### Stochastic dominant singular vectors method for variation-aware extraction

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1145/1837274.1837444">http://dx.doi.org/10.1145/1837274.1837444</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Association for Computing Machinery (ACM)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Tue Feb 27 05:04:16 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/72204">http://hdl.handle.net/1721.1/72204</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
ABSTRACT

In this paper we present an efficient algorithm for variation-aware interconnect extraction. The problem we are addressing can be formulated mathematically as the solution of linear systems with matrix coefficients that are dependent on a set of random variables. Our algorithm is based on representing the solution vector as a summation of terms. Each term is a product of an unknown vector in the deterministic space and an unknown direction in the stochastic space. We then formulate a simple nonlinear optimization problem which uncovers sequentially the most relevant directions in the combined deterministic-stochastic space. The complexity of our algorithm scales with the sum (rather than the product) of the sizes of the deterministic and stochastic spaces, hence it is orders of magnitude more efficient than many of the available state of the art techniques. Finally, we validate our algorithm on a variety of on-chip and off-chip capacitance and inductance extraction problems, ranging from moderate to very large size, not feasible using any of the available state of the art techniques.

1. INTRODUCTION

Advanced integrated circuit manufacturing steps such as etching, chemical mechanical polishing (CMP), electro-deposition, and photolithography produce large variabilities in the geometries of the manufactured interconnect structures. Such geometrical variabilities are not deterministic and are beginning to play a major role in determining the electrical characteristics of the interconnect structures. The development of efficient variation aware extraction algorithms is a critical step to promptly and accurately account for the effects of such variabilities. Recently, a few attempts have been made towards developing such solvers [1, 2, 3].

Variation-aware solvers can be divided in general into two different categories, namely, those based on sampling algorithms (sometimes identified as non-intrusive) and those based on a stochastic formulation (sometimes identified as intrusive). All such algorithms are better understood if we abstract the problem of extracting the input impedance of interconnect structures with a large number of random variations to one of solving a large linear system of the form \( A(p)x(p) = b(p) \), where \( p \) is a large vector of random variables. The “non-intrusive” algorithms are basically those algorithms that rely on sampling the parameter space vector \( p \), and then solving a deterministic system \( A(p_k)x(p_k) = b(p_k) \) for every sample \( p_k \) of the parameter space. The term “non-intrusive” refers to the fact that any standard deterministic solver can be used to compute the unknown vector \( x(p_k) \) in each sample. Examples of non-intrusive algorithms include the well-known Monte Carlo method, and the stochastic collocation method [4]. The computational complexity of standard sampling based methods is directly proportional to \( N_S \), the number of sample points in the parameter space \( O(N_S N^2) \). The main disadvantage of such methods is that generally a very large number of sample points \( N_S \) is required in order to accurately sample a high dimensional parameter space. In order to address this issue, the approach in [5] computes a small number of dominant basis vectors which span the solution space. The dominant basis vectors represent a reduced system and are constructed sequentially. At each step the reduced system is solved at a new point from a given set of quadrature points for the statistical moments of the output. The resulting residual is then used to determine if the current basis needs to be extended. The complexity of such algorithm is \( O(N_D N^2) \), where \( N_D \) is the total number of dominant basis vectors, hence \( N_D \ll N_S \).

On the other hand, “intrusive” algorithms solve stochastic linear systems using specialized techniques. Examples of intrusive algorithms are the Neumann expansion method [1, 6], the stochastic finite element method (SFE) [7], and the combined Neumann-Hermite expansion method (CNHE) [3]. Unfortunately, as for non-intrusive methods, also intrusive methods can be quite expensive. For instance, the complexity of the Neumann expansion method is \( O(N^4) \), the complexity of the stochastic Galerkin method is \( O(N_D^2 N^2) \) and the complexity of the combined Neumann-Hermite expansion is \( O(N_D^2 N^2) \).

