A Compact and Robust Block Structure Preserving Model Order Reduction Considering L-inverse Element

Hao Yu, Yiyu Shi and Lei He
Electrical Engineering Department
University of California, Los Angeles 90095
{hy255, yshi, lhe}@ee.ucla.edu

ABSTRACT

1. INTRODUCTION

As current technology advances from 90nm to 65nm with the increasing clock speed towards 10GHz, chips have more devices and I/O counts and flip-chip package becomes more widely used. To ensure the signal and power integrity for chip package co-design, a complete RLC model is required for accurate representations of interactions among package layers, C4 bumps, vias, on-chip power grids and all other signal traces. However, a detailed 3D extraction using the partial-element equivalent circuit (PEEC) model introduces densely coupled inductances that increase the model complexity. Sparsification and model order reduction (MOR) can be used to reduce the model complexity.

Because the partial inductance matrix (L) in the PEEC model is not diagonal dominant, simply truncating off-diagonal elements leads to negative eigenvalues and the truncated matrix loses passivity [1]. Instead of using L, the L-inverse matrix ($S = L^{-1}$) is found more strictly diagonal dominant, and therefore, enables the passive sparsification by directly truncating [2] or windowing [**?**]. In [3], the L-inverse element (S) is further derived from the vector potential equivalent circuit [4], where S is related to the drop of branch vector potential. Because the size of sparsified matrix could be still large, model order reduction is needed to further reduced the model complexity. Because AWE [**?**] explicitly matches the moment of the original system, it introduces the numerical instability and can not generate the high-order reduced model. Using the Arnoldi orthogonalization to find the Krylov subspace to span the moment space of the original system, and the congruencent transformation based projection, PRIMA [5] is widely used to efficiently generate an order reduced macro-model with the preserved passivity. However, this method only handles inductance that is stamped in the MNA (modified nodal analysis) matrix within the first-order form. As a result, it is computationally expensive when L is dense. Moreover, directly stamping S in the MNA matrix and applying PRIMA can not guarantee the passivity [].

Using nodal analysis (NA), ENOR [**?**] is proposed to passively reduce the RCS system in the second-order form, where the L-inverse element is represented by the nodal susceptance matrix $\Gamma = E_s^T L^{-1} E_s/s$ in the s-domain. As G, C, Γ matrices are symmetric positive definite (s.p.d.), congruencent transformation based projection preserves the passivity. Like AWE, ENOR needs explicit moment matching by recursive formula. Moreover, such a second-order form leads to the s-term in the right-hand side (RHS) of the system equation, and an auxiliary variable is introduced during the moment matching. Note that this auxiliary variable is not orthogonalized. As a result, above two aspects can both lead to the numerical error. To improve upon ENOR, SMOR [] eliminates the auxiliary variable and use an approximated Krylov subspace to span the moment space of the original system. Due to this approximation, the moment matching by SMOR is not accurate. Recently, SAPOR [**?**] is proposed to correctly obtain the orthogonalized Krylov subspace via a modified second-order Arnoldi method. However, this algorithm is prone to *break down* or can not generate compact model efficiently because of (i) the RHS s-term introduces nonzero initial correlation of input state; (ii) the low-rank Γ matrix has leads to strongly correlated spanning vectors.

In this paper, we present a compact and robust model order reduction considering L-inverse element. By introducing the *nodal vector potential* as the new state variable, the second-order formulation is obtained physically and no additional s-term is introduced in th right-hand-side (RHS). Moreover, the Singular-value-decomposition (SVD) and K-means-clustering based partitioning is applied to the moment matrix $X = colspan\{X_0, X_1, ..., X_n\}$ to group them into several uncorrelated blocks, and a block structure-preserving projection is applied accordingly. As a result, more poles can be closely approximated and less iteration is needed to achieve desired accuracy.

