

SPARE—A Scalable Algorithm for Passive, Structure Preserving, Parameter-Aware Model Order Reduction

Jorge Fernández Villena, *Student Member, IEEE*, and Luís Miguel Silveira, *Senior Member, IEEE*

Abstract— This paper describes a flexible and efficient new algorithm for model order reduction of parameterized systems. The method is based on the reformulation of the parameterized system as a perturbation-like parallel interconnection of the nominal transfer function and the nonparameterized transfer function sensitivities with respect to the parameter variations. Such a formulation reveals an explicit dependence on each parameter which is exploited by reducing each component system independently via a standard nonparameterized structure preserving algorithm. Therefore, the resulting smaller size interconnected system retains the structure of the original system with respect to parameter dependence. This allows for better accuracy control, enabling independent adaptive order determination with respect to each parameter and adding flexibility in simulation environments. It is shown that the method is efficiently scalable and preserves relevant system properties such as passivity. The new technique can handle fairly large parameter variations on systems whose outputs exhibit smooth dependence on the parameters, also allowing design space exploration to some degree. Several examples show that besides the added flexibility and control, when compared with competing algorithms, the proposed technique can, in some cases, produce smaller reduced models with potential accuracy gains.

Index Terms—Model order reduction (MOR), parameterized models, parameterized reduction, projection methods, simulation.

I. INTRODUCTION

MODEL order reduction (MOR) techniques are a set of numerical procedures which aim at replacing a large-scale model of a physical system by a lower dimensional (or reduced order) model which exhibits similar input–output behavior. This order reduction enables efficient simulation and verification of large systems [2], [3]. Since the first attempts in this area, the methods for linear model reduction have greatly evolved and can be broadly characterized into those based on transfer function matching [4], [5] and projection methods.

Manuscript received May 12, 2009; revised November 5, 2009. Date of current version May 21, 2010. Preliminary results appeared in [1]. This work was supported in part by the EU Information Society Technologies (IST) CHAMELEON RF Project (FP6/2004/IST/4-027378), and in part by the NXP Semiconductors Research, under EMonIC Project. This paper was recommended by Associate Editor R. Suaya.

The authors are with the Instituto Superior Técnico, Technical University of Lisbon, 1000-029 Lisbon, Portugal and also with the Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa, 1000-029 Lisbon, Portugal. Luís Miguel Silveira is also with Cadence Research Laboratories, Berkeley, CA 94704 USA. (e-mail: jorge.fernandez@inesc-id.pt; lms@inesc-id.pt).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCAD.2010.2048372

Among the former, it can be distinguished between moment matching [6], [7], those based on balancing techniques [8], [9], and sampling based methods [10]–[12].

Although previously ignored, variability can no longer be disregarded as it directly impacts system behavior and performance. Accounting for the effects of process and geometric parameters, temperature, etc., leads to parameterized models whose complexity must be tackled both during the design and verification phases. Therefore, parameterized MOR (pMOR) techniques, able to generate reduced models that accurately capture the effects of the variability, are being considered as essential in the determination of correct system behavior.

Several pMOR techniques have been developed for modeling large-scale parameterized systems. Although the first attempts were based on perturbation, such as [13], [14], the most common and effective ones appear to be extensions of the basic projection-based MOR algorithms [6], [10] to handle parameterized descriptions. These projection based pMOR techniques can be broadly cast into *multidimensional moment matching* [15]–[19], whose projector is built from a selection of the orthonormalized moments of the multidimensional transfer function, or into *sample-based techniques* [16], [18], [20], which generate the projector from samples in the multidimensional space.

Most of those pMOR methods are oriented toward handling small variations of process parameters, and seem particularly suited for simple interconnects, such as buses and clock trees, whose parameterized response can be captured by slightly increasing the dimension of the projection basis. On the other hand, in the performance of analog mixed-signal systems, designed-in passives, such as integrated spirals and capacitors, play a fundamental role. Here, detailed electromagnetic (EM) models are needed for taking into account all possible effects. Whenever the geometrical layout parameters are changed, a new EM model must be generated and simulated, which leads to efficiency loss, as design parameters may have a large range, and may completely modify the behavior of the system. In this scenario, although the variation on the output due to the parameters may be smooth, the subspace to match is highly increased, and so is the size of the reduced order model (ROM) given by standard pMOR methodologies.

A different approach, based on a Taylor series representation of the effect of parameters on the output of the system,

was proposed in [21]. This approach directly captures the parameterized dependence in an explicit sense, being able to tackle fairly large parameter variations in some scenarios. Unfortunately the technique does not guarantee passivity, and the parameterized dependence is lost after reduction. Therefore, the method is not efficiently scalable if higher order approximation on the parameters is required.

In this paper, we present an algorithm for pMOR that, similarly to [21], is based on the reformulation of the parameterized system, revealing an explicit dependence on each parameter. Unlike [21], however, here this dependence is directly exploited by reducing each component system independently. Therefore, in the proposed approach, the resulting reduced model retains the structure of the original with respect to parameter dependence. This allows for better accuracy control, enabling independent adaptive order determination with respect to each parameter and adding flexibility in simulation environments. Furthermore, the procedure is shown to preserve passivity and is efficiently scalable if higher accuracy, thus, higher order is required.

This paper is structured as follows. In Section II, an overview of pMOR and a discussion of existing techniques is presented. In Section III, the new scheme will be introduced, starting with a description of the underlying representation and the methodology for reduction. We argue that the reduction effort and the ROM size are fairly independent from the variation range envisioned. We also show that the procedure preserves passivity of the model. Section IV will introduce some implementation and computational considerations about the proposed methodology. In Section V, several examples are shown that illustrate the efficiency of the proposed technique, and in Section VI, the conclusion is drawn.

II. BACKGROUND

A. Parameterized Systems

Parameter aware modeling stages lead to parameterized state-space system representations with the next descriptor form

$$C(\lambda)\dot{x}(\lambda) + G(\lambda)x(\lambda) = Bu, \quad y(\lambda) = Ex(\lambda) \quad (1)$$

where $C, G \in \mathbb{R}^{n \times n}$ are, respectively, the dynamic and static matrix descriptors, $B \in \mathbb{R}^{n \times m}$ relates the input vector $u \in \mathbb{R}^m$ to the state vector $x \in \mathbb{R}^n$, and $E \in \mathbb{R}^{n \times p}$ links those inner states to the outputs $y \in \mathbb{R}^p$. The elements of C and G , as well as the states x , depend on a set of Q parameters $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_Q]$ which model the effects of the mentioned variability. Usually, the system is formulated such that the matrices related to the inputs and outputs (B and E) do not depend on the parameters. The associated parameter dependent frequency response modeled via the transfer function

$$H(s, \lambda) = E(sC(\lambda) + G(\lambda))^{-1}B \quad (2)$$

for which we seek to generate a reduced order approximation, able to accurately capture the input–output behavior of the system for any point in the joint frequency–parameter space

$$\widehat{H}(s, \lambda) = \widehat{E}(s\widehat{C}(\lambda) + \widehat{G}(\lambda))^{-1}\widehat{B}. \quad (3)$$

In general, one attempts to generate a ROM whose structure is as similar to the original as possible, i.e., exhibiting a similar parameter dependence. The most common procedure to achieve this goal is to use some form of projection scheme. Once a suitable subspace basis V is computed, the system can be projected into that subspace, and a reduced model such as (3) can be obtained, that captures the behavior of the system under parameter variations

$$\begin{aligned} \widehat{C}(\lambda) &= V^T C(\lambda) V & \widehat{G}(\lambda) &= V^T G(\lambda) V \\ \widehat{B} &= V^T B & \widehat{E} &= E V \\ & & x(\lambda) &= V \widehat{x}(\lambda) \end{aligned} \quad (4)$$

where $V \in \mathbb{R}^{n \times q}$ spans the projection subspace of reduced dimension q , and $\widehat{C}, \widehat{G} \in \mathbb{R}^{q \times q}$, $\widehat{B} \in \mathbb{R}^{q \times m}$, $\widehat{E} \in \mathbb{R}^{q \times p}$, and $\widehat{x} \in \mathbb{R}^q$ define the ROM.

B. Representation as a Matrix Taylor Series Approximation

The state space descriptor system in (1) is a mathematical representation of a physical system depending on a set of parameters. The matrices $G(\lambda)$ and $C(\lambda)$ usually represent a discrete characterization of the physical system in terms of electrical elements, which in this case depend on a set of physical parameters. Existing literature [15] advocates for a Taylor series expansion of the elements of these matrices with respect to the parameters, which leads to

$$\begin{aligned} G(\lambda) &= G_{0\dots 0} + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} G_{\psi_1 \dots \psi_Q} \\ C(\lambda) &= C_{0\dots 0} + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} C_{\psi_1 \dots \psi_Q} \end{aligned} \quad (5)$$

where $G_{0\dots 0}$ and $C_{0\dots 0}$ are the nominal values for the matrices, $G_{\psi_1 \dots \psi_Q}$ and $C_{\psi_1 \dots \psi_Q}$ are the joint sensitivities of order $(\psi_1 \dots \psi_Q)$ with respect to the Q parameters, and $\Lambda_{\psi_1 \dots \psi_Q} = \lambda_1^{\psi_1} \dots \lambda_Q^{\psi_Q}$. For simplicity and without loss of generality, we assume that each λ has a normalized variation around the nominal value to a maximum of $|\lambda| \leq 1$. The Taylor series can be extended up to the desired (or required) order, including cross terms, for the sake of accuracy. This representation of $G(\lambda)$ and $C(\lambda)$ can be included in (1) to generate a state space Taylor series representation, which provides some advantages. One of the most relevant comes from compatibility with respect to extraction methods, which usually take into account variability by using sensitivity analysis, and thus, directly provide the matrix terms of this representation (see, for example [22]–[24]). Another advantage is the fact that a projection based reduction leads to a ROM of the same form, i.e., a state space Taylor series representation whose matrices have smaller dimensions. This allows for an explicit parameter representation, and thus, the same ROM can be efficiently evaluated for different parameter settings (notice, however, that any parameter change implies solving the system again).