In this paper we present a novel “intrusive” algorithm to obtain sequentially a small number of dominant basis vectors, spanning the space of the solution, as in the “non-intrusive” algorithm [5]. In addition to being “intrusive”, there are several fundamental differences distinguishing the algorithm in this paper from the one in [5]. Instead of constructing the dominant basis vectors from a given set of quadrature points, in this paper the basis vectors are constructed...
using a global optimization based approach, resulting in an optimal selection and in a global error bound. While in [5] the dominant basis vectors span only a deterministic space, in this paper we represent the solution using two sets of basis vectors, one spanning the deterministic space and the other spanning the stochastic space. Such representation allows computing directly an efficient reduced expansion for the unknown solution, rather than just statistical moments. The intrusive algorithm in this paper has an unprecedented low complexity $O(N^2 + N_p^3)$. Since even in worst case applications $N_p$ is typically several orders of magnitude smaller than $N$, from a practical point of view the complexity is basically just $O(N^2)$, making our algorithm independent of the dimension of the parameter space. The details of the basic algorithm are given in section 3. A relaxation is then given in 4. Section 5 studies convergence properties. Section 6 presents an accelerated technique for computing the product of a stochastic matrix and a stochastic vector. Section 7 shows a detailed analysis of the computational complexity of the overall algorithm. Finally, in Section 8 we show the accuracy and speed of our algorithm on a variety of examples.

2. BACKGROUND

2.1 Impedance Extraction

In a general impedance extraction setting, we can use the standard mixed potential integral equation [8] to describe the relation between the volumetric current density $j(r)$ inside of the conductors, the surface charge distribution $p(r)$ on the surface of the conductors, and the electric potential $\phi(r)$:

\[
\frac{j(r)}{\sigma} + j_0 \frac{\mu}{4\pi} \int G(r,r')j(r')dr' = -\nabla \phi(r) \quad (1)
\]

\[
\int S G(r,r')p(r')dr' = \phi(r) \quad (2)
\]

\[
V \cdot j(r) = 0 \quad (3)
\]

\[
\hat{n} \cdot j(r) = j_0 p(r), \quad (4)
\]

where $G(r,r')$ is the Green’s function, $V$, $S$ are the conductors volume and surface area, respectively, $\sigma$ is the conductor conductivity, $\epsilon$ is the complex dielectric constant including dielectric losses, $\mu$ is the magnetic permeability, and $\omega$ is the angular frequency in radians. The special case of the standard electro- quasi-static (EQS) capacitance extraction problem is described by imposing just equation (2). The special case of the standard magneto-quasi-static (MQS) resistance and inductance extraction problem is described by imposing just equations (1) and (3).

A standard procedure for solving (1)-(4), or the corresponding EQS and MQS subsets, involves discretizing the current density $j(r)$ and the charge density $p(r)$ with piecewise constant basis functions and using Galerkin testing to obtain a linear system of equations of the form:

\[
\tilde{A}(\tilde{p}) \cdot \tilde{x}(\tilde{p}) = \tilde{b}(\tilde{p}) \quad (5)
\]

where $\tilde{A}(\tilde{p})$, $\tilde{x}(\tilde{p})$ and $\tilde{b}(\tilde{p})$ are described in detail in [9] for EQS, in [10] for MQS, and in [8] for EMQS and fullwave; $\tilde{p}$ is a vector of random geometrical variations described by a multivariate Gaussian probability density function with a Gaussian or exponential correlation function. Examples of random geometrical variations are width and thickness variations in on-chip interconnects or rough surface in on-package or on-board interconnects.

Following the general practice, the Karhunen-Loeve expansion (or in other words a truncated Singular Value Decomposition) [11] is used to transform the correlated physical parameters $\tilde{p}$ into independent Gaussian random variables $p$, such that $\tilde{p} = \sqrt{C}p$, where $C$ is the correlation matrix associated with the multivariate PDF. Substituting $\tilde{p} = \sqrt{C}p$ into (5) one can obtain the general statement of the problem

\[
A(p)x(p) = b(p) \quad (6)
\]

where $A(p) = \tilde{A}(\sqrt{C}p)$, $x(p) = \tilde{x}(\sqrt{C}p)$, $b(p) = \tilde{b}(\sqrt{C}p)$, and as mentioned previously $p$ is a vector of $N_p$ independent Gaussian random variables.

