Parameterized Model Order Reduction via a Two-Directional Arnoldi Process

Yung-Ta Li[*], Zhaojun Bai[†], Yangfeng Su[‡], and Xuan Zeng[§]

[*]Dept. of Mathematics, University of California, Davis, CA 95616 Email: ytli@math.ucdavis.edu
[†]Dept. of Computer Science, and Dept. of Mathematics, University of California, Davis, CA 95616 Email: bai@cs.ucdavis.edu
[‡]School of Mathematical Sciences, Fudan University, Shanghai 200433, China Email: yfsu@fudan.edu.cn
[§]ASIC & System State Key Lab., Fudan University, Shanghai 200433, China Email: xzeng@fudan.edu.cn

Abstract—**This paper presents a multiparameter moment-matching based model order reduction technique for parameterized interconnect networks via a novel two-directional Arnoldi process. It is referred to as a PIMTAP algorithm, which stands for Parameterized Interconnect Macromodeling algorithm via a Two-directional Arnoldi Process. PIMTAP inherits the advantages of previous multiparameter moment-matching algorithms and avoids their shortfalls. It is numerically stable and adaptive, and preserves the passivity of parameterized RLC networks.**

I. INTRODUCTION

With the decrease of IC feature size and increase of signal frequency, interconnect has become a dominant factor in the whole chip performance. Interconnect networks are represented by large scale system equations with respect to frequency parameter. However, during the circuit synthesis of large scale digital or analog applications, it is also crucial to evaluate the response of interconnect as functions of other design parameters, such as geometry and temperature. Parameterized model order reduction (PMOR) methods are considered as necessary techniques. In addition, the indetermination in the manufacturing of IC chips may cause system variations of critical dimensions and inter level dielectric thicknesses of interconnects, which could make the chip performance unpredictable and cause significant parametric yield lose. Therefore, it becomes necessary to use PMOR algorithms for the analysis of interconnects in the presence of process variations [1]–[8].

A number of PMOR methods have been developed. The perturbation technique [5] is one of the early work to capture small variation around the nominal circuit values. It becomes inefficient when modeling strong nonlinear effects caused by the intra-die variations [7]. Multiparameter moment-matching methods presented in [3], [9] use a subspace projection approach and guarantee the passivity. However, the proposed computational procedures are potentially numerical unstable due to the fact that basis vectors of underlying projection subspaces are calculated explicitly. The method in [10], [11] uses a multiseries expansion but is difficult to generalize to several parameters. The CORE algorithm [7] is numerically stable. Unfortunately, it does not provide the reduced system in state equation form and does not preserve the passivity.

In this paper, we pursue the multiparameter moment-matching approach and present a new algorithm, referred to as PIMTAP (Parameterized Interconnect Macromodeling via a Two-directional Arnoldi Process). The PIMTAP algorithm inherits the advantages of previous moment-matching based methods. However, it avoids their shortfalls. PIMTAP method is computationally stable and robust. PIMTAP model preserves the structure of the original state equations like PRIMA [12] for non-parameterized RLC networks.

All multiparameter moment-matching based PMOR techniques, including the one proposed in this paper, are generally designed for a low-dimensional parameter space. Variations in modern VLSI technologies may introduce a high-dimensional parameter space, and is referred to as the *curse of dimensionality*. In [13] a two-step approach is introduced to construct a compact model of parameterized system: (1) parameter dimension reduction, i.e. construct reduced-parameter interconnect models which reduce a high-dimensional parameter space to a low-dimensional parameter space. (2) parameterized model order reduction, i.e., construct reduced order models from reduced-parameter models by applying parameterized model reduction techniques. Moment-matching based PMOR techniques are intended to be used for the second step to address the challenge of the *curse of dimensionality*.

We note that it is beyond the scope of this paper to compare the PIMTAP method with other parameterized and variational MOR techniques, such as truncated balance realization based approaches [8], stochastic spectral Galerkin method [1], [14], and rational approximation methods [15], [16].

II. BACKGROUND

We consider a single-input and single-output (SISO) parameterized linear dynamical system of the form

$$\begin{cases} \mathbf{C}(\lambda)\dot{\mathbf{x}} + \mathbf{G}(\lambda)\mathbf{x} &= \mathbf{b}u \\ y &= \mathbf{l}^T\mathbf{x} \end{cases} \quad (1)$$

with initial conditions $\mathbf{x}(\lambda, 0) = \mathbf{x}_0(\lambda)$ where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_k)$ are parameters. The matrices $\mathbf{C}(\lambda)$ and $\mathbf{G}(\lambda)$ are of *affine* forms:

$$\begin{aligned} \mathbf{C}(\lambda) &= \mathbf{C}_0 + \lambda_1\mathbf{C}_1 + \cdots + \lambda_k\mathbf{C}_k, \\ \mathbf{G}(\lambda) &= \mathbf{G}_0 + \lambda_1\mathbf{G}_1 + \cdots + \lambda_k\mathbf{G}_k, \end{aligned} \quad (2)$$

where $\mathbf{C}_i$ and $\mathbf{G}_i$ are $N \times N$ constant matrices. $\mathbf{b}$ is an excitation vector and $\mathbf{l}$ is a selector vector of the output of

interest y. $\mathbf{x}$ is a state vector of the system. The affine model (2) arises from parameterized interconnect networks, where $\mathbf{C}(\lambda)$ and $\mathbf{G}(\lambda)$ are the contributions of parameter dependent memory and memoryless elements of the network [5], [9], [17], [18], λ_i represent the geometry of the network, such as the width and height of an interconnect line. In [19], FEA of parametric electrothermal simulation also gives rise to the affine model (2).

For the clarity of exposition, we only consider the affine system of a single parameter:

$$\begin{cases} (\mathbf{C}_0 + \lambda\mathbf{C}_1)\dot{\mathbf{x}} + (\mathbf{G}_0 + \lambda\mathbf{G}_1)\mathbf{x} &= \mathbf{b}u \\ y &= \mathbf{l}^T\mathbf{x} \end{cases} \quad (3)$$

and λ is a scalar. It is assumed that $\mathbf{G}_0$ is nonsingular. The technique we present in this paper can be extended to systems of multiple parameters, see example B in section V and [20].