A relevant issue to take into account for this kind of representation is the stability and passivity of the system. Any physical linear system, independently of the values of the parameters it depends on, is passive, as it is unable to generate energy. In mathematical terms, passivity is associated with positive realness of its transfer function. A matrix transfer function $H(s)$ is positive real if

$$\begin{aligned} H(s) &= \bar{H}(\bar{s}) \\ H(s) & \text{ is analytic in } \operatorname{Re}[s] > 0 \\ H(s) + H^*(s) & \text{ is PSD in } \operatorname{Re}[s] > 0 \end{aligned} \quad (6)$$

where Re stands for real part, $*$ for Hermitian (transpose conjugate), and PSD stands for positive semidefinite. These conditions are complied by electrical circuits under certain structural representations. In [6], it was shown that $B = E^T$ and C, G be PSD are sufficient conditions for passivity in an LTI system. In the case of parameterized systems, the positive realness should be extended to any parameter value in the domain of interest (let us denote Γ as the appropriate set of values for λ that makes sense for the application). Therefore $H(s, \lambda)$ should be analytic in $Re[s] > 0$ and $H(s, \lambda) + H^T(s^*, \lambda)$ be PSD $\forall Re[s] > 0$ and $\forall \lambda \in \Gamma$. In terms of state space descriptor, sufficient conditions for passivity are $B = E^T$ and $C(\lambda), G(\lambda)$ are PSD $\forall \lambda \in \Gamma$.

It is clear that passivity of a parameterized system depends on the parameter values, and the range of variation of these parameters is limited if passivity is to be assured. If a stamping methodology is used to build the matrices, it should be assured that the physical meaning of any stamped element remains the same for the range of variation of the parameters. This means that a conductance, capacitance, inductance, etc, should remain as such for any parameter value. Consider, for example, a conductance g as a function of a set of parameters. If the parameters represent physically meaningful variations, it should be guaranteed that $g(\lambda) \geq 0 \forall \lambda \in \Gamma$

$$g(\lambda) = g_{0\dots 0} + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} g_{\psi_1 \dots \psi_Q} \geq 0 \quad \forall \lambda. \quad (7)$$

If the sensitivities are built blockwise (e.g., via direct differentiation), a similar result should be enforced.

Regarding the reduction of such representation, the passivity of the ROM is guaranteed if a congruence projection, as the one in (4), is applied (see [6] or Section VI of [25]), with the additional advantage that an equivalent, yet reduced, Taylor series representation is generated.

C. Projection Based Reduction

Most of the techniques appearing in the literature extend the moment matching paradigm [6], [7] to the multidimensional case. They usually rely on the implicit or explicit matching of the moments of the parameterized transfer function (2). These moments depend not only on the frequency, but also on the set of parameters affecting the system, and thus, are denoted as multidimensional or multiparameter moments

$$x(s, \lambda) = \sum_{\psi_s \psi_1 \dots \psi_Q} M_{\psi_s \psi_1 \dots \psi_Q} s^{\psi_s} \Lambda_{\psi_1 \dots \psi_Q} \quad (8)$$

where $M_{\psi_s \psi_1 \dots \psi_Q}$ is a ψ th ($\psi = \psi_s + \psi_1 + \dots + \psi_Q$) order multiparameter moment corresponding to the coefficient term $s^{\psi_s} \Lambda_{\psi_1 \dots \psi_Q}$. Following the same idea used in the nominal moment matching techniques, a basis for the subspace formed from these moments can be built

$$\text{colspan}[V] = \text{span}\{M_{0,0\dots 0}, \dots, M_{\psi_s \psi_1 \dots \psi_Q}\} \quad (9)$$

and the resulting matrix V can be used as a projection matrix for reducing the original system. The generated parameterized ROM matches up to the ψ th order multiparameter moment of the original system. Different approaches differ in which moments are matched and how these moments are generated. The

straightforward approach, denoted multiparameter moment matching [15], is to match all the combinations. Some schemes try to improve this paradigm by low-rank approximations [16], others obtain the projector from an overall basis of multiple disjoint subspaces, built separately for each dimension, i.e., the frequency s (for which ψ_s block moments are obtained) and for each parameter λ_i (generating ψ_i block moments) [17]. Recent approaches [18], [19] rely on a recursive procedure to compute the frequency moments of different order approximation in the parameters. This adds flexibility as the number of moments to match with respect to each parameter and the frequency can be different.

In general, these methods, which rely on local matching, suffer from oversize of the models when the number of moments to match is high, either because of large frequency ranges, or because the number of parameters is large, and their variation can lead to vastly different dynamics changes in different frequency ranges.

Although based on moment matching paradigms, some methods rely on sampling, either to generate Krylov subspaces for the systems when evaluated at different parameter sets ($\lambda = \lambda_k$), as advocated in [16], or to select the most relevant or probable moments, as in [18]. A different sampling based approach presented in [20] extends the poor man's truncated balanced realization (PMTBR) [10] algorithm to include variability by means of a statistical interpretation. The method relies in a multidimensional sampling of the joint frequency plus parameter space. It can also use statistical information available for λ to guide the sampling or apply a weighting scheme in the parameter space. After the sampling, the most relevant vectors are selected via singular value decomposition (SVD) in order to build a projection matrix. Noticeable advantages are that easy to compute proxies for error bounds estimates can be obtained (via analysis of the eigenvalues of the SVD), and the size of the reduced model is less sensitive to the number of parameters, or how the parameter dependence is modeled. On the other hand, the issue of sample selection, already an important one in the nonparameterized version, becomes even more relevant, since a much higher-dimensional space must be sampled.

D. Explicit Parameter Matching

A different approach was previously presented in the compact order reduction for parameterized extraction (CORE) algorithm [21], which proposes an explicit moment matching with respect to the parameters. The system is reformulated by expanding it in Taylor series, supposedly accurate enough in the range of variation of the parameter set $[\lambda_1 \dots \lambda_Q]$. This is achieved by a Taylor series representation of the matrices $C(\lambda)$ and $G(\lambda)$ as in (5), plus an expansion of the state vector in Taylor series with respect to the parameters, but not with respect to the frequency

$$x(s, \lambda) = x_{0\dots 0}(s) + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} x_{\psi_1 \dots \psi_Q}(s) \quad (10)$$

where the subscript indicates parameter dependence and frequency dependence is explicitly stated. Again, and although in [21] only first order was used, it can be extended up to

the desired (or required) order, including cross derivatives, for the sake of accuracy. Mimicking the procedure of matching the coefficients of the same powers, an augmented system can be obtained. As an illustration, let us suppose two parameters, λ_1 and λ_2 , with second order expansion, and one cross term [following (5)], subindex ψ , ϑ indicates ψ th order with respect to the first parameter and ϑ th order with respect to the second)

$$\begin{aligned}
x_{00}(s) &= (G_{00} + sC_{00})^{-1}Bu \\
x_{10}(s) &= -(G_{00} + sC_{00})^{-1}(G_{10} + sC_{10})x_{00}(s) \\
x_{01}(s) &= -(G_{00} + sC_{00})^{-1}(G_{01} + sC_{01})x_{00}(s) \\
x_{20}(s) &= -(G_{00} + sC_{00})^{-1}(G_{10} + sC_{10})x_{10}(s) \\
&\quad -(G_{00} + sC_{00})^{-1}(G_{20} + sC_{20})x_{00}(s) \\
x_{02}(s) &= -(G_{00} + sC_{00})^{-1}(G_{01} + sC_{01})x_{01}(s) \\
&\quad -(G_{00} + sC_{00})^{-1}(G_{02} + sC_{02})x_{00}(s) \\
x_{11}(s) &= -(G_{00} + sC_{00})^{-1}(G_{11} + sC_{11})x_{00}(s) \\
&\quad -(G_{00} + sC_{00})^{-1}(G_{10} + sC_{10})x_{01}(s) \\
&\quad -(G_{00} + sC_{00})^{-1}(G_{01} + sC_{01})x_{10}(s).
\end{aligned} \tag{11}$$

Using the state vector $\mathbf{x} = [x_{00}^T \ x_{10}^T \ x_{01}^T \ x_{20}^T \ x_{02}^T \ x_{11}^T]^T$, we can generate a nonparameterized augmented system, in which the states no longer depend on the parameters, and the parameter dependence is shifted into the output matrix (see [21] for details). This system is then used as initial formulation for the parameterized model. CORE uses this new representation to generate a q -order basis for the new state vector \mathbf{x} , via nominal moment matching approaches [6]. The system is then reduced via orthogonal projection $\mathbf{x} = V\hat{\mathbf{x}}$.

Advantages of this approach are that in this representation the states only depend on the frequency, and thus, any nominal MOR approach can be applied. Another advantage is the fact that by projecting this system with a Krylov moment matching approach to a given reduced order q , q frequency moments of the \mathbf{x} vector are matched, and thus, the parameter effect of the augmented system is captured (for any parameter value).

On the other hand, there are some relevant disadvantages to the CORE approach. Projection of the matrices leads to full reduced matrices, in which all the structure is lost. Moreover, the parameter dependence is shifted to the output matrix, and thus, the parameter dependence is lost after reduction due to the projection, making model evaluation for different parameter values expensive and decreasing flexibility. Also, the method is not scalable, as the expansion order with respect to the parameters must be decided *a priori*. Furthermore, accuracy considerations for the frequency moment matching based on a single expansion point may require high orders in large frequency ranges. A last critical drawback is the passivity preservation. Orthogonal projection on the augmented system guarantees stability preservation (see [21] for details), but on the other hand the $B = E^T$ condition is no longer complied, which means that no guarantee can be given with respect to passivity preservation.

III. SCALABLE PARAMETER AWARE REDUCTION

In this section we outline scalable parameter aware reduction (SPARE), which exhibits the following properties.