2.2 Hermite Expansion

The Hermite expansion of a given scalar function $f(p)$ is

\[
f(p) = \sum_{i=1}^{K} f_i H_i(p) \quad (7)
\]

where $H_i(p)$ is a set of multivariate Hermite orthogonal polynomials [7] and $K$ is a function of both $N_p$, the order of the orthogonal polynomial expansion, and $N_p$, the dimension of the parameter vector $p$ [12]:

\[
K = 1 + \sum_{n=1}^{N_p} \left( n + N_p - 1 \right) \quad (8)
\]

Consequently, for a second order expansion $N_p = 2$, the dependence of $K$ on $N_p$ is quadratic $K = O(N_p^2)$. The coefficients of the expansion $f_i$ can be efficiently computed as suggested in [3]. If the Hermite expansion is applied to a vector or a matrix then, it is understood that, it is applied element-wise.

2.3 Stochastic Inner Product and Norm

We define the stochastic inner product between two vectors $u(p)$ and $v(p)$ to be

\[
\langle u(p), v(p) \rangle = \mathbb{E} \left[ u(p)^T v(p) \right] = \int_{p} u(p)^T v(p) \exp\left(-\frac{0.5p^T Cp}{(2\pi)^{N_p/2} SV_p} \right) dp \quad (9)
\]

where $\mathbb{E}[\cdot]$ is the expectation operator. Consequently, the stochastic norm of a vector $u(p)$ is $\|u(p)\|_S = \sqrt{\mathbb{E}[\|u(p)\|^2]}$.

2.4 Representation of Parameter Dependence

The unknown $x(p)$ lives in the $N$ dimensional stochastic space. One way to describe the dependence of the vector $x(p)$ on the parameters is to expand $x(p)$ in terms of multivariate orthogonal polynomials spanning the parameter space:

\[
x(p) = \sum_{i=1}^{K} x_i H_i(p) \quad (10)
\]

This last expansion can be expressed in the more compact form

\[
x(p) = Xp \quad (11)
\]

where $X$ is a deterministic matrix of size $N \times K$ with columns $x_i$ and $p = (H_0(p), \ldots, H_K(p))^T$ is a vector of $K$ orthogonal polynomials of the stochastic variable $p$. This decomposition represents the unknown vector as a product of two different spaces, namely, the deterministic space, and the stochastic space. The idea behind both the standard stochastic collocation and Galerkin methods is to fix the deterministic component $h(p)$, and solve only for the deterministic component $X$.

In particular in the non-intrusive stochastic collocation method (SCM) [4], the parameter space is sampled at a point using the delta
function and the solution is found at that particular point. The challenge with the SCM is that a very large number of sample points is needed in order to accurately sample the parameter space. Consequently, solving the linear system at every sample point in the parameter space is computationally very expensive $O(N^2N_p^2)$.

On the other hand, in the intrusive stochastic Galerkin Method (SGM), also known as stochastic finite element (SFE) [7], the solution matrix $X$ is found such that the residual is orthogonal to the stochastic space spanned by $h(p)$. The SGM is computationally very expensive since solving simultaneously for the entire unknown matrix $X$ is $O(N^2N_p^2)$, i.e., even more expensive than the SCM.