The *transfer function* of the system (3) is given by

$$h(s, \lambda) = \mathbf{l}^T(\mathbf{G}_0 + \lambda\mathbf{G}_1 + s(\mathbf{C}_0 + \lambda\mathbf{C}_1))^{-1}\mathbf{b}, \quad (4)$$

where $s = 2\pi f\mathtt{i}$, f is referred to as the *frequency* and $\mathtt{i} = \sqrt{-1}$. A power series expansion of $h(s, \lambda)$ about $(s, \lambda) = (0, 0)$ via a *symbolic* way [21] is given as

$$h(s, \lambda) = \sum_{j=0}^{\infty}\left(\sum_{i=0}^{\infty} m_i^j s^i\right)\lambda^j,$$

where $m_i^j = \mathbf{l}^T\mathbf{r}_i^j$ are referred to as *(multiparameter) moments*. The vectors $\mathbf{r}_i^j$, are *moment generating vectors* defined by the following two-directional recurrence

$$\mathbf{r}_i^j = -\mathbf{G}_0^{-1}(\mathbf{C}_0\mathbf{r}_{i-1}^j + \mathbf{G}_1\,\mathbf{r}_i^{j-1} + \mathbf{C}_1\,\mathbf{r}_{i-1}^{j-1}). \quad (5)$$

The initial vector is $\mathbf{r}_0^0 = \mathbf{G}_0^{-1}\mathbf{b}$. Note that $\mathbf{r}_i^j = \mathbf{0}$ if $i < 0$ or $j < 0$.

We pursue a moment-matching based algorithm under the framework of a subspace projection technique [4], [12], [22]. Specifically, let $\mathbf{V}$ be an orthonormal basis of a proper defined projection subspace $\mathcal{V}$ of dimension n, $n \le N$, then the reduced model of order n is given by

$$\begin{cases} (\widehat{\mathbf{C}}_0 + \lambda\widehat{\mathbf{C}}_1)\dot{\mathbf{z}} + (\widehat{\mathbf{G}}_0 + \lambda\widehat{\mathbf{G}}_1)\mathbf{z} &= \widehat{\mathbf{b}}u \\ \widehat{y} &= \widehat{\mathbf{l}}^T\mathbf{z}, \end{cases} \quad (6)$$

where $(\widehat{\mathbf{C}}_0, \widehat{\mathbf{C}}_1, \widehat{\mathbf{G}}_0, \widehat{\mathbf{G}}_1) \equiv \mathbf{V}^T(\mathbf{C}_0, \mathbf{C}_1, \mathbf{G}_0, \mathbf{G}_1)\mathbf{V}$ and $(\widehat{\mathbf{b}}, \widehat{\mathbf{l}}) \equiv \mathbf{V}^T(\mathbf{b}, \mathbf{l})$.

Given desired approximation orders p and q of frequency and geometric parameters s and λ, respectively, a proper choice of the projection subspace $\mathcal{V}$ should produce a reduced system (6) matching the following $(p+1)(q+1)$ moments:

$$m_i^j = \widehat{m}_i^j \quad (7)$$

for $i = 0, 1, 2, \ldots, p$ and $j = 0, 1, 2, \ldots, q$. This implies that the transfer function $\widehat{h}(s, \lambda)$ of the reduced system (6) is an order-(p, q) approximation of the original transfer function $h(s, \lambda)$, i.e., $h(s, \lambda) - \widehat{h}(s, \lambda) = \mathcal{O}(s^{p+1}) + \mathcal{O}(\lambda^{q+1})$.

To this end, it is clear that the gists of a moment-matching PMOR are (1) a proper choice of projection subspaces $\mathcal{V}$ and

(2) a stable and efficient algorithm to compute an orthonormal basis $\mathbf{V}$ of $\mathcal{V}$.

III. PREVIOUS WORK

A multivariable Taylor series method is proposed in [9]. By introducing an auxiliary parameter $\mu = s\lambda$, and using the multivariable Taylor series expansion, the transfer function (4) can be written as

$$h(\lambda, s, \mu) = \sum_{k=0}^{\infty}\sum_{j=0}^{k}\sum_{i=0}^{k-j} m_{i,j,k}\lambda^{k-(i+j)}s^i\mu^j,$$

where $m_{i,j,k} = \mathbf{l}^T\mathbf{f}_{i,j,k}$ are the *multiparameter moments*, and

$$\mathbf{f}_{i,j,k} = -\mathbf{G}_0^{-1}(\mathbf{C}_0\mathbf{f}_{i-1,j,k-1} + \mathbf{G}_1\mathbf{f}_{i,j,k-1} + \mathbf{C}_1\mathbf{f}_{i,j-1,k-1})$$
$$(8)$$

with the initial $\mathbf{f}_{0,0,0} = \mathbf{G}_0^{-1}\mathbf{b}$, where $\mathbf{f}_{i,j,k} = \mathbf{0}$, if $i, j, i+j \notin \{0, 1, \ldots, k\}$. To match the $(p+1)(q+1)$ moments defined in (7), the dimension of the projection subspace $\text{span}\{\mathbf{f}_{i,j,k}\}$ has to be $\mathcal{O}((p+q)^3)$.