- **Explicit structural dependence on the parameter set is maintained** which implies that reevaluation of the model for different parameter values is very efficient.
- **Model evaluation is efficient** as the reduced model is described by a sparse block lower triangular form, that allows to apply recursive procedures.
- Model is **accurate** for smooth output dependence on the parameters, as any range of variation is perfectly matched as long as the underlying output Taylor series formulation is accurate enough.
- **Accuracy control** is enhanced as **independent adaptive order determination** with respect to each parameter (i.e., for each parameter related transfer function) is possible.
- Algorithm and model are **scalable** as terms can be added or removed at any stage to modify the order, depending on the accuracy required, with data reuse. The accuracy/order trade off can be easily determined from the output sensitivities, and model size increases linearly the number of terms.
- **Passivity is preserved** in the reduction process, as long as the proposed projection based methodology is applied.
- **Model construction is flexible** in the sense that any nominal (nonparameterized) MOR can be used for reducing each parameter related transfer function. Even different approaches can be used for different transfer functions, although it can have consequences on the efficiency. As we shall see either moment matching or sampled techniques can be efficiently applied.
- **Design parameters** can be handled up to some degree, as long as they have a smooth effect on the output. This feature in combination with the efficient model evaluation allow a fast first approach to optimal values in optimization steps.

A. SPARE Representation

As mentioned, the approach followed is akin to the one developed in [21] in terms of representation, but, as we shall see, it deviates from it afterward. Therefore, we follow the same steps so that we arrive to the representation of multiple states as in (11). In a general case

$$H(s, \lambda) = H_{0\dots 0}(s) + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} H_{\psi_1 \dots \psi_Q}(s) \tag{12}$$

where $H_{\psi_1 \dots \psi_Q} = Ex_{\psi_1 \dots \psi_Q}$, making the parameterized dependence clearly explicit. In other words, the parameterized transfer function can be written as the contribution of the nominal transfer function plus the one of the nonparameterized transfer function sensitivities with respect to the parameters, i.e., a linear combination of the multiple nonparameterized transfer functions weighted by the parameter variation. Recall that this representation can be extended up to any order on the parameters, including cross-terms and, as we shall see, additional terms can be added at a later stage with little extra cost.

Every transfer function in (12) can be easily represented in its state space descriptor starting from the nominal matrices and the sensitivities of the Taylor series expansion (5). Again, we will illustrate the representation using the two parameter

outline. To apply PMTBR to the multiple interconnected, individual systems can be very expensive due to the sample effort required. However, careful consideration of the structure of the system matrices and of the computation involved can lead to considerable savings

$$H_{10} = -E(G_{00} + sC_{00})^{-1}(G_{10} + sC_{10})(G_{00} + sC_{00})^{-1}B.$$

In order to sample this function in the PMTBR scheme we need to obtain the sample vectors

$$z_{10k} = (G_{00} + s_k C_{00})^{-1}(G_{10} + s_k C_{10})(G_{00} + s_k C_{00})^{-1}B.$$

The term $(G_{00} + s_k C_{00})^{-1}$ is common to the nominal and all sensitivity transfer functions. Thus, it only needs to be generated once for each sample, factored and then reused in ensuing computations. This computation requires care in order to ensure numerical stability. In fact, a single LU factorization at each sample point can allow us to obtain the sampled vectors of all the transfer functions in (12) with much less computational effort. Once the samples are obtained for each transfer function (i.e., for every $s_{k\psi\vartheta}$), an SVD is applied to orthonormalize the vectors, and the vectors whose associated singular value falls below a settled tolerance are dropped. The remaining (dominant) vectors generate an orthonormal projector, $V_{\psi\vartheta}$. A BSP-type projector [27] \mathcal{V} , generated from the block projectors $V_{\psi\vartheta}$, is then obtained

$$\mathcal{V} = \text{diag}\{V_{00}, V_{10}, \dots, V_{\psi\vartheta}\}$$

and applied to the augmented system (15). The reduced matrices have the same structure as the ones in (15), with the same block-sparsity pattern, and in which each block is independently reduced (here subindex i, j indicate block position, not sensitivity index)

$$\begin{aligned} \widehat{C}_{ij} &= V_i^T C_{ij} V_j & \widehat{B}_i &= V_i^T B \\ \widehat{G}_{ij} &= V_i^T G_{ij} V_j & \widehat{E}_j &= E V_j. \end{aligned} \quad (16)$$

It is important to notice that if a sensitivity block has only zero entries, it will remain empty after the reduction, increasing the matrix sparsity.

Under a good sampling scheme, a small number of vectors is sufficient for accurately characterizing the system (whereas more samples are required to cover a multidimensional space). This approach also allows us to control the frequency sampling range, focusing on the more interesting areas or the more affected by the variation. *A posteriori* error bound can be used to control the error in every transfer function (see [10] for details). An algorithm is presented as Algorithm 1.

C. Stability and Passivity Preservation

An important property of any model reduction algorithm is the preservation of stability and passivity, which implies that the original model satisfies these constraints itself.

As already presented, a sufficient condition for passivity of an arbitrary system is that $B = E^T$ and C, G are positive semidefinite (PSD). The proposed representation for SPARE complies with the first condition. With respect to the second condition, for a real matrix G to be PSD, a necessary and sufficient condition is that its symmetric part $G^{\text{Sym}} = (G + G^T)/2$

Algorithm 1 SPARE-(PMTBR-based version)

From matrices $C_{00}, G_{00}, B, E, C_{\psi\vartheta}, G_{\psi\vartheta}$.

- 1: Select a quadrature rule of K points in the frequency space, and for each frequency point s_k :
- 2: Compute the LU decomposition: $LU = (s_k C_{00} + G_{00})$
- 3: For each transfer function, obtain the sampled vector

$$\begin{aligned} z_{00k} &= U^{-1}L^{-1}B \\ z_{10k} &= -U^{-1}L^{-1}(s_k C_{10} + G_{10})z_{00k} \\ z_{20k} &= -U^{-1}L^{-1}[(s_k C_{20} + G_{20})z_{00k} + \\ &\quad + (s_k C_{10} + G_{10})z_{10k}] \\ &\dots \end{aligned}$$

- 4: For each Transfer Function compose the matrices, and perform SVD on each one

$$\begin{aligned} Z_{\psi\vartheta} &= [z_{\psi\vartheta 1} \dots z_{\psi\vartheta k}] \\ Z_{\psi\vartheta} &= V_{\psi\vartheta} S_{\psi\vartheta} U_{\psi\vartheta} \end{aligned}$$

- 5: For each matrix $V_{\psi\vartheta}$ drop the columns whose singular values fall below the desired tolerance.
- 6: Build a Block Structure Preserving Projector from the remaining columns

$$\mathcal{V} = \text{diag}\{V_{00} \ V_{10} \ \dots \ V_{\psi\vartheta}\}$$

- 7: Apply \mathcal{V} in a congruence transformation on the augmented system $\mathcal{C}, \mathcal{G}, \mathcal{B}, \mathcal{L}$ in (15)

$$\widehat{C} = \mathcal{V}^T \mathcal{C} \mathcal{V}, \quad \widehat{G} = \mathcal{V}^T \mathcal{G} \mathcal{V}, \quad \widehat{B} = \mathcal{V}^T \mathcal{B}, \quad \widehat{E} = \mathcal{E} \mathcal{V}$$

be PSD. For a symmetric matrix to be PSD, a necessary and sufficient condition is that all its eigenvalues be nonnegative (see Section VI-C of [25]). We will use these results to prove that the SPARE matrices are PSD by construction.

Let us take a single parameterized electric element $a(\lambda)$ dependent upon a set of parameters λ , each of which is normalized to a maximum value $|\lambda_k| \leq 1$. The representation of the element in terms of λ is $a(\lambda) = a_{0\dots 0} + \sum_{\psi_1 \dots \psi_Q} \Lambda_{\psi_1 \dots \psi_Q} a_{\psi_1 \dots \psi_Q}$, with $\Lambda_{\psi_1 \dots \psi_Q} = \lambda_1^{\psi_1} \dots \lambda_Q^{\psi_Q}$. We will further assume that for every element, the following conditions are always met:

$$\begin{aligned} a_{00\dots 0} &\geq 0 \\ a_{00\dots 0} &\geq \sum_{\psi_1 \dots \psi_Q} |a_{\psi_1 \dots \psi_Q}|. \end{aligned} \quad (17)$$

While the first condition is obvious from a physical perspective, the second condition merely states that for any element, the perturbation caused by parameter variability has magnitude smaller than the nominal value. This is a reasonable assumption that ensures that under any parameter setting all element values remain positive, thus, physical.

Lemma 3.1: A matrix corresponding to a SPARE stamping of a single element, for arbitrary order and number of parameters, is PSD if the conditions in (17) hold for the element.

Proof: If we consider a single element a between two arbitrary nodes κ and γ , the symmetric part of the SPARE matrix \mathcal{A} for an arbitrary order and number of parameters, related to the stamping of this element is of the form

$$\mathcal{A}^{\text{Sym}}(a, \kappa, \gamma) = \begin{bmatrix} 2A_{11} & A_{21} & A_{31} & & \\ A_{21} & 2A_{11} & A_{32} & \dots & \\ A_{31} & A_{32} & 2A_{11} & & \\ \vdots & & & \ddots & \end{bmatrix} \quad (18)$$

where subindexes ij indicates block position, not sensitivity index. Each block ij has the corresponding stamping (nominal in the diagonal blocks, or sensitivity in the off diagonals). As an illustration, for the diagonal block

$$A_{11} = \begin{bmatrix} \kappa & \gamma \\ a_{0\dots 0} & -a_{0\dots 0} \\ -a_{0\dots 0} & a_{0\dots 0} \end{bmatrix} \quad \begin{matrix} \kappa \\ \gamma \end{matrix} \quad (19)$$

where $a_{0\dots 0} \geq 0$ is the nominal value of the element. The block is symmetric with eigenvalues $a_{0\dots 0}$ and zero, and therefore PSD. All the blocks in (18) have the same structure, with one zero eigenvalue, and another equal to the stamped value. Notice that in the SPARE formulation, the number of nonzero off-diagonal blocks in a column is at most equal to the number of sensitivities of (5), with each nonzero off-diagonal block corresponding to a different sensitivity.