3. STOCHASTIC DOMINANT SINGULAR VECTORS METHOD

To avoid the complexities associated with both the SCM and the SGM we suggest the use of a nonlinear representation of $x(p)$, where both the deterministic and stochastic components are assumed unknown. To better understand the idea behind our work let us consider expressing $x(p)$ in terms of its dominant basis by expressing $X$ in terms of its singular value decomposition (SVD)

$$X = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

where $\sigma_i v_i = v_i$ and $r$ is the total number of dominant basis. Note that $v_i^T h(p)$ is a scalar polynomial and that (10) can therefore be expressed as

$$x(p) = \sum_{i=1}^{r} v_i^T h(p) u_i$$

where $u_i$ is an unknown vector of length $N$ and $v_i$ is an unknown vector of length $K$ representing a direction in the stochastic space. It should be noted that the representation of the unknown in (11) has been previously and independently suggested in [13]. Our algorithm (presented in the rest of the paper) is however significantly different than that in [13].

The Key Idea. We propose to decompose the solution vector $x[p]$ in the form of the summation (11) and find its components sequentially, i.e., at every iteration $n$ of our algorithm we solve only for $u_n$ and $v_n$. These vectors are computed such that they minimize the norm of the residual at iteration $n$.

To achieve our objective we first substitute (11) in (6) to obtain

$$A(p) \sum_{i=1}^{r} v_i^T h(p) u_i = b(p)$$

Assume that at step $n$ of our algorithm we know all the vectors $(u_i, v_i : i = 1, \dots, n - 1)$ and define

$$x_n(p) = \sum_{i=1}^{n} v_i^T h(p) u_i = x_{n-1}(p) + v_n^T h(p) u_n$$

and

$$r_n(p) = b(p) - A(p)x_n(p)$$

Equation (14) can be put in a recursive form by using (13)

$$r_n(p) = b(p) - A(p)x_{n-1}(p) - A(p)v_n^T h(p) u_n$$

$$= r_{n-1}(p) - A(p)v_n^T h(p) u_n$$

where $x_0(p) = 0$ and $r_0(p) = b$.

As mentioned above, at step $n$ of our algorithm we find $u_n$ and $v_n$ such that the stochastic norm of the residual $r_n(p)$ is minimized

$$\min_{u_n, v_n} ||r_n(p)||^2$$

where the objective function is

$$f = ||r_n(p)||^2 = ||r_{n-1}(p)||^2 - 2E[u_n^T A(p)^T v_n^T h(p) u_{n-1}(p)] + E[u_n^T A(p)^T v_n^T h(p) A(p) u_{n-1}(p)]$$

One approach to minimizing (16) is to set the gradient $f'$ of the objective function $f$ to zero, which can be achieved using Newton’s method, solving at each iteration the linear system

$$f' = -f''dz$$

$f''$ is the Hessian of the objective function. Detailed expressions for $f'$ and $f''$ are given in Appendix A. Notice that (16) can also be efficiently solved using a fixed point iteration.

The number of the free optimization parameters (length of both $u_n$ and $v_n$) is $O(N + K)$, which means that at every iteration we only need to solve a linear system of size $O(N + K)$ for a complexity $O(N + K)^2$. Such minimization is performed only $r$ times, where $r$ is the number of dominant basis of $x(p)$. Note, typically $r \ll K \approx N$. Consequently, the complexity of our algorithm scales with just $O(N^2)$, practically independent of the number of parameters. A more detailed complexity analysis will follow in Section 7.

The norm of the residual is used as an indicator of the accuracy of the solution. In other words, we keep looking sequentially for dominant basis until the norm of the residual becomes smaller than a given threshold. Algorithm 1 presents a complete summary of our proposed algorithm. Notice that both unknowns $u_n$ and $v_n$ are combined into a single vector $z$.