An explicit-and-implicit moment-matching technique, referred to as Compact Order Reduction for parameterized Extraction (CORE) algorithm, is proposed [7]. CORE has two steps. At step 1: CORE defines the following linear system to *explicitly* approximate the geometric parameter λ to the qth order:

$$\begin{cases} (\mathbf{G}_{[q+1]} + s\mathbf{C}_{[q+1]})\widetilde{\mathbf{x}}_{[q+1]} &= \widetilde{\mathbf{b}}_{[q+1]}\tilde{u} \\ \widetilde{y}_{[q+1]} &= (\mathbf{l}_{[q+1]}(\lambda))^T\widetilde{\mathbf{x}}_{[q+1]}, \end{cases} \quad (9)$$

where $\mathbf{G}_{[q+1]}$ and $\mathbf{C}_{[q+1]}$ are $(q+1)$ by $(q+1)$ block lower bidiagonal matrices and $\widetilde{\mathbf{b}}_{[q+1]}$ and $\mathbf{l}_{[q+1]}(\lambda)$ are column vectors. The transfer function of (9) is an order-q approximation of the transfer function of the original system (3) in λ. At step 2, CORE applies a subspace projection method with the Krylov subspace

$$\mathcal{V} = \mathcal{K}_{p+1}\left(-\mathbf{G}_{[q+1]}^{-1}\mathbf{C}_{[q+1]}, \mathbf{G}_{[q+1]}^{-1}\widetilde{\mathbf{b}}_{[q+1]}\right)$$

to *implicitly* match the moments of s. A so-called *recursive Arnoldi algorithm* is used to generate an orthonormal basis $\mathbf{V}$ of $\mathcal{V}$, which effectively utilizes the block lower triangular structures of $\mathbf{G}_{[q+1]}$ and $\mathbf{C}_{[q+1]}$.

The CORE reduced-order model is defined by

$$\begin{cases} (\widehat{\mathbf{G}}_{[q+1]} + s\widehat{\mathbf{C}}_{[q+1]})\widetilde{\mathbf{z}}_{[q+1]} &= \widehat{\mathbf{b}}_{[q+1]}\tilde{u} \\ \widehat{\widetilde{y}}_{[q+1]} &= (\widehat{\mathbf{l}}_{[q+1]}(\lambda))^T\widetilde{\mathbf{z}}_{[q+1]}, \end{cases} \quad (10)$$

where $(\widehat{\mathbf{G}}_{[q+1]}, \widehat{\mathbf{C}}_{[q+1]}) \equiv \mathbf{V}^T(\mathbf{G}_{[q+1]}, \mathbf{C}_{[q+1]})\mathbf{V}$, $\widehat{\mathbf{b}}_{[q+1]} \equiv \mathbf{V}^T\mathbf{b}_{[q+1]}$ and $\widehat{\mathbf{l}}_{[q+1]}(\lambda) \equiv \mathbf{V}^T\mathbf{l}_{[q+1]}(\lambda)$. The total number of matched-moments is $(p+1)(q+1)$. But the dimension n of the reduced linear system (10) is $p+1$, independent of q. This is an advantage of the CORE method. However, due to the geometric parameter λ appears in a truncated power form of the reduced transfer function, it could lead to numerical instability illustrated by example A in section V. Moreover, CORE method does not preserve the structure of the original system, and it is not a passivity-preserving PMOR scheme.

IV. PIMTAP ALGORITHM

To match the moments defined in (7), a natural choice of the projection subspace $\mathcal{V}$ is

$$\mathcal{V} = \text{span}\{\mathbf{r}_i^j : i = 0, 1, \ldots p, \; j = 0, 1, \ldots, q\} \quad (11)$$

$$= \text{span}\left\{ \begin{array}{ccccc} \mathbf{r}_0^0, & \mathbf{r}_1^0, & \mathbf{r}_2^0, & \ldots, & \mathbf{r}_p^0, \\ \mathbf{r}_0^1, & \mathbf{r}_1^1, & \mathbf{r}_2^1, & \ldots, & \mathbf{r}_p^1, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{r}_0^q, & \mathbf{r}_1^q, & \mathbf{r}_2^q, & \ldots, & \mathbf{r}_p^q. \end{array} \right\}, \quad (12)$$

where the horizontal direction is for the approximation orders in frequency parameter s, and the vertical direction is for geometric parameter λ. A proof of the moment-matching property (7) with the projection subspace (11) is given in [20]. The dimension n of the reduced-order system (6) is no greater than $(p+1)(q+1)$. It could be smaller if deflations occur.

The critical question is how to stably and efficiently compute an orthonormal basis $\mathbf{V}$ of $\mathcal{V}$. The algorithm we propose in this section is to construct $\mathbf{V}$ incrementally in two directions. If $\mathbf{V}_j$ is an orthonormal basis of the subspace spanned by $\mathbf{r}_i^j$ of the first j rows of the $\mathbf{r}$-vector array (12), then to compute $\mathbf{V}_{j+1}$, we first form a basis $\mathbf{L}_{j+1}$ of the subspace spanned by the $(j+1)$st row vectors, i.e.,

$$\text{span}\{\mathbf{L}_{j+1}\} = \text{span}\{\mathbf{r}_0^j, \mathbf{r}_1^j, \ldots, \mathbf{r}_p^j\}, \quad (13)$$

and then combine $\mathbf{V}_j$ and $\mathbf{L}_{j+1}$, we have

$$\mathbf{V}_{j+1} = \text{orth}([\; \mathbf{V}_j \quad \mathbf{L}_{j+1} \;]), \quad (14)$$

where $\text{orth}(\mathbf{X})$ stands for an orthonormal basis of the range of $\mathbf{X}$. This process is done iteratively for $j = 1, 2, \ldots, q$ and $\mathbf{V} = \mathbf{V}_{q+1}$.

Specifically, we first compute an orthonormal basis $\mathbf{V}_1$ of the subspace spanned by the first row vectors $\mathbf{r}_i^0$. It is easy to see that the vectors $\mathbf{r}_i^0$ satisfy the recurrence

$$\mathbf{r}_i^0 = \mathbf{A}_{[1]}\mathbf{r}_{i-1}^0,$$

where $\mathbf{A}_{[1]} = -\mathbf{G}_0^{-1}\mathbf{C}_0$ and $\mathbf{r}_0^0 = \mathbf{G}_0^{-1}\mathbf{b} \equiv \mathbf{b}_{[1]}$. Hence the vectors $\mathbf{r}_i^0$ span the Krylov subspace $\mathcal{K}_{p+1}(\mathbf{A}_{[1]}, \mathbf{b}_{[1]})$. An orthonormal basis $\mathbf{V}_1$ of $\text{span}\{\mathbf{r}_i^0\}$ can be computed by the Arnoldi procedure [23].