The matrix in (18) can be rewritten as

$$\mathcal{A}^{\text{Sym}} = \mathcal{A}^D + \sum_{ij} \mathcal{A}_{ij}^O \quad (20)$$

where \mathcal{A}^D is a block diagonal matrix with

$$\text{diag}(\mathcal{A}^D, j, j) = A_{jj}^D = 2A_{11} - \sum_i A_{ij}^+ \quad (21)$$

where $A_{ij}^+ = A_{ij}$ if the block A_{ij} in (18) is PSD (i.e., the sensitivity is positive), or $A_{ij}^+ = -A_{ij}$ if the block A_{ij} is negative SD (i.e., the sensitivity is negative). It is straightforward that A_{ij}^+ is PSD, with its nonzero eigenvalue positive and equal to the module of the corresponding sensitivity value a_{ij} . The block matrices \mathcal{A}_{ij}^O in (20) are

$$\mathcal{A}_{ij}^O = \begin{matrix} & \begin{matrix} i & j \end{matrix} \\ \begin{bmatrix} A_{ij}^+ & A_{ij} \\ A_{ij} & A_{ij}^+ \end{bmatrix} & \begin{matrix} i \\ j \end{matrix} \end{matrix} \quad (22)$$

By the hypothesis of the conditions in (17), the blocks A_{ij}^D are PSD (the related element is always positive, and thus, the nonzero eigenvalue is positive), and thus, \mathcal{A}^D is PSD. Regarding \mathcal{A}_{ij}^O , the matrices are symmetric and singular. A block row and column elimination leads to a matrix with a single diagonal block, equal to A_{ij}^+ . The nonzero eigenvalues of this matrix are the same as the eigenvalues of A_{ij}^+ , which are nonnegative. Since the matrix is symmetric with nonnegative eigenvalues, we can state that it is PSD.

Therefore, the symmetric part of the stamping matrix, shown in (18), is as a sum of PSD matrices, and thus, a PSD matrix. ■

Lemma 3.2: The SPARE matrices \mathcal{G} and \mathcal{C} are PSD for an arbitrary order and number of parameters, for any number of elements, as long as every element complies with the conditions in (17).

Proof: For an arbitrary number of components the SPARE matrix can be generated by the sum of the matrices with the individual stamping of each component. Each one of these individual matrix is PSD, and thus, their sum is a PSD matrix. ■

The above results can be applied to any formulation that complies with the assumptions presented. Some methods

may include incidence blocks [e.g. modified nodal analysis (MNA) stamping of an inductor]. In this case the sensitivities of the incidence values are zero (the parameters affect the value of the element, not its connections). Furthermore, such blocks are zeroed out when computing the symmetric part of the diagonal blocks (see (26) in [6]).

The consequences of these results, subject to the conditions in (17) limit the acceptable range of variation for the parameters so that the maximum perturbation of an element is still smaller than the nominal value. While this could theoretically be conceived as a limitation of the present technique, in most cases this condition is reasonable if we take into account the physical nature of the problem (perturbations larger than the nominal value could lead to negative elements). Furthermore, most extraction methods only provide first order sensitivities, and thus, this limitation is no more restrictive than the one proposed in Section II-B.

With respect to the passivity of the SPARE ROM, the congruence projection preserves the PSD character of the matrices. Therefore the reduced matrices maintain the same structure and properties as the original, and since $\hat{\mathcal{B}} = \hat{\mathcal{E}}^T$, the ROM is guaranteed passive subject to the same restrictions as the original SPARE representation.

IV. COMPUTATIONAL CONSIDERATIONS

A. Reduced Model Generation

The cost of the procedure depends on the underlying nominal MOR scheme chosen. We will study the case of PMTBR based approaches. For K frequency sampling points, we must perform K factorizations (one for each frequency point s_k) of the matrix $A_k = Cs_k + G$. For sparse matrices, the cost of each factorization is $O(n^\beta)$, with n the size matrix, and typically $1.1 \leq \beta \leq 1.5$. Once the matrix factors are obtained, we must perform a set of matrix operations to generate the nominal vector x_0 . These operations have an associated cost of $O(n^\alpha)$, with $1 \leq \alpha \leq 1.2$ for sparse matrices. For each extra vector x_i we must perform P (with P the number of transfer functions sensitivities at the output) similar matrix operations, with a cost of $O(n^\alpha)$. Once all the vectors are computed for the various transfer functions, the SVD is applied, at a cost of $O(nq_i^2)$, with q_i the reduced order for each transfer function. Thus, the total cost of the procedure is

$$O(Kn^\beta + K(P+1)n^\alpha + \sum_{i=0}^P nq_i^2). \quad (23)$$

Let us compare this result with the total cost of the variational PMTBR (VPMTBR) approach [20], which is the closest methodology. In the VPMTBR, \tilde{K} samples must be taken in frequency plus parameters subspace. For each point, a factorization and solve is required to generate the vectors, with a cost $O(n^\beta + n^\alpha)$. A SVD must be applied to the vectors to obtain the basis, with cost $O(nq^2)$, in which q is the global number of vectors. The global cost is

$$O(\tilde{K}n^\beta + \tilde{K}n^\alpha + nq^2). \quad (24)$$

When comparing (23) and (24) we must notice that for the VPMTBR approach, the number of samples \tilde{K} is larger

($\tilde{K} > K$ as we must sample a higher dimensional space), and the size of the ROM, q , is (as we will see in the next sections) at least equal to the size of the SPARE global matrix ($\sum_{i=0}^p q_i$). Knowing that the dominant factor in the cost of MOR algorithms comes from the factorizations, it is clear that the cost of the SPARE algorithm is smaller than for the VPMTBR approach. This assessment will be validated with experimental results in the following sections.

We would also point out, although beyond the scope of this paper, that the procedure is highly parallelizable, in particular the PMTBR-based approach, as the computation of the vectors at different sampling points is completely independent. Furthermore, the recursive procedure is independent for every parameter (as long as no cross terms are taken into account). Therefore, different levels of parallelism can be taken into account. For example the frequency sampling level, as the vectors x_i can be generated independently for every one of the K frequency points s_k . Another is the order level, as for each sampling point, the vector of each transfer function level only depends on previous levels (see Fig. 1). Parallelization can also be applied on the orthonormalization stage, since independent SVD are applied to the multiple basis generated, in order to obtain the orthonormalized vectors.

Orthogonalization between the different sets of vectors (i.e., V_0, V_1, \dots) is not an issue, as they are applied to different transfer functions. However, for high orders in the sensitivities with respect to the parameters, differences in the condition number of the nominal matrices and the sensitivities can lead to numerical errors. To avoid such pitfalls, a simple scheme for scaling the matrices can be performed. Another consideration we have to take into account is the fact that we are treating every transfer function H_i independently, and thus, the parameter and block related to such transfer function can be scaled independently without consequences both before and after SPARE reduction.

A different issue may arise from the use of Taylor series approximation of the parameterized behavior. Taylor series were chosen here for simplicity and because it is the de facto standard in pMOR, for which lower orders are usually sufficient for capturing the parameterized behavior. However, it is important to notice that the methodology proposed can be combined with any representation of the parameterized behavior. For example, it can be combined with functions depending on several parameters

$$\begin{aligned} G(\lambda) &= G_0 + f_1(\lambda)G_1 + f_2(\lambda)G_2 + \dots \\ C(\lambda) &= C_0 + f_1(\lambda)C_1 + f_2(\lambda)C_2 + \dots \\ H(s, \lambda) &= H_0(s) + f_1(\lambda)H_1(s) + f_2(\lambda)H_2(s) + \dots \end{aligned} \quad (25)$$

where $f_1(\lambda)$ and $f_2(\lambda)$ can be any function depending on the set (of subsets) of parameters λ . This fact opens the possibility of using different representations of the parameterized systems, based on approximations of quasi-linear behavior at the output via regression, orthogonal basis, kernel functions, etc. Good representations could lead to better accuracy for lower order at the output, or to parameter clustering (capture the behavior of several physical parameters via a single numerical function). Although interesting for future studies, these issues are beyond the scope of this paper.

B. Model Evaluation

As it has been already pointed out, the preservation of the structure in the reduction step allows for faster evaluation of the model. The typical Taylor series reduced models are described by a set of full reduced matrices of size q . To solve the system for each frequency (or time step in time-domain simulations), we must evaluate the system for the parameter set λ_k , and then solve it for the frequency value s_k (or equivalent time point). This implies solving (3) with $\{s_k, \lambda_k\}$, which has an associated cost dominated by $O(q^3)$ since the model is now essentially full. Every time a parameter or the frequency change, a complete solve must be done.

The block lower triangular matrices in the SPARE models can be exploited in a recursive fashion (although diagonal blocks are different), and thus, for each frequency (or time step in time-domain simulations), we must solve a number of nominal transfer functions P , each of order q_i . For instance, in the frequency domain and for first order with one parameter

$$\begin{aligned} \hat{x}_{00}(s_k) &= (\hat{G}_{00} + s_k \hat{C}_{00})^{-1} \hat{B}_{00} \\ \hat{H}_{00}(s_k) &= \hat{E}_{00} \hat{x}_{00} \\ \hat{H}_{10}(s_k) &= \hat{E}_{10} (\hat{G}_{00} + s_k \hat{C}_{00})^{-1} (\hat{G}_{10} + s_k \hat{C}_{10}) \hat{x}_{00} \end{aligned} \quad (26)$$

which requires two solves with full block matrices and a matrix-vector multiplication: $O(q_0^3 + q_1^3 + q_0 q_1)$, with q_0 and q_1 the sizes of the blocks, respectively. In the general case the associated cost is dominated by $O(\sum_{i=0}^p q_i^3)$, in which the sizes q_i are smaller. For the same matrix size (i.e., $q = \sum_{i=0}^p q_i$), it is clear that the cost of using the SPARE models is smaller than for the traditional full Taylor series models: $O(\sum_{i=0}^p q_i^3) < O\left(\left(\sum_{i=0}^p q_i\right)^3\right)$.