Algorithm 1 Stochastic Dominant Singular Vectors Method (SDSV)

1: $x(p) \leftarrow 0$, $r(p) \leftarrow h(p)$
2: $u_0 \leftarrow x_0$, the solution of the nominal problem
3: $v_0 \leftarrow e_1$
4: $z \leftarrow [u_0^T, v_0^T]^T$
5: while $||r(p)||^2 > \text{Threshold}$ do
6:   repeat
7:     form first derivative $f'$ as in (22)
8:     form Hessian $f''$ as in (23)
9:     solve linear system $f''dz = -f'$
10:   repeat
11:     $z \leftarrow z + dz$
12:     $u_n \leftarrow z(1:N, 1)$, $v_n \leftarrow z(N + 1 : N + K, 1)$
13:   until $||f'|| < \text{Threshold}$
14:   $x(p) \leftarrow x(p) + x_n v_n^T h(p)$
15:   $r(p) \leftarrow r(p) - A(p) u_n v_n^T h(p)$
16:   $u_n \leftarrow x_0$
17:   $v_n \leftarrow e_1$
18: $z \leftarrow [u_n^T, v_n^T]^T$
19: end while

For efficient implementation we use the Hermite expansion to express both the system matrix $A(p) = \sum_{i=1}^{K} A_i H_i(p)$ and the residual vector $r_n(p) = R_n h(p)$. The implementation details are omitted, however, one example of the resulting expressions can be found in Appendix A.
4. RELAXING THE OPTIMIZATION

Consider a large dimensional problem, e.g. one in which the original system size is \( N = 20,000 \) and the size of uncorrelated parameters \( N_P = 300 \) (the total number of orthogonal polynomials for a second order Hermite expansion \( K \leq 50,000 \)). The size of the linear system to be solved is \( O(N + K) \approx O(70,000) \), which might become prohibitive. To manage such large size problems, we propose the following relaxation. First, the index \( k_{\text{max}} \) corresponding to the vector coefficient of the residual Hermite expansion with the largest norm is determined

\[
k_{\text{max}} = \arg \max_k \left| \left| \mathbf{r}_{\text{res}}(k) \right| \right|
\]

where \( \mathbf{r}_{\text{res}}(k) = \mathbf{r}_{\text{res}}(1:k) \). Then the indices of the elements of \( \mathbf{v}_n \) that have a contribution in the \( k_{\text{max}} \)-direction is determined. In other words, we find all the indices \( j \) of the elements of the vector \( \mathbf{v}_n \), such that \( v_n(j) \) multiplied with the system matrix will result in some component in the \( k_{\text{max}} \)-direction:

\[
j : \langle A(p)^T \mathbf{h}(p), H_{\text{max}}(k) \rangle \neq 0.
\]

The number of such indices is \( O(N_P) \). We then solve a relaxed version of the optimization problem in which only the \( N_P \) identified components of \( \mathbf{v}_n \) are allowed to vary, while all the other components are fixed to 0. Consequently, the reduced problem is of size \( O(N + N_P) \approx O(20,300) \).

5. CONVERGENCE ANALYSIS

The following theorem summarizes the convergence properties of our algorithm.

**Theorem 5.1.** For any nonsingular matrix \( A(p) \) the sequence of the norms of the residuals \( \left| \left| r_k(p) \right| \right| \) is strictly decreasing for any non-zero update of the solution \( u_k, v_k^2 \mathbf{h}(p) \).

**Proof.** The optimal \( u_k \) and \( v_k \) occur when the derivative of the nonlinear objective function in (17) is zero

\[
0 = -\mathbb{E} \left[ A(p)^T v_k^2 \mathbf{h}(p) r_{n-1}(p) + A(p)^T A(p) u_k v_k^2 \mathbf{h}(p) v_k^2 \mathbf{h}(p) \right]
\]

Multiplying (19) by \( u_k \) and substituting the resulting expression in (17)

\[
\left| \left| r_k(p) \right| \right|_2^2 = \left| \left| r_{n-1}(p) \right| \right|_2^2 - \left| \left| A(p) u_k v_k^2 \mathbf{h}(p) \right| \right|_2^2
\]

Since \( A(p) \) is nonsingular, the vector \( A(p) u_k v_k^2 \mathbf{h}(p) \) is non-zero for any non-zero update to the solution \( u_k, v_k^2 \mathbf{h}(p) \). Consequently, \( A(p) u_k v_k^2 \mathbf{h}(p) \) is orthogonal to \( r_{n-1}(p) \). That implies that \( \left| \left| r_k(p) \right| \right|_2^2 < \left| \left| r_{n-1}(p) \right| \right|_2^2 \).