The rest of the paper is organized as follows. In Section II, we first review the background for L-inverse element and its macro-modeling approach. In Section III, by introducing the *nodal vector potential variable*, we derive our approach to reduce L-inverse element in the second-order form without RHS s-term. In Section IV, we discuss a further accuracy improvement by using block structure-preserving model reduction. We present the experimental results in Section V, and concludes the paper with discussion in Section VI.

2. PRELIMINARY

2.1 Stamp-in RCL^{-1} Element

Consider the branch resistive, capacitive, and inductive elements in the admittance form, with the external excitation current source $\mathbf{I}(t)$, the KCL and KVL equations are

$$\mathbf{E}\mathbf{i}_b^T = 0, \qquad \mathbf{E}^T\mathbf{v}_n = \mathbf{v}_b, \tag{1}$$

where $\mathbf{E} = [\mathbf{E}_g \quad \mathbf{E}_c \quad \mathbf{E}_l \quad \mathbf{E}_i]^T$ is incident matrix, $\mathbf{i}_b = [\mathbf{i}_g \quad \mathbf{i}_c \quad \mathbf{i}_l \quad \mathbf{i}_i]^T$, $\mathbf{v}_b = [\mathbf{v}_g \quad \mathbf{v}_c \quad \mathbf{v}_l \quad \mathbf{v}_i]^T$, are branch current and voltage, and $\mathbf{v}_n$ is nodal voltage. These branch

currents are determined by

$$\mathbf{i}_g = \mathbf{G}\mathbf{v}_g, \mathbf{i}_c = \mathbf{C}\frac{d\mathbf{v}_c}{dt}, \mathbf{v}_l = \mathbf{L}\frac{d\mathbf{i}_l}{dt}, \mathbf{i}_i = -\mathbf{I}(t) \quad (2)$$

where $\mathbf{G}$, $\mathbf{C}$, $\mathbf{L}$ are *branch-admittance form* of conductance, capacitance and inductance matrices.

Only reserving the nodal voltage $\mathbf{v}_n$ and branch-inductive-current $\mathbf{i}_l$ as the state variables, (1) and (2) become

$$G\mathbf{v}_n + C\frac{d\mathbf{v}_n}{dt} + \mathbf{i}_l = \mathbf{E}_i\mathbf{I}(t) \quad (3)$$

$$\mathbf{L}^{-1}(\mathbf{E}_l\mathbf{v}_n) = \frac{d\mathbf{i}_l}{dt}. \quad (4)$$

where $G = \mathbf{E}_g^T\mathbf{G}\mathbf{E}_g$, and $C = \mathbf{E}_c^T\mathbf{C}\mathbf{E}_c$ are *nodal-admittance form* of conductance and capacitance matrices. As such, (4) can be written in the *first-order form* in s-domain

$$\begin{aligned}(\mathcal{G} + s\mathcal{C})x(s) &= \mathcal{B}\mathbf{I}_0(s)\\ u(s) &= [\mathbf{E}_i^T \quad 0]x(s),\end{aligned} \quad (5)$$

where

$$x(s) = \begin{bmatrix}\mathbf{v}_n & \mathbf{i}_l\end{bmatrix}, \mathcal{B} = \begin{bmatrix}\mathbf{v}_n & \mathbf{i}_l\end{bmatrix}$$
$$\mathcal{G} = \begin{bmatrix} G & \mathbf{E}_l \\ -\mathbf{L}^{-1}\mathbf{E}_l^T & 0\end{bmatrix}, \mathcal{C} = \begin{bmatrix} C & 0 \\ 0 & I\end{bmatrix}. \quad (6)$$

Because the state matrix $\mathcal{G}$ does not satisfy: $\mathcal{G} + \mathcal{G}^T \succ 0$, directly using congruencent transformation based projection does not preserve passivity []. Moreover, since only nodal voltage $\mathbf{v}_n$ is of interests in most applications, the branch-inductive-current $\mathbf{i}_l$ can be viewed as intermediate variable that can be eliminated. As a result, (4) can be further written in the *second-order form* in s-domain

$$\begin{aligned}(sC + G + \Gamma/s)x(s) &= \mathbf{E}_i\mathbf{I}(s)\\ u(s) &= \mathbf{E}_i^T x(s),\end{aligned} \quad (7)$$

where

$$x(s) = \mathbf{v}_n \quad \Gamma = \mathbf{E}_l L^{-1}\mathbf{E}_l^T. \quad (8)$$

