp\{j\xi m - \frac{1}{2}\xi^2 \boldsymbol{a}' \Omega^{-1} \boldsymbol{a}\}$ and the eigenvalue decomposition can be eliminated.

IV. EXPERIMENTAL RESULTS

The proposed capacitance variability modeling approach has been implemented in C/C++ language. All experiments are executed on a Pentium(R) 4 CPU 1.4GHz machine with 1GB RAM. Monte Carlo simulation with $10,000$ runs is used for comparison purpose.

First, for the 2×2 bus crossing problem, probability density functions (PDF) obtained from the canonical linear model and the quadratic model are shown in Fig. 7 and compared with that from Monte Carlo simulation. It is illustrated that there is a significant accuracy improvement by using the second order quadratic model instead of the canonical model. The accuracy improvement of the quadratic model is mostly due to the probability distribution region corresponding to larger capacitance values, which is actually more critical for circuit performance and timing analysis. The canonical model will tend to underestimate the possible capacitance value in the high probability region. This underestimation, in reality, will result in optimistic design and excessive chip failure. This example clearly shows the necessity of the quadratic model in today's technology where process variation can no longer be ignored..

In the second experiment, the CDFs and PDFs of the second order quadratic models with different number of dominant factors are compared. Without applying PFA, the number of random variables is equal to the total number of leaf panel, which is 1126 for bus 2×2. In practice, how many number of dominant factors need to be preserved is determined by the Gaussian correlation length in Eq. 19. The setup of Gaussian correlation length depends on the detailed processing techniques and the local layout characteristics. For different regions and different panel orientations, we can assign different correlation lengths. In this test, PFA with only ten factors is very close to the result CDF and PDF from Monte Carlo simulation, so that ninety percent random variable reduction has been achieved by PFA. And in this case, the error in CDF compared with Monte Carlo is less than 3%. Furthermore, as

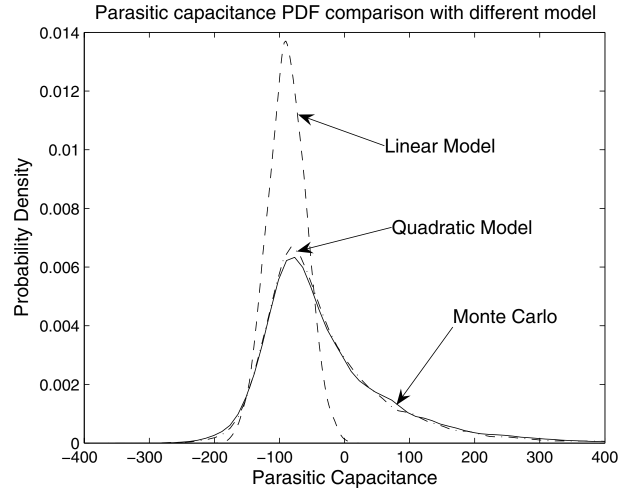


Fig. 7. First and second order capacitance models and their comparisons with Monte Carlo method for the bus 2×2 benchmark ($\sigma = 20\%$).

the number of factors increases, the CDFs and PDFs from the quadratic models quickly converge to those from Monte Carlo simulation.

In table IV, the run times of Monte Carlo method and the quadratic model with 10 dominant factors for different bus crossing benchmarks are compared. It is clear that the quadratic model exhibits over $100\times$ speedup compared with Monte Carlo simulation. Statistical distribution-related parameters, such as mean value, standard deviation, and skewness are normally within 3% errors. Combined with the results from previous experiments, We can safely conclude that, currently, the second order approximation is accurate enough for variational parasitic capacitance modeling.

2 × 2 Bus

Method	Time	Mean μ	Std Variation σ	Skewness η
M.C.	1826	-78.56	106.01	1.868
QuadMod	9.78	-81.43	103.64	1.927
Speedup/Err	186.7×	3.7%	2.2%	3.2%

4 × 4 Bus

Method	Time	Mean μ	Std Variation σ	Skewness η
M.C.	4673	-194.89	85.62	-1.78
QuadMod	16.88	-192.45	83.78	-1.72
Speedup/Err	276.8×	1.3%	2.1%	3.4%

6 × 6 Bus

Method	Time	Mean μ	Std Variation σ	Skewness η
M.C.	8568	-195.49	89.34	-1.42
QuadMod	69.56	-190.71	85.52	-1.37
Speedup/Err	123.2×	2.4%	4.3%	3.5%

TABLE I

SIMULATION RUNTIME COMPARISON FOR BUS CROSSING BENCHMARK. (1) MONTE CARLO (M.C.); (2) QUADRATIC MODEL (QUADMOD).

V. CONCLUSION

This paper presents an efficient methodology for generating explicit statistical representations of parasitic capacitances. Our method applies principle factor analysis to reduce the