6. FAST STOCHASTIC MATRIX VECTOR PRODUCT

The computational complexity of our algorithm is mainly determined by a matrix vector product of the form \( A(p) \mathbf{u}_n \) (see (17)). The complexity of computing such a matrix vector product using the traditional Hermite expansion of the system matrix

\[
A(p) \mathbf{u}_n = \sum_{i=1}^{K} A_i h_i(p) \mathbf{u}_n = \sum_{i=1}^{K} A_i \mathbf{u}_n h_i(p) = \mathbf{T} \mathbf{p},
\]

would be \( O(N^2 K) \), since the columns of \( \mathbf{T} \) are computed successively by computing the products \( A_i \mathbf{u}_n \). Instead, we compute the rows of \( \mathbf{T} \) successively, leading to a final complexity of \( O(N^2) \). In the following, we summarize the key steps allowing such reduced complexity. We note that the \( k^{th} \) row of \( \mathbf{T} \) is formed by taking the product of the \( k^{th} \) row of each \( \mathbf{A}_i \) with \( \mathbf{u}_n \). All such rows are collected in a single matrix \( \mathbf{A}_k \in \mathbb{R}^{K \times N} \) in which the \( q^{th} \) row is the \( k^{th} \) row of matrix \( \mathbf{A}_i \), i.e. \( \mathbf{A}_k(i,:) = \mathbf{A}_i(k,:) \). The SVD of \( \mathbf{A}_k \) is formed by successively, leading to a final complexity of \( O(N^2) \).
8.1 Small On-Chip 2-D Capacitance Extraction Example

The first example is a 16-conductor on-chip 2-D structure. The conductors are arranged in a 4 × 4 grid typically used as a template for building capacitance tables. Each conductor is discretized using 30 panels, resulting in a total of 481 unknowns. We assume that the width of each conductor is an independent Gaussian random variable, resulting in \( N_P = 16 \). The total number of terms required for a second order Hermite expansion is \( K = 153 \). We set our algorithm to terminate when the norm of the residual became less than \( 10^{-4} \).

We observed that a total of 64 directions are sufficient to represent the solution to the desired accuracy. Solving this with a Matlab implementation of our algorithm required 25 sec which is a factor of 8 \times\) faster than the SMOR [5], a factor of 18 \times\) faster than the SCM [4], and a factor of 32 \times\) faster than the SGM [7].

Using the relaxed formulation in Section 4 a total of 78 directions are needed to obtained the same accuracy. In other words, the number of iterations (dominant search directions) increases by about 20\% compared to the original formulation (see Figure 1), however, the total simulation time is reduced by a factor of 2 \times\) or just 12.5 sec.

8.2 Large On-Chip 2-D Capacitance Extraction Example

The second example is a larger 2-D structure consisting of 100 conductors, arranged on a 10 × 10 grid. In this new setup \( N = 5000 \), \( N_P = 100 \) and \( K = 5151 \). A total of 142 dominant bases are required to reach a residual of norm \( 10^{-4} \). Using the computed \( x(p) \) we are able to estimate the capacitance of 99\% of a 1000 randomly generated structures with accuracy better than 1\%. The complete problem (142 search directions) is solved in the time required to solve a total of just 2000 solves. Our algorithm has about the same time complexity as the SMOR [5], which (like our algorithm) can exploit the sparsity of the coefficient matrices of the system matrix Hermite expansion. Our algorithm is 10 \times\) faster than the second order sparse grid SCM [4] which requires more than 20,000 solves for the same accuracy. The problem cannot be solved using the SGM [7] due to memory constrains (system matrix requires about 200TB).
dom variables. Using a Matlab implementation of our algorithm we simulated the structure in 4 hours, while at least 8 hours were required for the SMOR, 32 hours were required for the combined Neumann Hermite expansion (CNHE), a total of 150 hours were required for the SCM, and finally an estimated total of 828 hours would be required for the standard Neumann expansion [7]. The SGM fails in this example due to excessive memory requirements (800,000 TB). In conclusion, our algorithm is a factor of 2× faster than the SMOR, a factor of 8× faster than the CNHE, a factor of 100× faster than the standard Neumann expansion and a factor of 37× faster than the best sampling-based approach.