The parameterized response is obtained by linear combination of the multiple transfer functions, as shown in (12), the cost of which is negligible. Furthermore, any change of a parameter (value setting) simply requires evaluating such linear combination, with no need for system solving.

V. EXPERIMENTAL RESULTS

A. Subspace Growth

In this section, we illustrate the effect that parameters have on the subspaces generated for projection. For several different parameterized systems a fairly exhaustive sampling is performed. First, the nominal case is considered, and only the frequency is sampled, with all the parameters set to their nominal value. Second, we consider the parameter effect, and thus, sampling is done in the complete frequency plus parameter space (including cross samples). In each case, for each system, an SVD is applied to the resulting set of vectors, and a threshold is set allowing truncation of the respective subspace. All vectors associated with singular values within $1e-6$ of the largest singular value (σ_1) are kept. The size of the resulting dominant subspace is then compared. The examples used in this experiment are given below.

- 1) Two electromagnetically coupled lossy lines on the same metal layer. They are subject to the variation of six geometrical parameters, modifying the width and thickness

TABLE I
SUBSPACE GROWTH WITH PARAMETER EFFECT

Example	Coupled Lines (Size 6002)	PEEC (Size 304)	$L_{SP} + C_{MIM}$ (Size 11005)
Nominal	26	38	58
Parameterized	54 (six parameter)	63 (one parameter)	446 (two parameter)

of the metal, each parameter inducing a variation up to $\pm 30\%$ element-wise. Each line and the couplings are modeled via distributed elements, with a total model order of 6002 states.

- 2) A well known 304 states partial element equivalent circuit (PEEC) based example [7], in which the conductive, capacitive and inductive parts are perturbed up to 30% with a single artificial parameter.
- 3) A full wave EM model of a typical industrial analog planar spiral inductor connected in series with a metal-insulator-metal capacitor (CMIM), including surrounding and substrate. The model, obtained via finite integration technique, FIT [28], has a size of 11 005, and is dependent on two parameters. One is the length of the side of the square that forms the spiral, and the other is the insulator width. First order sensitivities relative to these parameters are obtained for G and C [22], [24].

Table I shows (for each of the examples) the size of the relevant subspace for the nominal system, and for the complete parameterized system. It can be seen that perturbations on simple examples, such as the coupled lines and PEEC, although they may largely affect the output response, generate few extra vectors to add to the subspace, and thus, such subspace can be efficiently matched with standard pMOR techniques. However, for the EM example, the effect of the perturbations drastically increases the number of vectors of the relevant subspace. In this scenario, the standard pMOR techniques must match this subspace in order to maintain the accuracy for the complete range of variation of the parameters. Therefore, the size of the ROM is drastically increased (it should be at least of the same size), and its full matrices may lead to small advantage in simulation steps over the original unreduced model.

In the next sections, we benchmark the proposed methodology against competing algorithms: the procedures presented in [20], denoted as VPMTBR, in [21], denoted as CORE, and in [19], denoted as parameterized interconnect macromodeling via a two-directional Arnoldi process (PIMTAP). The advocated methodology will be combined with either moment matching or PMTBR underlying schemes for generating the projector, and will be denoted as K-SPARE and BT-SPARE, respectively. The nonreduced model response, formulated as in (5), will be denoted as nominal or parameterized, depending on whether a variation of the parameters has been applied. In all cases, first order Taylor series sensitivities will be used, both for matrices and for the transfer functions.

B. First Order Approximation: Coupled Lines Example

The first example will be the two coupled lossy metal lines. The figure of merit selected for this benchmark is to capture the cross-talk between lines, i.e., the transfer function between the input of the first line, and the output of the second line.

TABLE II
 $n = 6002$ -STATES COUPLED LINES EXAMPLE: pMOR FEATURES

	NNZ (G C) & Sparsity	Gen. Effort	Speed-Up (One Eval)
Original size 6002	16 000 8000 4e-4 2e-4	– –	1×
VPMTBR [20] size 28	784 1.00	100 Spl.(w+ λ) SVD($n \times 28$)	$\approx 20\times$
PIMTAP [19] size 28	784 1.00	2 BM QR($n \times 28$)	$\approx 20\times$
CORE [21] size 28	784 1.00	14 BM QR($7n \times 28$)	$\approx 20\times$
K-SPARE size 98 (14)	2548 0.265	7 BM $7 \times \text{QR}(n \times 14)$	$\approx 20\times$
BT-SPARE size 98 (14)	2548 0.265	25 Spl.(w) $7 \times \text{SVD}(n \times 14)$	$\approx 20\times$

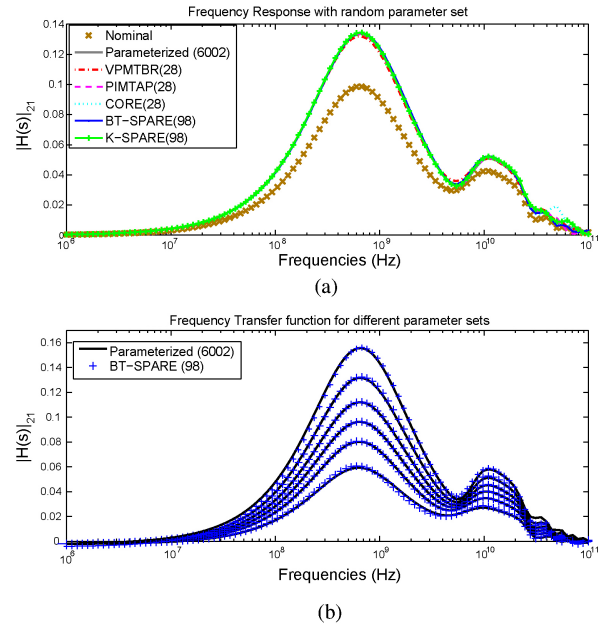


Fig. 2. Coupled lines example. (a) Transfer function for nominal and parameterized models, and ROMs for a single parameter point. Curves are virtually indistinguishable, although Krylov based models lose accuracy at high frequencies. (b) Transfer function for the parameterized and BT-SPARE models for different parameter settings.

The geometrical parameters largely affect this coupling effect between lines.

For this first comparison, we fix the order for the VPMTBR to obtain a model with very good accuracy (absolute error less than $1e-3$). In this case, the order is set to 28. The orders for the rest of the pMOR models are obtained so that the speed-up for a single evaluation is approximately the same. In the case of the PIMTAP and CORE approaches, the size of the ROM must be the same, whereas for the SPARE based approaches we fix the block size to 14 for each transfer function, which generates a system with 98 states and matrices with a block lower triangular (BLT) structure. The effort in evaluate these models for a single point is approximately the same than for an order 28 full matrix (see Section IV-B for details). Table II shows the relevant characteristics of the models and algorithms used, namely the number of nonzero elements in the model (NNZ), the sparsity factor (ratio of NNZ to size), the cost of model generation (Gen. Effort, in which BM stands for block moments, and Spl. stands for samples) and evaluation speed-

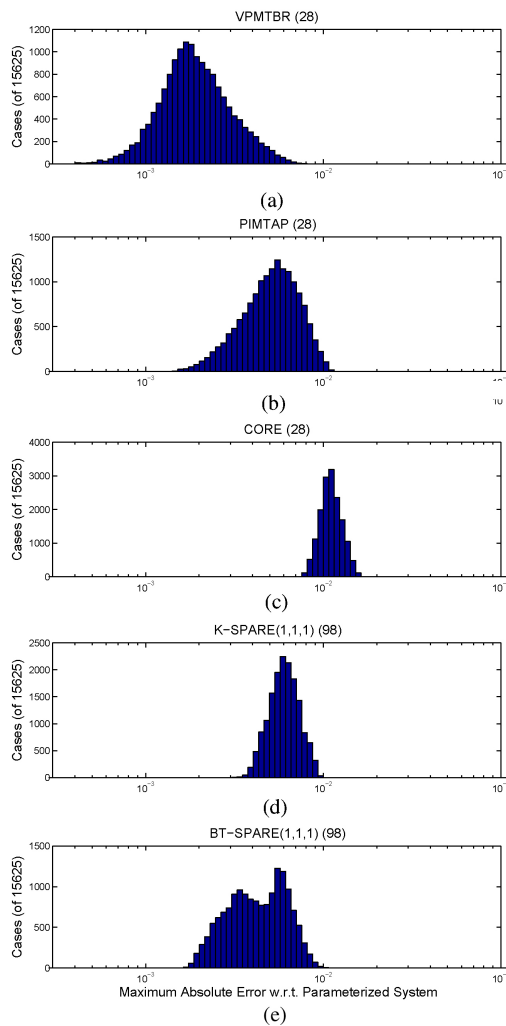


Fig. 3. Coupled lines example. Monte Carlo analysis on 15 625 parameter settings. On the vertical axis are the number of occurrences, whereas on the horizontal axis are the maximum relative error for the complete frequency range with respect to the original parameterized model. From (a) to (e), the models are VPMTBR, PIMTAP, CORE, K-SPARE, and BT-SPARE, respectively.

up (a ratio of the time spent evaluating the reduced versus the original models, and which is related to discussion on Section IV-B). In the case of PIMTAP, CORE, K-SPARE, and BT-SPARE, the ROMs are obtained for first order with respect to the parameters (linear approximation). The top plot (a) in Fig. 2 shows the frequency transfer function for the nominal, the perturbed Taylor series original system and various ROMs, for a single parameter setting. All the pMOR methodologies capture the parameterized behavior with good accuracy for this setting. The bottom plot (b) in the same figure shows how the transfer function varies for different parameter settings, and how the BT-SPARE approach is able to capture this behavior.