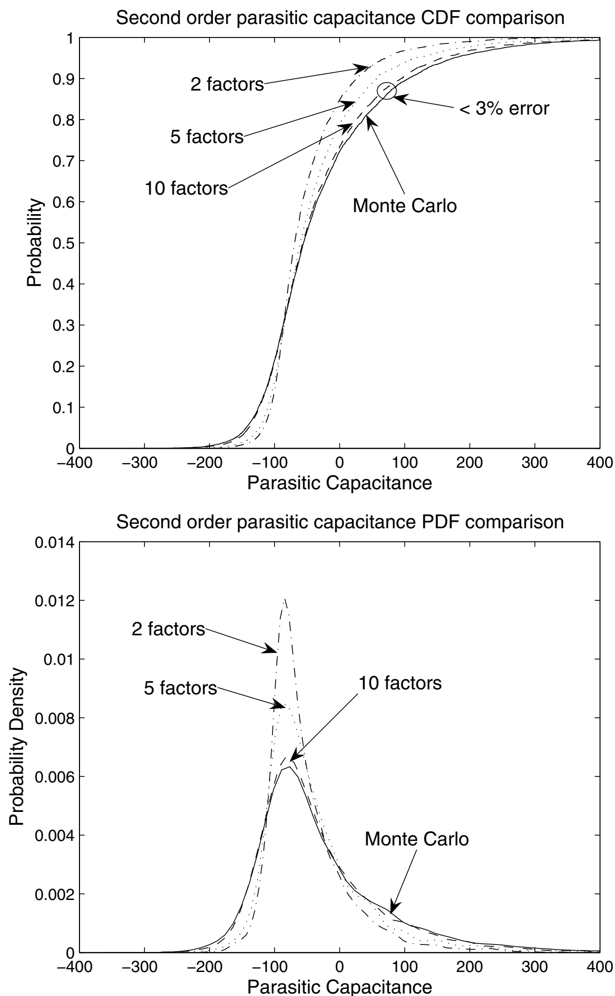


Fig. 8. Second order parasitic capacitance modeling with different number of factors and the comparison with Monte Carlo method for bus 2×2 benchmark.

number of random variables while preserving the dominant global/local factors that induce the conductor surface fluctuation due to process variations. The obtained quadratic form can not only be used to directly generate parasitic capacitance probability distribution to locate design corners, but it is also fully compatible with statistical model order reduction and statistical timing analysis tools.

ACKNOWLEDGMENT

This work was partially funded by Intel, TSMC, UMC, Faraday, SpringSoft, National Science Foundation under grants CCR-0093309 & CCR-0204468 and National Science Council of Taiwan, R.O.C. under grant NSC 92-2218-E-002-030.

REFERENCES

[1] P. Wrschka, J. Hernandez, G. S. Oehrlein, and J. King, "Chemical mechanical planarization of copper damascene structures," *Journal of The Electrochemical Society*, pp. 706–712, 2000.

[2] P. Li, F. Liu, X. Li, L. T. Pileggi, and S. R. Nassif, "Modeling interconnect variability using efficient parametric model order reduction," *Design Automation and Test in Europe*, pp. 958–963, 2005.

[3] E. Chiprout and T. Nguyen, "Survey of model reduction techniques for analysis of package and interconnect models of high-speed designs," *IEEE 6th Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 251–254, 1997.

[4] B. N. Sheehan, "Branch merge reduction of rlcm networks," *Proc. ICCAD*, pp. 658–664, 2003.

[5] H. Chang and S. S. Sapatnekar, "Statistical timing analysis considering spatial correlations using a single pert-like traversal," *Proc. ICCAD*, pp. 621–625, 2003.

[6] L. Daniel, C. S. Ong, S. C. Low, K. H. Lee, and J. White, "A multiparameter moment matching model reduction approach for generating geometrically parameterized interconnect performance models," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 23, no. 5, pp. 678–693, May 2004.

[7] X. Li, J. Le, P. Gopalakrishnan, and L. T. Pileggi, "Asymptotic probability extraction for non-normal distributions of circuit," pp. 2–9, 2004.

[8] K. Nabors and J. White, "Fastcap: a multipole accelerated 3-d capacitance extraction program," *IEEE Trans. on CAD*, pp. 1447–1459, 1991.

[9] W. Shi, J. Liu, N. Kakani, and T. Yu, "A fast hierarchical algorithm for 3-d capacitance extraction," *IEEE Trans. on CAD*, pp. 330–336, 2002.

[10] S. Yan, V. Sarin, and W. Shi, "Sparse transformations and preconditioners for hierarchical 3-d capacitance extraction with multiple dielectrics," *Proc. DAC*, pp. 788–793, 2004.

[11] M. Beattie and L. Pileggi, "Electromagnetic parasitic extraction via a multipole method with hierarchical refinement," *Proc. ICCAD*, pp. 437–444, 1999.

[12] S. Kapur and D. E. Long, "Ies3: A fast integral equation solver for efficient 3-dimensional extraction," *Proc. ICCAD*, pp. 448–455, 1997.

[13] B. Krauter, X. Yu, A. Dengi, and L. Pileggi, "A sparse image method for bem capacitance extraction," *Proc. DAC*, pp. 357–362, June 1996.

[14] J. Tausch and J. White, "A multiscale method for fast capacitance extraction," *Proc. DAC*, pp. 537–542, 1999.

[15] J. R. Phillips and J. White, "A precorrected fft method for capacitance extraction of complicated 3-d structures," *IEEE Trans. CAD*, pp. 1059–1072, 1997.

[16] Z. Zhu, A. Demir, and J. White, "A stochastic integral equation method for modeling the rough surface effect on interconnect capacitance," *Proc. ICCAD*, pp. 887–891, 2004.

[17] R. Jiang, Y.-H. Chang, and C. C.-P. Chen, "Iccap: A linear time sparse transformation and reordering algorithm for 3d bem capacitance extraction," *Proc. DAC*, pp. 163–166, 2005.

[18] R. L. Gorsuch, *Factor Analysis*. Hillsdale, NJ, 1974.

[19] B. V. Gnedenko, *Theory of Probability*. Gordon and Breach Science Publishers, 1997.

[20] A. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*. New York Marcel Dekker, 1992.