9. CONCLUSION

In this paper we have presented a new algorithm for solving stochastic linear systems of equations, i.e. linear systems in which the system matrix and right hand side are functions of a multivariate random vector. In our approach we express the solution in terms of its dominant bases in the combined deterministic-stochastic space. Such bases are computed sequentially, such that at every step the stochastic norm of the residual is minimized. The iteration terminates when the norm of the residual becomes smaller than a preset threshold, thereby providing control on the accuracy vs. computational time trade-off. Our algorithm has an unprecedented low threshold, thereby providing control on the accuracy vs. computational time trade-off. Our algorithm has an unprecedented low threshold, thereby providing control on the accuracy vs. computational time trade-off.

The Mathematical formulation is basically just the complexity independent of the number of parameters. Furthermore, we have presented a relaxation of the problem which significantly simplifies the complexity to $O(N^2)$, while typically only increasing by less than 20% the number of dominant bases, required for the same accuracy. The proposed algorithm has been validated on a variety of on-chip and off-chip capacitance and inductance extraction problems ($N$ up to 20,000 and $N_p$ up to 400 uncorrelated parameters). Such examples are significantly larger than any other examples available in the literature. We were able to solve such large examples in few hours on a 4GB machine. In all examples we have observed factors of $8 \times$ to $100 \times$ reduction in the computational time (for the same accuracy) compared to the best available techniques in literature, such as the stochastic collocation method, the stochastic Galerkin method, the Neumann expansion or the combined Neumann Hermite expansion.

Acknowledgment

This work has been partially supported by Mentor Graphics, AMD, IBM, the Semiconductor Corporations and the Interconnect Focus Center, one of five research centers funded under the Focus Center Research Program, a DARPA and Semiconductor Research Corporation program.

APPENDIX

A. EXPRESSIONS FOR EQUATION (18)

\[
\begin{align*}
\frac{d f}{d u_k} &= -2E [d(p) A(p)^T r_{n-1}(p)] + 2E [d(p)^T A(p) A(p) u_k] \\
\frac{d^2 f}{d u_i d u_j} &= 2E [d(p)^T A(p)^T A(p) u_i] + 2E [d(p)^T A(p)^T A(p) u_j] \\
\frac{d^2 f}{d u_i d u_j} &= 2E [u_i^T A(p)^T A(p) u_j] \\
\frac{d^2 f}{d u_i d u_j} &= -2E [h(r_{n-1})^T A(p)] + 4E [d(p) h(p) u_i^T A(p) A(p) u_j]
\end{align*}
\]

The above expressions can be efficiently expressed using the Hermite expansions of the system matrix $A(p)$ and the residual $r(p)$. As an example, the $i^{th}$ row of the second order derivative matrix $\frac{d^2 f}{d u_i d u_j}$ can be expressed as

\[\sum_{k=1}^{K} \left( -2 \sum_{j=1}^{M} R_{n-1}(j;k) A_i + \sum_{j=1}^{M} u_j^T A^T A u_i \right) v_{n}(j) \]

where we use the notation $\sum_{j=1}^{M}$ to indicate a summation over all the indices $i, j : \langle H_i H_j, H_k \rangle \neq 0$, the notation $\sum_{i=1}^{I}$ to indicate a summation over all $i : \langle H_i H_j, H_k \rangle \neq 0$, the Matlab notation $R_{n-1}(j;k)$ to indicate the $k^{th}$ column of the matrix $R_{n-1}$ and the notation $v_{n}(j)$ to indicate the $j^{th}$ element of vector $v_n$.

B. REFERENCES