Let us now study the behavior of the ROMs when the parameters are varied to different settings. To this end, we perform a Monte Carlo (MC) analysis with 15 625 parameter settings covering all the space (and for each parameter setting a fine frequency sweep is done). For each ROM we calculate the transfer function for the current parameter setting, and obtain the maximum absolute error for the complete frequency sweep with respect to the original parameterized model. Fig. 3

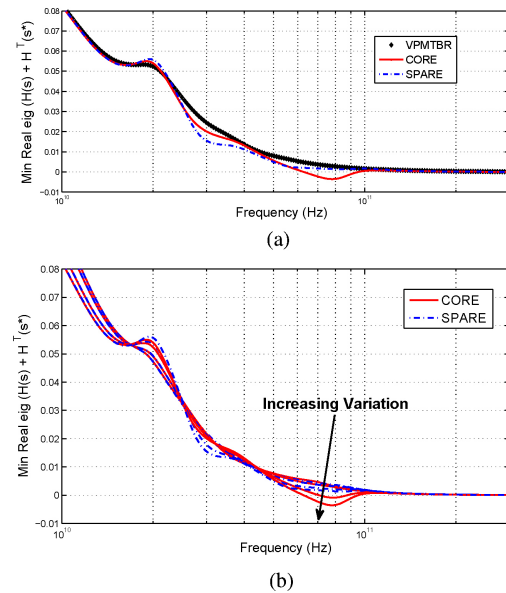


Fig. 4. Coupled lossy lines example. (a) Minimum real part of eigenvalues of $(H(s) + H^T(s^*))$ along the frequency for a single parameter set. (b) Minimum real part of eigenvalues of $(H(s) + H^T(s^*))$ as the parameter values increase.

shows, for each methodology, the number of parameter settings against the maximum absolute error of the ROM transfer function in the whole frequency sweep. It can be seen that VPMTBR is the most reliable framework, with a very low error for all the samples. BT-SPARE also generates good results, due to the sampling scheme that provides better frequency-wise overall models. Krylov based approaches lose accuracy at high frequencies due to the inherent locality of the methodologies, although for the frequency range studied the accuracy provided by the number of moments matched is good. K-SPARE and CORE approaches rely on capturing the transfer function parameterized behavior by a Taylor series approximation (linear in this case). For this reason, the minimum error will be the one generated by the nominal, and for any parameter variation, the error will increase by the difference between the real output behavior and the approximation. This is why the largest error in these approaches is higher. However, the accuracy of the K-SPARE approach is better than the one of the CORE model, even though the number of block moments generated is smaller (see Table II). PIMTAP model applies the projector with multidimensional moments on the parameterized system, and in this case it provides a good agreement even for a small number of block moments.

C. Passivity Preservation

In this section, we will present results that validate the theoretic results for passivity preservation. To this end, we use some of the models generated in the previous section, that is, the ROMs obtained via VPMTBR, CORE and SPARE for the coupled lossy lines example. The VPMTBR retains the original matrix Taylor series formulation, and in consequence its passivity attributes.

For these models, we vary the parameter values and check the passivity of the resulting model. For zero and small variations all the models, including CORE, remain passive.

TABLE III
 $n = 11005$ -STATES EM EXAMPLE: PMOR FEATURES

	NNZ (G C) and Sparsity	Gen. Effort	Speed-Up (One Eval)
Original size 11 005	48 708 13 510 4e-4 1.1e-4	— —	1×
VPMTBR size 150	22 500 22 500 1.00 1.00	75 Spl.(w+λ) SVD($n \times 150$)	≈250×
PIMTAP size 149	22 201 22 201 1.00 1.00	25 BM QR($n \times 149$)	≈250×
CORE size 150	22 500 22 500 1.00 1.00	75 BM QR($3n \times 150$)	≈250×
K-SPARE size 150 (50)	12 500 12 500 0.55 0.55	25 BM 3 × QR($n \times 50$)	≈800×
BT-SPARE size 150 (50)	12 500 12 500 0.55 0.55	25 Spl(w) 3 × SVD($n \times 50$)	≈800×

However, as we increase the parameter variation, we have noticed that the CORE models quickly become nonpassive. This could be understood by noticing the larger the parameter value, the larger the difference between B and E matrices for the CORE model (recall that CORE included the parameter effect in the E matrix). On the other hand, both SPARE and VPMTBR models remain passive for any perturbation (as long as they comply with the presented theoretical results).

From the conditions of the positive real lemma (6), for a passive system the real part of the eigenvalues of $H(s) + H^T(s^*)$ should be positive for any $s \in \mathbb{R}(s) > 0$. Fig. 4 shows the minimum real part of the eigenvalues of $(H(s) + H^T(s^*))$ for all the frequency points. The top plot shows this curve for the three perturbed models. It is clear that the CORE curve falls below zero, and thus, it is nonpassive. The bottom plot presents the same curve for SPARE and CORE models for different parameter sets, starting at zero variation (nominal) and increasing until maximum variation. It can be seen that the CORE model becomes nonpassive as the parameter values are increased.

D. Models of Same Order: EM-Based Example

As a second benchmark we use the spiral inductor and CMIM EM-based model, which depends on two parameters. The same algorithms are used here: VPMTBR, PIMTAP, CORE, and SPARE based methodologies. In all cases the ROMs are obtained for first order with respect to each parameter (linear approximation). Table III shows the same relevant characteristics of the models and algorithms used. The sizes of all the ROMs are set to 150 for comparison. The effort needed for generating such models varies: Krylov based techniques, such as CORE and K-SPARE, are cheaper. However, the number of BM computed for K-SPARE is one third of that computed in CORE (see Table III), as the basis is expanded. The cost of BT-SPARE is lower than that of VPMTBR, as it only samples the frequency, and the SVD is applied on a much smaller number of vectors. The VPMTBR model is computed using the same frequency samples as BT-SPARE, plus parameter samples around those frequency points. The evaluation of the models is again much faster in the SPARE-based models, as their matrices are sparser and block lower triangular (see Table III, where the ROM, and the blocks sizes in the SPARE case are shown below the method name) whereas VPMTBR, PIMTAP, and CORE yield full models. Furthermore, any

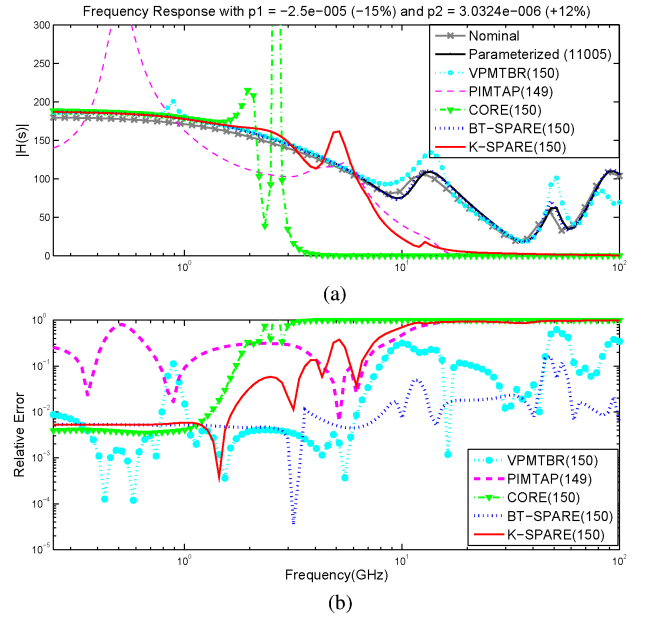


Fig. 5. Spiral and CMIM example. $|Z_{11}|$ versus frequency for a single parameter set. (a) Transfer function for nominal, perturbed TS and ROMs. (b) Relative error with respect to TS for the ROMs.

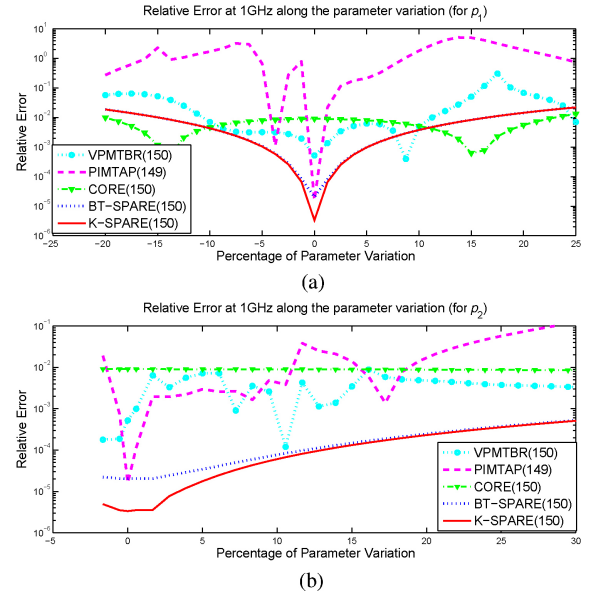


Fig. 6. Spiral and CMIM example. relative error of the ROMs versus the parameter variation for (a) p_1 and (b) p_2 at 1 GHz.

change in the parameter values implies a complete reevaluation of the model for all techniques but SPARE. For SPARE this cost is negligible and, therefore, for N different parameter settings, the speed-up increases by a factor of N . With respect to the accuracy of the models, Fig. 5 shows the frequency transfer function computed by all methods, for a given setting of the parameters. PIMTAP is unable to capture the parameterized behavior, as the first order approximation is not enough to produce all the moments the Taylor series system needs for being accurately modeled. On the other hand, CORE and K-SPARE, which rely on an approximation of the output, show good agreement for low frequencies, but with increasing error as frequency rises (more frequency moments would improve