We now consider how to compute the basis $\mathbf{L}_{j+1}$, $1 \leq j \leq q$. By stacking the vectors $\mathbf{r}_i^j$ columnwise and defining

$$\mathbf{r}_{[i]}^{[j+1]} = \left[\begin{array}{c} \mathbf{r}_{i-1}^0 \\ \mathbf{r}_{i-1}^1 \\ \vdots \\ \hline \mathbf{r}_{i-1}^j \end{array} \right] \equiv \left[\begin{array}{c} \mathbf{r}_{[i]}^{[j]} \\ \hline \mathbf{r}_{i-1}^j \end{array} \right],$$

it can be shown that we have the following recurrence

$$\mathbf{r}_{[i]}^{[j+1]} = \mathbf{A}_{[j+1]}\mathbf{r}_{[i-1]}^{[j+1]},$$

where

$$\mathbf{A}_{[j+1]} = \left[\begin{array}{cc} \mathbf{A}_{[j]} & \\ \mathbf{A}_{[j+1,:]} & \mathbf{A}_{j+1} \end{array} \right], \quad (15)$$

and

$$\mathbf{A}_{[j+1,:]} = -\mathbf{G}_0^{-1}\left([\; \mathbf{0} \quad \mathbf{G}_1 \;]\mathbf{A}_{[j]} + [\; \mathbf{0} \quad \mathbf{C}_1 \;]\right),$$

$$\mathbf{A}_{j+1} = -\mathbf{G}_0^{-1}\mathbf{C}_0.$$

The initial vector is

$$\mathbf{r}_{[1]}^{[j+1]} = \left[\begin{array}{c} \mathbf{b}_{[j]} \\ \mathbf{r}_0^j \end{array} \right] \equiv \mathbf{b}_{[j+1]} \quad (16)$$

where

$$\mathbf{r}_0^j = -\mathbf{G}_0^{-1}[\; \mathbf{0} \quad \mathbf{G}_1 \;]\mathbf{b}_{[j]}.$$

Thus the vectors $\mathbf{r}_{[i]}^{[j+1]}$ span the Krylov subspace

$$\text{span}\left\{\mathbf{r}_{[1]}^{[j+1]}, \mathbf{r}_{[2]}^{[j+1]}, \ldots, \mathbf{r}_{[p+1]}^{[j+1]}\right\} = \mathcal{K}_{p+1}(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]}).$$

We have two ways to generate the basis $\mathbf{L}_{j+1}$. The first way is to apply the Arnoldi procedure with $\mathbf{A}_{[j+1]}$ and $\mathbf{b}_{[j+1]}$ to obtain an orthonormal basis $\mathbf{Q}_{p+1}^{[j+1]}$ of $\mathcal{K}_{p+1}(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]})$. Let $\mathbf{Q}_{p+1}^{[j+1]}$ be partitioned into

$$\mathbf{Q}_{p+1}^{[j+1]} = \left[\begin{array}{c} \widetilde{\mathbf{Q}}_1 \\ \widetilde{\mathbf{Q}}_2 \end{array} \right] \begin{array}{c} jN \\ N \end{array}. \quad (17)$$

Then we have

$$\text{span}\{\widetilde{\mathbf{Q}}_2\} = \text{span}\{\mathbf{L}_{j+1}\} = \text{span}\{\mathbf{r}_0^j, \mathbf{r}_1^j, \ldots, \mathbf{r}_p^j\}. \quad (18)$$

Alternatively, an efficient way is to exploit the relation between the Krylov subspaces $\mathcal{K}_{p+1}(\mathbf{A}_{[j]}, \mathbf{b}_{[j]})$ and $\mathcal{K}_{p+1}(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]})$. By (15) and (16), we note that $\mathbf{A}_{[j+1]}$ and $\mathbf{b}_{[j+1]}$ are the bordered matrix and vector of $\mathbf{A}_{[j]}$ and $\mathbf{b}_{[j]}$, respectively, Therefore, if $\mathbf{Q}_{p+1}^{[j+1]}$ is an orthonormal basis of $\mathcal{K}_{p+1}(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]})$, where $\mathbf{Q}_{p+1}^{[1]} = \mathbf{V}_1$, and partitioned as in equation (17), then we have the following two facts:

1) $\text{span}\{\mathbf{r}_{[1]}^{[j]}, \mathbf{r}_{[2]}^{[j]}, \ldots, \mathbf{r}_{[p+1]}^{[j]}\} = \text{span}\{\widetilde{\mathbf{Q}}_1\}$.
2) $\text{span}\{\mathbf{r}_0^j, \mathbf{r}_1^j, \ldots, \mathbf{r}_p^j\} = \text{span}\{\widetilde{\mathbf{Q}}_2\}$.

Fact 1) is due to the relation (15). Fact 2) is given by the same argument used for equation (18). Therefore $\mathbf{Q}_{p+1}^{[j+1]}$ can be written in the form

$$\mathbf{Q}_{p+1}^{[j+1]} = \left[\begin{array}{c} \mathbf{Q}_{p+1}^{[j]}\mathbf{R}_{j+1} \\ \mathbf{L}_{j+1} \end{array} \right], \quad (19)$$

where $\mathbf{R}_{j+1}$ is upper triangular and $\mathbf{L}_{j+1}$ satisfies (13). As a result, we can exploit the relation (19) and compute $\mathbf{L}_{j+1}$ directly. Such a computational procedure is given in Appendix, referred to as a Two-directional Arnoldi Process (TAP). In the TAP, only the submatrices $\mathbf{A}_{[j+1,:]}$ and $\mathbf{A}_{j+1}$ are referenced through matrix-vector products, instead of $\mathbf{A}_{[j+1]}$.

Once the basis $\mathbf{L}_{j+1}$ is available, an orthonormal basis matrix $\mathbf{V}_{j+1}$ is computed by (14). In summary, the following algorithm is for computing the reduced system (6) satisfying the moment-matching property (7):

PIMTAP ALGORITHM

1) Run Arnoldi algorithm with $\mathbf{A}_{[1]}$ and $\mathbf{b}_{[1]}$ to construct the orthonormal basis $\mathbf{V}_1 \equiv \mathbf{Q}_{p+1}^{[1]}$.
2) For $j = 1, 2, \ldots, q$,

a) Run TAP (Appendix) with $\mathbf{A}_{[j+1,:]}$, $\mathbf{A}_{j+1}$ and $\mathbf{r}_0^j$ to construct the basis matrix $\mathbf{L}_{j+1}$.

b) $\mathbf{V}_{j+1} = \mathrm{orth}\left(\begin{bmatrix} \mathbf{V}_j & \mathbf{L}_{j+1} \end{bmatrix}\right)$.

c) Construct $\mathbf{Q}_{p+1}^{[j+1]}$ by (19).