2.2 RCL^{-1} MOR in Second-order Form

Define $P = s_0C + G + \Gamma s_0$, where s_0 is the expansion point that keeps P non-singular. By substituting $s = s_0(1-z)$ and introducing an auxiliary quantity $y(z) = \frac{x(z)}{1-z}$, ENOR obtains following recursive relation to explicitly generate the moment X_k, Y_k, $\mathbf{I}_k$ of x, y, $\mathbf{I}_0$,

$$\begin{aligned}PX_k &= s_0CX_{k-1} - \frac{1}{s_0}\Gamma Y_{k-1} + \mathbf{E}_i\mathbf{I}_k\\ Y_k &= X_k + Y_{k-1}, \quad X_{-1} = Y_{-1} = 0.\end{aligned} \quad (9)$$

The generated of X_k are actually the scaled nodal voltage moments expanded at s_0. An orthonormalization using Gram-Schmidt method is further applied to obtain the subspace $colspan\{X_0, ..., X_q\}$. Note that the frequency scaling is prone to numerical instability, and the un-orthonormalized auxiliary quality $y(z)$ during explicitly moment calculation can grow too rapidly to lead numerical instability as well. As a result, ENOR is not accurate to generates high-order reduced model.

SMOR [6] improves upon ENOR by replacing $Y_k = \sum_{i-1}^{k} X_i$, for $k \geq -1$, and it results in

$$\begin{aligned}PX_k &= s_0CX_{k-1} - \tfrac{1}{s_0}\Gamma\textstyle\sum_{i-1}^{k} X_i \qquad k \geq 1\\ X_{-1} &= 0, \quad X_0 = P^{-1}\mathbf{E}\mathbf{I}_0.\end{aligned} \quad (10)$$

where the unit impulse current input is assumed to calculate moments, i.e., $\mathbf{I}_k = 0$. However, the summation in the recurrence relation still presents the numerical problem as the errors can be accumulated when matching high-order moments. Therefore, SMOR only uses the first three terms in (10) to generate moment X'_k that approximates X_k

$$PX'_k = s_0CX_{k-1} - \frac{1}{s_0}\Gamma X'_{k-1} - \frac{1}{s_0}\Gamma X'_{k-1} \qquad k \geq 1. \quad (11)$$

As a result, the subspace $colspan\{X'_0, ..., X'_q\}$ is only an approximation of the space spanned by the moments of the original system . Therefore, the reduced model by SMOR can not exactly match the original system as well.

In SAPOR [?], the RCL^{-1} system can be rewritten explicitly in

$$\begin{aligned}(s^2C + sG + \Gamma)x(s) &= s\mathbf{E}_i\mathbf{I}(s)\\ u(s) &= \mathbf{E}_i^T x(s).\end{aligned} \quad (12)$$

Expanding (12) at $s = s_0 + \sigma$, it becomes

$$(\sigma^2C + \sigma D + K)x(\sigma) = b_0 + b_1\sigma \quad (13)$$

where $D = 2s_0C + G$, $K = s_0^2C + s_0G + \Gamma$, $b_0 = s_0\mathbf{E}_i\mathbf{I}_0$, and $b_1 = \mathbf{E}_i\mathbf{I}_0$.