TABLE IV

 $n = 11\,005$ -STATES EM EXAMPLE: PMTBR PMOR FEATURES

	NNZ (G C) and Sparsity	Gen. Effort	Speed-Up (One Eval)
VPMTBR size 256	65 536 1.00	75 Spl.(w+ λ) SVD($n \times 256$) $O(8.13e8)$	$\approx 55 \times$
VPMTBR size 377	142 129 1.00	125 Spl.(w+ λ) SVD($n \times 377$) $O(1.71e9)$	$\approx 15 \times$
BT-SPARE(1, 1) size 120 (57, 33, 30)	8829 0.613	25 Spl.(w) SVD($n \times 57$) SVD($n \times 33$) SVD($n \times 30$) $O(9.18e7)$	$\approx 850 \times$
BT-SPARE(2, 1) size 158 (57, 33, 38, 30)	11 527 0.461	25 Spl.(w) +SVD($n \times 38$) $O(1.09e8)$	$\approx 640 \times$
BT-SPARE(3, 1) size 193 (57, 33, 38, 35, 30)	14 082 0.378	25 Spl.(w) +SVD($n \times 35$) $O(1.24e8)$	$\approx 540 \times$

their accuracy). VPMTBR is not able to maintain the accuracy for parameter variations whereas BT-SPARE of first order exhibits the best accuracy for the whole frequency range and for a fairly large variation on the parameters. This behavior can be seen in Fig. 6, which shows the transfer function variation with respect to parameters (a) p_1 and (b) p_2 , for a fixed frequency point. It is clear that the SPARE based models are able to capture the parameterized behavior for fairly large variations. CORE exhibits a similar behavior, but loses accuracy as frequency rises. The accuracy of VPMTBR varies and deteriorates when the variations increase (for the same ROM size). p_2 has a less relevant effect, and all the algorithms are able to capture it accurately, although the SPARE based techniques display a *smoother* behavior.

E. PMTBR Based Approach: Automatic Order Selection

In this section, we are going to compare the PMTBR based approaches, i.e., the VPMTBR [20] approach versus the PMTBR based SPARE (BT-SPARE). In this case, the size is determined by singular values. This means that we set a defined tolerance (in this case, a relative tolerance of $5e-5$), and the singular vectors whose associated singular value falls below such tolerance, are dropped.

For VPMTBR two models are generated. The difference between them is the number of samples taken for the model generation. In both cases, a logarithmic sampling of 25 points in the frequency is done (with parameter values set to nominal). For the first VPMTBR model, around those nominal frequency points, two random samples are taken in the parameter space [this gives $25 \times (1 + 2) = 75$ samples]. After the SVD, a 256-vector basis is obtained. This means a 256 full model. For the second VPMTBR model, around those nominal frequency points, four random samples are taken in the parameter space [this gives $25 \times (1 + 4) = 125$ samples]. After the SVD, a 377 full model is generated.

For BT-SPARE, three models are generated, with different orders with respect to each parameter. In the three cases we use the same 25 nominal samples of the VPMTBR scheme. Each model is obtained by increasing the output order with respect to one parameter. The first model, denoted as BT-SPARE(1, 1)

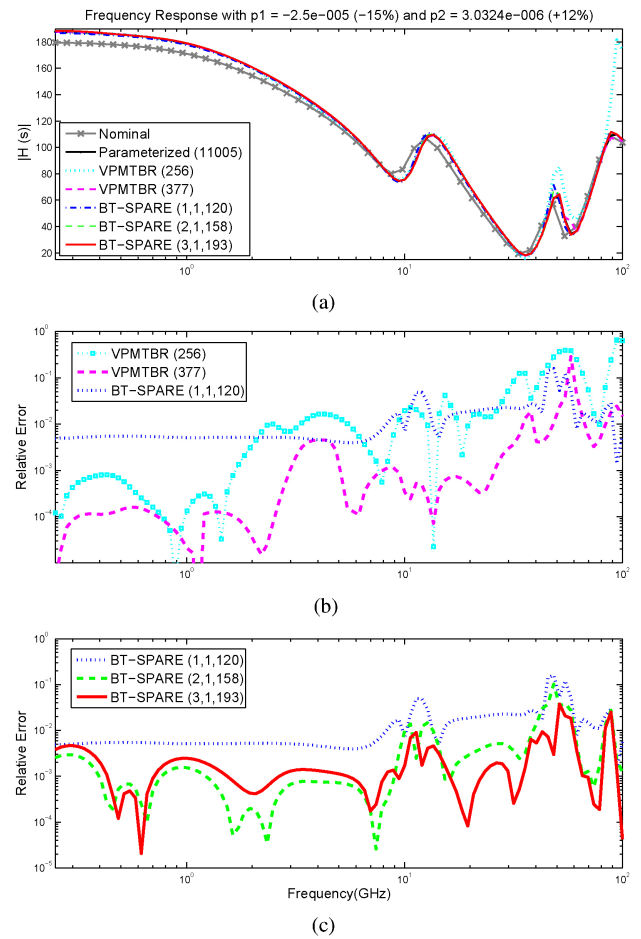


Fig. 7. SPIRAL and CMIM EM example. (a) $|Z_{11}|$ for the nominal, the perturbed TS, and the ROMs, along the frequency for a single parameter set. (b) Relative error along the frequency of the VPMTBR (256), the VPMTBR (377) and BT-SPARE(1, 1) ROMs with respect to the original TS, for the same parameter set. (c) Relative error along the frequency of BT-SPARE (1, 1), BT-SPARE(2, 1) and BT-SPARE(3, 1) ROMs with respect to the original TS, for the same parameter set.

is a first order approximation with respect to each parameter. After the SVD is applied to each set of vectors, the number of retained vectors are: 57 for the nominal transfer function, 33 for the first order with respect to the first parameter, and 30 for the first order with respect to the second parameter. This generates BLT G and C matrices of size 120. The second model, denoted as BT-SPARE(2, 1) is a second order w.r.t to the first parameter and a first order with respect to the second parameter approximation. The model is incrementally obtained from the previous by including the second order with respect to the first parameter. This adds a 38-order block, and thus, complete BLT G and C matrices of size 158. For the third model, denoted as BT-SPARE(3, 1), we incrementally add the third order w.r.t to the first parameter. This adds a 35-order block, and generates complete BLT G and C matrices of size 193.

Table IV shows the same characteristics of the models as in the previous examples, for the current study case. In the column related to generation cost, we have included the incremental cost for increasing the order, and we have added a row with the theoretical cost for each of the ROMs, by simply evaluating (24) and (23). It is clear that the VPMTBR cost is higher than the SPARE methodology.

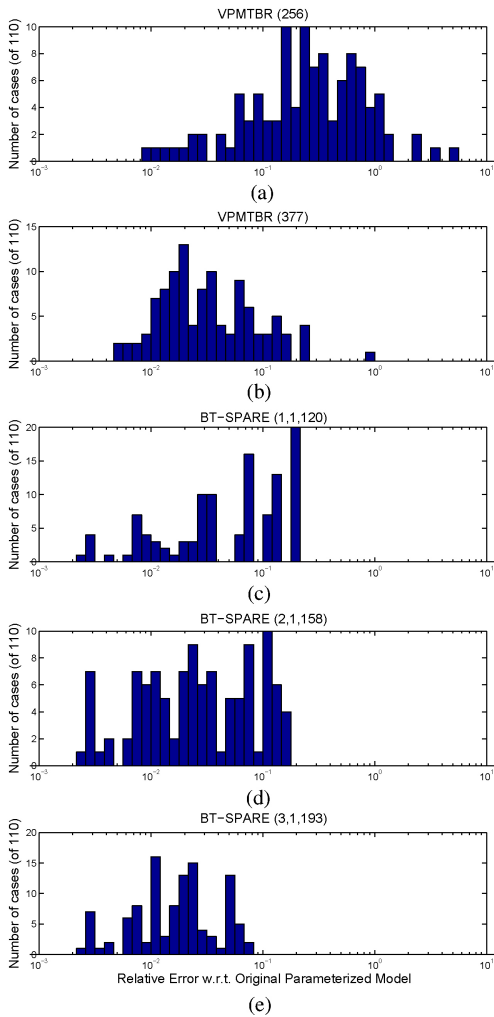


Fig. 8. SPIRAL and CMIM EM example. MC analysis on 110 parameter settings. On vertical axis are the number of occurrences, whereas on horizontal axis are the maximum relative error with respect to the original TS for the complete frequency range. ROMs are, from (a) to (e), VPMTBR (size 256), VPMTBR (size 377), BT-SPARE(1, 1) (size 120), BT-SPARE(2, 1) (size 158), and BT-SPARE(3, 1) (size 193).

Fig. 7 shows the frequency response of the ROMs’ transfer functions with respect to the original TS, for a single parameter set. The smaller (256) VPMTBR model does not have sufficient accuracy. VPMTBR of size 377 and BT-SPARE approaches are apparently indistinguishable. However, increasing the order with respect to the parameter leads to an accuracy improvement in SPARE based approaches, which are better than those of the VPMTBR models. Relative errors [(b) middle and (c) bottom plots of Fig. 7] confirm these results.

The advantage of SPARE in this scenario comes from the fact that the system output behavior when the parameters are modified is smooth, and thus, can be well approximated by a Taylor series representation, for which the nominal and sensitivities transfer functions are well approximated by low order models. On the other hand, the VPMTBR model needs to capture the complete subspace (which is much larger due to the effect of the parameters).

To see the behavior of such models, a Monte Carlo like simulation is applied and 110 different parameter settings, covering the range of interest, are simulated. For each set, the

frequency transfer function is computed for each ROM, and the maximum relative error with respect to the original Taylor series is computed. Fig. 8 shows, for each ROM, the number of parameter settings against the maximum relative error of the ROMs’ transfer function for the complete frequency range. It can be seen that for the two VPMTBR models (top two figures in Fig. 8) the error is high. Increasing the number of samples leads to accuracy improvement, but the 125 samples are not sufficient. Increasing the number of samples, or performing an optimal sample, would lead to better results. Regrettably, there is no optimal methodology for multidimensional sampling, and the size of the ROM would still make the model uncompetitive if compared to SPARE.