3) Set $\mathbf{V} = \mathbf{V}_{q+1}$.

4) Compute the reduced-order model (6).

The costs of floating point operations of PIMTAP and CORE are essentially the same. However, PIMTAP has the following advantages:

- It is a stable process to generate an orthonormal basis $\mathbf{V}$ due to the stability of the 2D Arnoldi process.
- It produces a reduced system in the same form of the original system (3).
- It preserves important properties such as passivity and stability.

PIMTAP is an adaptive process with respect to the increase of the approximation order of geometric and/or frequency parameters, shown in Example B in Section V.

V. NUMERICAL RESULTS

In this section we present two examples to demonstrate the accuracy and efficiency of a Matlab implementation of the PIMTAP algorithm on a PC with 1.6 GHz Intel Core Duo T2050 processor.

EXAMPLE A. We consider an interconnect circuit, consisting of an 8-bit bus and two shielding lines from an industrial application. An RLC MNA formulation is used to model the capacitive and magnetic coupling effects between any two of these lines. The description matrices are

$$\mathbf{C}(\lambda) = \begin{bmatrix} (1+\lambda)\mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{L} \end{bmatrix}, \ \mathbf{G}(\lambda) = \begin{bmatrix} (1+\lambda)\mathbf{G} & \mathbf{E} \\ -\mathbf{E}^T & \mathbf{0} \end{bmatrix},$$

where $\mathbf{C}$, $\mathbf{L}$ and $\mathbf{G}$ are capacitance, inductance and conductance matrices, respectively. $\mathbf{E}$ is the incident matrix associated with the connectivity of the circuit. λ represents the fabrication variation, and is assumed to be varied within $\pm 15\%$. The order of $\mathbf{C}$ and $\mathbf{G}$ is 330. The order of $\mathbf{L}$ is 160.

Fig. 1 shows the relative errors of the transfer functions computed by CORE and PIMTAP at $\lambda = 0.06$. It clearly shows that PIMTAP with $(p,q) = (40,1)$ is more accurate than CORE with $(p,q) = (80,1)$. Note that the order of PIMTAP model is 76, instead of 82, due to the deflation among the basis vectors $\mathbf{r}_i^j$. The order of the CORE model is 81.

One way to improve the approximation accuracy is to increase the order q of λ. By Fig. 2, we see that the CORE model becomes unstable at high frequencies . On the other hand, such instability does not occur in the PIMTAP model (the PIMTAP curve is visually indistinguishable from the original one).

To compare the CPU elapsed time, the geometric parameter interval $[-0.15, 0.15]$ is divided into 10 equal length subintervals, and the frequency range $[0, 10]$ GHz into 300 equal length subintervals. Thus we have 11×301 grid points (λ_k, s_ℓ). The following table shows that for about the same accuracy, PIMTAP is about 30% faster than the CORE:

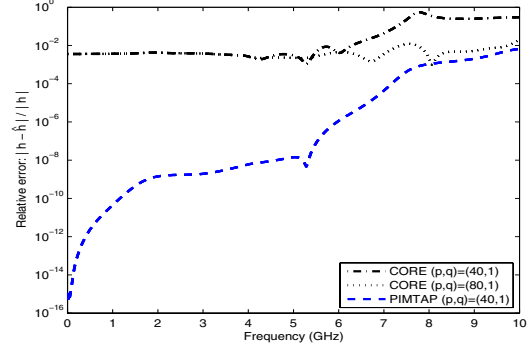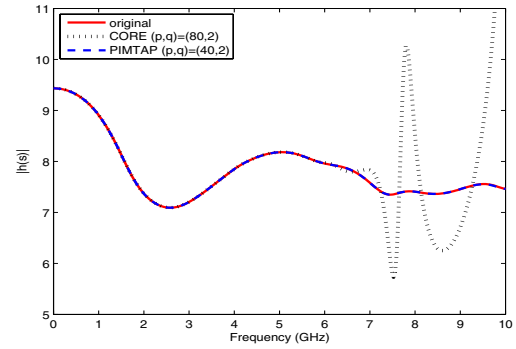Fig. 1.   The relative errors of CORE and PIMTAP.

Fig. 2.   Frequency responses.

	CORE		PIMTAP
(p,q)	(40,1)	(80,1)	(40,1)
Reduced-order n	41	81	76
Max relative-error	10^0	10^{-2}	10^{-2}
CPU-total (s)	2.578	10.490	7.172

Note that CORE with $(p,q) = (40,1)$ theoretically matches the same number of moments as PIMTAP with $(p,q) = (40,1)$. Unfortunately, the accuracy is not acceptable.

EXAMPLE B. We consider a multi-parameter model from a electrothermal simulation of a MEMS device described in [19]. After discretization in space by a FEM, we have a 3-parameter linear system:

$$\begin{cases} \mathbf{E}\dot{\mathbf{t}}(t) + (\mathbf{K} + \sum_{i \in \{t,s,b\}} \lambda_i \mathbf{K}_i)\mathbf{t}(t) &= \mathbf{b}u(t) \\ \mathbf{y}(t) &= \mathbf{C}^T\mathbf{t}(t), \end{cases} \quad (20)$$

where $\mathbf{t}(t)$ is the vector of unknown temperatures at 4257 nodes. $\mathbf{E}$ and $\mathbf{K}$ are of order 4257 representing the heat capacity and conductivity. λ_t, λ_s, and λ_b are film coefficients describing the heat flow between the device and three boundaries. $\mathbf{K}_t$, $\mathbf{K}_s$ and $\mathbf{K}_b$ specify the contributions of the film coefficients to the global system matrix $\mathbf{K}$. $u(t)$ is the heat source. $\mathbf{b}$ is a column vector and $\mathbf{C}$ is a 4257×7 output matrix. As discussed in [19], we consider an expansion of the

transfer function $\mathbf{h}(s, \lambda)$ at $(s, \lambda_t, \lambda_s, \lambda_b) = (0, 10, 10, 10)$:

$$\mathbf{h}(s, \lambda) = \sum_{|\alpha|=0}^{\infty} \sum_{i=0}^{\infty} (\mathbf{C}^T \mathbf{r}_i^\alpha) s^i \widetilde{\lambda}^\alpha,$$

where $\mathbf{r}_i^\alpha$ are defined by the recurrence

$$\mathbf{r}_i^\alpha = \mathbf{r}_i^{\alpha_t, \alpha_s, \alpha_b} = -\widetilde{\mathbf{K}}^{-1} \left(\mathbf{K}_t \, \mathbf{r}_i^{\alpha_t - 1, \alpha_s, \alpha_b} \right.$$
$$\left. + \mathbf{K}_s \, \mathbf{r}_i^{\alpha_t, \alpha_s - 1, \alpha_b} + \mathbf{K}_b \, \mathbf{r}_i^{\alpha_t, \alpha_s, \alpha_b - 1} + \mathbf{E} \, \mathbf{r}_{i-1}^{\alpha_t, \alpha_s, \alpha_b} \right)$$

with the initial $\mathbf{r}_0^{0,0,0} = \widetilde{\mathbf{K}}^{-1}\mathbf{b}$, where $\widetilde{\mathbf{K}} = \mathbf{K} + 10\mathbf{K}_t + 10\mathbf{K}_s + 10\mathbf{K}_b$ and $\widetilde{\lambda} = (\widetilde{\lambda}_t, \widetilde{\lambda}_s, \widetilde{\lambda}_b) = (\lambda_t - 10, \lambda_s - 10, \lambda_b - 10)$. $\alpha = (\alpha_t, \alpha_s, \alpha_b)$ and $|\alpha| = \alpha_t + \alpha_s + \alpha_b$. By *a priori* physical knowledge that λ_t plays a major role, a choice of the projection subspace is to include only those vectors $\mathbf{r}_i^\alpha$ corresponding to s and λ_t. PIMTAP has the *flexibility* to do so. For example, we can require to match 28 moments in s and 8 moments in λ_t by using the following projection subspace

$$\mathcal{V} = \text{span} \left\{ \begin{array}{l} \mathbf{r}_0^{0,0,0}, \mathbf{r}_1^{0,0,0}, \mathbf{r}_2^{0,0,0}, \ldots\ldots\ldots, \mathbf{r}_{27}^{0,0,0}, \\ \mathbf{r}_0^{1,0,0}, \mathbf{r}_1^{1,0,0}, \ldots, \mathbf{r}_7^{1,0,0}. \end{array} \right\}.$$

As a result, the order of PIMTAP model is 36. The following table shows the maximum relative errors of the PIMTAP model on the frequency range $[0, 100]$ Hz for 12 selected geometric parameters $(\lambda_t, \lambda_s, \lambda_b)$. It has about the same accuracy as reported in [19].

λ_t	λ_s	λ_b	max rel.-err
1	1	1	7.05×10^{-3}
1	1	10^4	7.05×10^{-3}
1	10^4	1	7.05×10^{-3}
1	10^4	10^4	7.05×10^{-3}
10^3	1	1	6.82×10^{-3}
10^3	1	10^4	6.82×10^{-3}
10^3	10^4	1	6.82×10^{-3}
10^3	10^4	10^4	6.82×10^{-3}
10^4	1	1	9.43×10^{-3}
10^4	1	10^4	9.34×10^{-3}
10^4	10^4	1	9.44×10^{-3}
10^4	10^4	10^4	9.36×10^{-3}

VI. Concluding Remarks

The main contributions of this paper are twofold. We first give a rigorous definition of the projection subspace for the multiparameter moment-matching property defined in (7). Then we present a stable and efficient numerical procedure to generate an orthonormal basis of the projection subspace. The new Arnoldi-like procedure exploits the two directional recurrence relationship among the moment generating vectors, and therefore is referred to as Two-directional Arnoldi Process (TAP). The resulting PIMTAP reduced order model has the same form of the original system and preserves the passivity. Finally, we point out that as illustrated by Example B, PIMTAP is flexible in matching selected numbers of moments corresponding to different parameters.

Appendix Two-directional Arnoldi Process

Let $\mathbf{A}_{[j]}$ and $\mathbf{b}_{[j]}$ be a sequence of block lower triangular matrices and vectors with conformal dimensions defined as

$$\mathbf{A}_{[j]} = \begin{bmatrix} \mathbf{A}_{[j-1]} & \mathbf{0} \\ \mathbf{A}_{[j,:]} & \mathbf{A}_j \end{bmatrix}, \quad \mathbf{b}_{[j]} = \begin{bmatrix} \mathbf{b}_{[j-1]} \\ \mathbf{b}_j \end{bmatrix} \quad (21)$$

for $j = 2, 3, \ldots$, with the initials $\mathbf{A}_{[1]} = \mathbf{A}_1$ and $\mathbf{b}_{[1]} = \mathbf{b}_1$, where $\mathbf{A}_j$ are $N_j \times N_j$ square matrices and $\mathbf{A}_{[j,:]}$ are matrices of conformal dimensions. The space

$$\mathcal{K}_k(\mathbf{A}_{[j]}, \mathbf{b}_{[j]}) = \text{span} \left\{ \mathbf{b}_{[j]}, \mathbf{A}_{[j]}\mathbf{b}_{[j]}, \ldots, \mathbf{A}_{[j]}^{k-1}\mathbf{b}_{[j]} \right\}$$

is the kth Krylov subspace induced by $\mathbf{A}_{[j]}$ and $\mathbf{b}_{[j]}$, referred to as the (j,k)th Krylov subspace for short.