The SPARE models have a much better behavior. It can be seen that the increment of order with respect to the first parameter improves the worst case accuracy [for the BT-SPARE(3, 1) model, the worst relative error was approximately 7%]. A noticeable effect of the SPARE approaches is a relatively large deviation in the error plot. This is because SPARE relies on capturing the transfer function parameterized behavior by a Taylor series approximation. Thus, the smaller error will be the one generated by the nominal, and for any parameter variation, the error will increase by the difference between the real output behavior and the approximation. Therefore, the error for small perturbations will be lower than for large ones, but still the error will grow gradually with the increment in the perturbation, as long as the behavior on the output is smooth.

VI. CONCLUSION

In this paper, a flexible and efficient pMOR algorithm was presented. The method is based on a reformulation of the original system as a parallel interconnection of the nominal transfer function and the nonparameterized transfer function that describe the effect of the various parameters. This formulation reveals an explicit dependence on the parameters which is exploited and preserved during reduction. The structure of the reduced model leads to efficient simulation and re-evaluation of the parameterized response, with smaller time and memory resource requirements in simulations stages. The explicit parameter dependence also allows for better accuracy control, enabling independent adaptive order determination with respect to each parameter and adding flexibility in simulation environments. The new technique can handle fairly large parameter variations on systems whose outputs exhibit smooth dependence on the parameters. This can be used as a fast linear approximation to explore design parameters in optimization steps using EM models. The procedure has also been shown to preserve passivity, and examples show that besides the added flexibility and control, the proposed algorithm can produce smaller reduced models with potential accuracy gains in comparison with competing methods.

On the other hand, the technique is less competitive when dealing with nonsmooth output behavior with respect to the parameters, due to the Taylor series approximation. Future work should explore different output representations and techniques for parameter clustering, in order to obtain accurate parameterized output representation with a minimum number of parameterized terms.

ACKNOWLEDGMENT

The authors would like to thank the Chameleon RF Consortium, and in particular, W. Schilders and R. Janssen from NXP, Eindhoven, The Netherlands, and G. Ciuprina and D. Ioan from Numerical Methods Laboratory, Politehnica University of Bucharest, Romania, for many helpful discussions and for providing some of the simulation examples. They would also like to thank the reviewers for their detailed comments and helpful suggestions, as well as C. Coelho from Cadence Research Labs, Berkeley, CA, and M. Rewienski from Synopsys, Mountain View, CA, for helping with several results in this paper.

REFERENCES

- [1] J. F. Villena and L. M. Silveira, "SPARE: A scalable algorithm for passive, structure preserving, parameter-aware model order reduction," in *Proc. Design Autom. Test Eur. Conf. Exhibition*, Mar. 2008, pp. 586–591.
- [2] W. H. A. Schilders, H. A. Van der Vorst, and J. Rommes, Eds., "Model Order Reduction: Theory, research aspects and applications," *Mathematics in Industry* (The European Consortium for Mathematics in Industry Subseries), vol. 13. New York: Springer, 2008.
- [3] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.
- [4] C. P. Coelho, J. R. Phillips, and L. M. Silveira, "A convex programming approach for generating guaranteed passive approximations to tabulated frequency-data," *IEEE Trans. Comput.-Aided Design*, vol. 23, no. 2, pp. 293–301, Feb. 2004.
- [5] K. C. Sou, A. Megretski, and L. Daniel, "A quasi-convex optimization approach to parameterized model order reduction," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 3, pp. 456–469, Mar. 2008.
- [6] A. Odabasioglu, M. Celik, and L. T. Pileggi, "PRIMA: Passive reduced-order interconnect macromodeling algorithm," *IEEE Trans. Comput.-Aided Design*, vol. 17, no. 8, pp. 645–654, Aug. 1998.
- [7] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 14, no. 5, pp. 639–649, May 1995.
- [8] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Automatic Control*, vol. 26, no. 1, pp. 17–32, Feb. 1981.
- [9] J. Phillips, L. Daniel, and L. M. Silveira, "Guaranteed passive balancing transformations for model order reduction," *IEEE Trans. Comput.-Aided Design*, vol. 22, no. 8, pp. 1027–1041, Aug. 2003.
- [10] J. R. Phillips and L. M. Silveira, "Poor man's TBR: A simple model reduction scheme," *IEEE Trans. Comput.-Aided Design*, vol. 24, no. 1, pp. 43–55, Jan. 2005.
- [11] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *Am. Inst. Aeronautics Astronautics J.*, vol. 40, no. 11, pp. 2323–2330, Nov. 2002.
- [12] J. R. Singler, "Approximate low rank solutions of Lyapunov equations via proper orthogonal decomposition," in *Proc. Am. Control Conf.*, Jun. 2008, pp. 267–272.
- [13] Y. Liu, L. T. Pileggi, and A. J. Strojwas, "Model order reduction of RC(L) interconnect including variational analysis," in *Proc. 36th Assoc. Comput. Machinery/IEEE Design Autom. Conf.*, Jun. 1999, pp. 201–206.
- [14] P. Heydari and M. Pedram, "Model reduction of variable-geometry interconnects using variational spectrally-weighted balanced truncation," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2001, pp. 586–591.
- [15] L. Daniel, O. C. Siong, S. C. Low, K. H. Lee, and J. K. White, "A multiparameter moment matching model-reduction approach for generating geometrically parameterized interconnect performance models," *IEEE Trans. Comput.-Aided Design*, vol. 23, no. 5, pp. 678–693, May 2004.
- [16] P. Li, F. Liu, X. Li, L. Pileggi, and S. Nassif, "Modeling interconnect variability using efficient parametric model order reduction," in *Proc. Design Autom. Test Eur. Conf. Exhibition*, Feb. 2005, pp. 958–963.
- [17] P. Gunupudi, R. Khazaka, M. Nakhla, T. Smy, and D. Celso, "Passive parameterized time-domain macromodels for high-speed transmission-line networks," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 12, pp. 2347–2354, Dec. 2003.
- [18] Z. Zhu and J. Phillips, "Random sampling of moment graph: A stochastic krylov-reduction algorithm," in *Proc. Design Autom. Test Eur. Conf. Exhibition*, Apr. 2007, pp. 1502–1507.
- [19] Y. Li, Z. Bai, Y. Su, and X. Zeng, "Model order reduction of parameterized interconnect networks via a 2-D Arnoldi process," *IEEE Trans. Comput.-Aided Design*, vol. 27, no. 9, pp. 1571–1582, Sep. 2008.
- [20] J. Phillips, "Variational interconnect analysis via PMTBR," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2004, pp. 872–879.
- [21] X. Li, P. Li, and L. Pileggi, "Parameterized interconnect order reduction with explicit and implicit multiparameter moment matching for inter/intra-die variations," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2005, pp. 806–812.
- [22] Y. Bi, K.-J. Van der Kolk, D. Ioan, and N. van der Meijs, "Sensitivity computation of interconnect capacitances with respect to geometric parameters," in *Proc. IEEE Int. Conf. Electr. Performance Electron. Packag.*, Oct. 2008, pp. 209–212.
- [23] T. El-Moselhy, I. Elfadel, and L. Daniel, "A capacitance solver for incremental variation-aware extraction," in *Proc. IEEE/Assoc. Comput. Machinery Int. Conf. Comput.-Aided Design*, Nov. 2008, pp. 662–669.
- [24] G. Ciuprina, D. Ioan, D. Niculae, J. F. Villena, P. Flores, and L. Silveira, "Parametric models based on sensitivity analysis for passive components," *Intelligent Computer Techniques in Applied Electromagnetics, Studies in Computational Intelligence*. Berlin, Germany: Springer, vol. 119, 2008, pp. 231–239.
- [25] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. New York: Academic, 1980.
- [26] R. W. Freund, "Sprim: Structure-preserving reduced-order interconnect macro-modeling," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2004, pp. 80–87.
- [27] H. Yu, L. He, and S. X. D. Tan, "Block structure preserving model order reduction," in *Proc. IEEE Behavioral Modeling Simulation Workshop*, Sep. 2005, pp. 1–6.
- [28] T. Weiland, "A discretization method for the solution of Maxwell's equations for 6 component fields," *AEÜ, Electron. Commun.*, vol. 31, pp. 116–120, Mar. 1977.



Jorge Fernández Villena (S'06) was born in Avilés, Spain. He received the Engineer's degree in telecommunications from Cantabria University, Cantabria, Spain, in 2005. He is currently working toward the Ph.D. degree from the Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC ID), and Instituto Superior Técnico (IST), Technical University, Lisbon, Portugal.

His current research interests include integrated circuit interconnect modeling and simulation, with emphasis in parameterized model order reduction.



Luís Miguel Silveira (S'85–M'95–SM'00) was born in Lisbon, Portugal. He received the Engineer's (summa cum laude) and Master's degrees in electrical and computer engineering from the Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal, in 1986 and 1989, respectively, and the M.S., E.E., and Ph.D. degrees in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 1990, 1991, and 1994, respectively.

He is currently a Professor of Electrical and Computer Engineering with the Instituto Superior Técnico, Technical University of Lisbon, a Senior Researcher with Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa (INESC ID), Lisbon, Portugal, and a Founding Member of the Lisbon Center of the Cadence Research Laboratories, Berkeley, CA. His current research interests include various aspects of computer-aided design of integrated circuits, with emphasis on interconnect modeling and simulation, parallel computer algorithms, and the theoretical and practical issues concerning numerical simulation methods for circuit design problems.

Dr. Silveira has received notable paper citations at various conferences, including the Electronic Components and Technology Conference, the International Conference on Computer Aided Design, the Design Automation Conference, the Conference on Design, Automation and Test in Europe. He is a Member of Sigma Xi.