The Arnoldi process [23], [24, sec.6.3] computes an orthonormal basis of the Krylov subspace with a fixed index j. Here we consider the situation where both k and j increase. Let $\mathbf{Q}_k^{[j]}$ be an orthonormal basis of the (j,k)th Krylov subspace $\mathcal{K}_k(\mathbf{A}_{[j]}, \mathbf{b}_{[j]})$, then the question is how to stably and efficiently compute an orthonormal basis $\mathbf{Q}_k^{[j+1]}$ of the $(j+1, k)$th Krylov subspace $\mathcal{K}_k(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]})$.

By (21), we have

$$(\mathbf{A}_{[j+1]})^i \mathbf{b}_{[j+1]} = \begin{bmatrix} (\mathbf{A}_{[j]})^i \mathbf{b}_{[j]} \\ \mathbf{v} \end{bmatrix},$$

where $\mathbf{v}$ is a vector of the length N_{j+1}. Therefore, we can conclude that $\mathbf{Q}_k^{[j]}$ and $\mathbf{Q}_k^{[j+1]}$ satisfy the following relation

$$\mathbf{Q}_k^{[j+1]} = \begin{bmatrix} \mathbf{Q}_k^{[j]}\mathbf{R}_{j+1} \\ \mathbf{L}_{j+1} \end{bmatrix}, \quad (22)$$

where $\mathbf{R}_{j+1}$ is a $k \times k$ nonsingular upper triangular matrix and $\mathbf{L}_{j+1}$ is an $N_{j+1} \times k$ matrix.

By the relation (22), we can derive a procedure to compute $\mathbf{R}_{j+1}$ and $\mathbf{L}_{j+1}$ directly. To limit the length of this paper, the derivation detail is omitted. In the following, we present a pseudocode of such a procedure. On input, we have an orthogonal matrix $\mathbf{Q}_{k+1}^{[j]}$ and a upper Hessenberg matrix $\widehat{\mathbf{H}}_k^{[j]}$ for the order-(j, k) Arnoldi decomposition induced by $\mathbf{A}_{[j]}$ and $\mathbf{b}_{[j]}$:

$$\begin{aligned} \mathbf{A}_{[j]}\mathbf{Q}_k^{[j]} &= \mathbf{Q}_k^{[j]}\mathbf{H}_k^{[j]} + h_{k+1,k}^{[j]}\mathbf{q}_{k+1}^{[j]}\mathbf{e}_k^T \\ &= \mathbf{Q}_{k+1}^{[j]}\widehat{\mathbf{H}}_k^{[j]}. \end{aligned}$$

On output, we have the matrices $\mathbf{R}_{j+1}$ and $\mathbf{L}_{j+1}$ and the upper Hessenberg matrix $\widehat{\mathbf{H}}_k^{[j+1]}$ for the order-$(j+1, k)$ Arnoldi decomposition induced by $\mathbf{A}_{[j+1]}$ and $\mathbf{b}_{[j+1]}$. Note that in the pseudocode the script $j+1$ of $\mathbf{R}_{j+1}$ and $\mathbf{L}_{j+1}$ and the superscript $[j+1]$ of $\widehat{\mathbf{H}}_k^{[j+1]}$ are omitted for clarity.

TAP: Two-directional Arnoldi Process

1) Set
   a) $\tau = \left\| \begin{bmatrix} (\mathbf{q}_1^{[j]})^T & (\mathbf{b}_{j+1}/\gamma_j)^T \end{bmatrix}^T \right\|$
   b) $\gamma_{j+1} = \gamma_j \tau$
   c) $\mathbf{R} = [\mathbf{r}_1] = [1/\tau]$
   d) $\mathbf{L} = [\mathbf{l}_1] = [\mathbf{b}_{j+1}/\gamma_{j+1}]$

e) $\widehat{\mathbf{H}}_0 = \emptyset$

2) For $i = 1, 2, \ldots, k$

    a) $\mathbf{x}_t = \mathbf{H}_i^{[j]} \mathbf{r}_i$

    b) $\mathbf{v}_b = \mathbf{A}_{[j+1,:]} \mathbf{Q}_i^{[j]} \mathbf{r}_i + \mathbf{A}_{j+1} \mathbf{l}_i$

    c) $\mathbf{h}_i = \mathbf{R}^T \mathbf{x}_t + \mathbf{L}^T \mathbf{v}_b$

    d) $\alpha_t = h_{i+1,i}^{[j]} (\mathbf{e}_i^T \mathbf{r}_i)$

    e) $\mathbf{w} = \mathbf{x}_t - \mathbf{R}\,\mathbf{h}_i$

    f) $\mathbf{z} = \mathbf{v}_b - \mathbf{L}\,\mathbf{h}_i$

    g) $h_{i+1,i} = \left(\left\| \begin{bmatrix} \mathbf{w} \\ \alpha_t \end{bmatrix} \right\|^2 + \|\mathbf{z}\|^2 \right)^{1/2}$

    h) $\mathbf{r}_{i+1} = \begin{bmatrix} \mathbf{w} \\ \alpha_t \end{bmatrix} / h_{i+1,i}$

    i) $\mathbf{l}_{i+1} = \mathbf{z} / h_{i+1,i}$

    j) set
$$\mathbf{R} = \begin{bmatrix} \mathbf{R} & \mathbf{r}_{i+1} \\ 0 & \end{bmatrix};$$
$$\mathbf{L} = \begin{bmatrix} \mathbf{L} & \mathbf{l}_{i+1} \end{bmatrix};$$
$$\widehat{\mathbf{H}}_i = \begin{bmatrix} \widehat{\mathbf{H}}_{i-1} & \mathbf{h}_i \\ 0 & h_{i+1,i} \end{bmatrix}$$

Two remarks are in order:

1) The scalar γ_j at step 1)-a) is a scaling factor such that the vector $\widehat{\mathbf{b}}_{[j+1]} = \left[(\mathbf{q}_1^{[j]})^T \; (\mathbf{b}_{j+1}/\gamma_j)^T \right]^T$ is parallel to the vector $\mathbf{b}_{[j+1]}$ and thus $\mathcal{K}_{k+1}(\mathbf{A}_{[j+1]}, \widehat{\mathbf{b}}_{[j+1]}) = \mathcal{K}_{k+1}(\mathbf{A}_{[j+1]}, \mathbf{b}_{[j+1]})$. The sequence $\{\gamma_j\}$ is defined by $\gamma_{j+1} = \|\widehat{\mathbf{b}}_{[j+1]}\| \gamma_j > 0$ with the initial $\gamma_1 = \|\mathbf{b}_{[1]}\|$ and γ_{j+1} is computed at step 1)-b).

2) If the matrix $\mathbf{A}_{[j+1]}$ and the vector $\mathbf{b}_{[j+1]}$ are as defined (15) and (16), we do not explicitly formulate $\mathbf{A}_{[j+1,:]}$, $\mathbf{A}_{j+1}$ and $\mathbf{b}_{j+1}$. We can use a recursive algorithm similar to the one presented at [7] for the vector $\mathbf{b}_{j+1}$ at steps 1)-a) and d), and the matrix-vector multiplication at step 2)-b).

ACKNOWLEDGMENT

REFERENCES

[1] J. Wang, P. Ghanta, and S. Vrudhula, "Stochastic analysis of interconnect performance in the presence of process variations," in *IEEE/ACM ICCAD*, 2004, pp. 880–886.

[2] J. M. Wang, O. A. Hafiz, and J. Li, "A linear fractional transform (LFT) based model for interconnect parametric uncertainty," in *IEEE/ACM DAC*, 2004, pp. 375–380.

[3] P. Gunupudi, R. Khazaka, and M. Nakhla, "Analysis of transmission line circuits using multidimensional model reduction techniques," *IEEE Trans. on Advanced Package*, vol. 25, no. 2, pp. 174–180, May 2002.

[4] P. K. Gunupudi, R. Khazaka, M. S. Nakhla, T. Smy, and D. Celo, "Passive parameterized time-domain macromodels for high-speed transmission-line networks," *IEEE Trans. on Microwave Theory and Techniques*, vol. 51, no. 12, pp. 2347–2354, December 2003.

[5] Y. Liu, L. T. Pileggi, and A. J. Strojwas, "Model order-reduction of RC(L) interconnect including variational analysis," in *IEEE/ACM DAC*, 1999, pp. 201–206.

[6] P. Li, F. Liu, X. Li, L. T. Pileggi, and S. R. Nassif, "Modeling interconnect variability using efficient parametric model order reduction," in *IEEE/ACM DATE*, 2005, pp. 958–963.

[7] X. Li, P. Li, and L. T. Pileggi, "Parameterized interconnect order reduction with explicit-and-implicit multi-parameter moment matching for inter/intra-die variations," in *IEEE/ACM ICCAD*, 2005, pp. 806–812.

[8] J. R. Phillips and L. M. Silveira, "Poor man's TBR: a simple model reduction scheme," *IEEE Trans. CAD*, vol. 24, no. 1, pp. 43–55, 2005.

[9] L. Daniel, O. C. Siong, L. S. Chay, K. H. Lee, and J. White, "A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models," *IEEE Trans. CAD*, vol. 23, no. 5, pp. 678–693, 2004.

[10] L. Feng, "Parameter independent model order reduction," *Mathematics and Computers in Simulation*, vol. 68, no. 3, pp. 221–234, 2005.

[11] L. H. Feng, E. B. Rudnyi, and J. G. Korvink, "Preserving the film coefficient as a parameter in the compact thermal model for fast electrothermal simulation," *IEEE Trans. CAD*, vol. 24, no. 12, pp. 1838–1847, 2005.

[12] A. Odabasioglu, M. Celik, and L. T. Pileggi, "PRIMA: passive reduced-order interconnect macromodeling algorithm," *IEEE Trans. CAD*, vol. 17, no. 8, pp. 645–654, August 1998.

[13] Z. Feng and P. Li, "Performance-oriented statistical parameter reduction of parameterized systems via reduced rank regression," in *IEEE/ACM ICCAD*, 2006, pp. 868–875.

[14] P. Ghanta, S. Vrudhula, R. Panda, and J. Wang, "Stochastic power grid analysis considering process variations," in *IEEE/ACM DATE*, 2005, pp. 964–969.

[15] K. C. Sou, A. Megretski, and L. Daniel, "A quasi-convex optimization approach to parameterized model order reduction," in *IEEE/ACM DAC*, June 2005.

[16] S. Grivet-Talocia, S. Acquadro, M. Bandinu, F. G. Canavero, I. Kelander, and M. Rouvala, "A parameterization scheme for lossy transmission line macromodels with application to high speed interconnects in mobile devices," *IEEE Trans. on Elec. Comp.*, vol. 49, no. 1, pp. 18–24, Feb 2007.

[17] E. Acar, S. Nassif, Y. Liu, and L. T. Pileggi, "Time-domain simulation of variational interconnect models," in *ISQED*, March 2002, pp. 419–424.

[18] P. Heydari and M. Pedram, "Model-order reduction using variational balanced truncation with spectral shaping," *IEEE Trans. CAS-I*, vol. 53, no. 4, pp. 879–891, April 2006.

[19] E. B. Rudnyi, L. H. Feng, M. Salleras, S. Marco, and J. G. Korvink, "Error indicator to automatically generate dynamic compact parametric thermal models," in *Proceedings of Therminic 2005*, September 2005, pp. 139–145.

[20] Y.-T. Li, Z. Bai, Y. Su, and X. Zeng, "Model Order Reduction of Parameterized Interconnect Networks via a Two-Directional Arnoldi Process," Universtiy of California, Davis, Department of Computer Science, Tech. Rep. CSE-2007-18, 2007.

[21] A. Cuyt, "How well can the concept of Padé approximant be generalized to the multivariate case?" *J. Comp. Appl. Math.*, vol. 105, no. 1-2, pp. 25–50, 1999.

[22] R.-C. Li and Z. Bai, "Structure-preserving model reduction using a Krylov subspace projection formulation," *Comm. Math. Sci.*, vol. 3, no. 2, pp. 179–199, 2005.

[23] W. E. Arnoldi, "The principle of minimized iteration in the solution of the matrix eigenvalue problem," *Quart. Appl. Math.*, vol. 9, pp. 17–29, 1951.

[24] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia: Society for Industrial and Applied Mathematics, 2003.