Introducing a new auxiliary variable $z(\sigma)$, the system can be linearized as

$$\begin{bmatrix}x \\ z\end{bmatrix} = (I - \sigma\mathcal{A})\begin{bmatrix}X_0 \\ Z_0\end{bmatrix} \quad (14)$$

where

$$\mathcal{A} = \begin{bmatrix}-K^{-1}D & K^{-1} \\ -C & 0\end{bmatrix}, \quad X_0 = K^{-1}b_0, \quad Z_0 = b_1. \quad (15)$$

Obviously, the k-th moment of x is

$$X_k = \begin{bmatrix}I & 0\end{bmatrix}A^{k-1}\begin{bmatrix}X_0 \\ Z_0\end{bmatrix}. \quad (16)$$

Using a modified *second-order Arnoldi* orthonormalization, a projection matrix $X = colspan\{X_0, X_1, ..., X_n\}$ is found to span the generalized Krylov space $\mathcal{K}(A, B, X_0)$, where $A = K^{-1}D$ and $B = K^{-1}$.

Note that (i) due to the s-term in right-hand-side, the Z_0 term is nonzero and linearly dependent on X_0; (ii) as dimension of nodal voltage variable $\mathbf{v}_n$ is usually much larger than branch-inductive-current variable $\mathbf{i}_l$, boosting $\mathbf{L}$ to Γ resulting in a low-rank matrix that has strongly correlated rows/columns. As a result, the generalized Krylov space $\mathcal{K}(\mathcal{A}, \mathcal{B}, X_0)$ has strong linear dependent column-span vectors $\{X_0, X_1, ..., X_n\}$. Therefore, this algorithm is prone to *break down* or can not generate compact model efficiently.

3. SECOND-ORDER FORMULATION WITH NODAL VECTOR-POTENTIAL VARIABLE

To effectively apply the second-order Arnoldi orthonormalization, we need find a second-oder formulation of RCL^{-1} system such that the auxiliary initial input p_0 at RHS is zero. This can be achieved by introducing the nodal vector potential variable.

3.1 Nodal Vector Potential Variable

As shown in [3], the vector potential equivalent is connected with L^{-1} by

$$\mathbf{L}^{-1}\mathbf{A}_l = \mathbf{I}_l \quad \frac{d\mathbf{A}_l}{dt} = \mathbf{E}_l\mathbf{v}_n, \quad (17)$$

where $\mathbf{A}_l$ is the branch vector potential associated with branch-inductance current $\mathbf{I}_l$, and all current are assumed flowing along z-axis.

Define the *nodal vector potential*

$$\mathbf{A}_l = \mathbf{E}_l \mathbf{A}_n \tag{18}$$

and rewrite (4)

$$\begin{aligned}(s^2C + sG + \Gamma)x(s) &= \mathbf{E}_i \mathbf{I}(s)\\ u(s) &= \mathbf{E}_i^T x(s),\end{aligned} \tag{19}$$

where $x(s) = \mathbf{v}_n$. Note that $\frac{d\mathbf{A}_n}{dt} = \mathbf{v}_n$. Let $x = [\mathbf{A}_n \quad \mathbf{v}_n]^T$. (**??**) can be linearized

$$\begin{aligned}\mathcal{G}x(s) + s\mathcal{C}x(s) &= \mathcal{B}\mathbf{I}_0(s)\\ u(s) &= \mathcal{L}x(s),\end{aligned} \tag{20}$$

where

$$\mathcal{G} = \begin{bmatrix}\Gamma & 0\\ 0 & C\end{bmatrix} \quad \mathcal{C} = \begin{bmatrix}G & C\\ -C & 0\end{bmatrix} \quad \mathcal{L} = \begin{bmatrix}0 & \mathbf{E}_i^T\end{bmatrix} \tag{21}$$

Note that the output port matrix is to select the nodal voltage $\mathbf{v}_n$.

Clearly, $\mathcal{G} + \mathcal{G}^T \succ 0$ and so does $\mathcal{C}$. The first-order based MOR like PRIMA, i.e., Arnoldi orthonormalization and congruencent transformation, therefore, leads to a passive model reduction of (20). As a result, the pole/residue can be easily obtained by eigen-decomposition and a MIMO macro-model is constructed accordingly for both time and frequency domain simulation. We call this approach as fVMOR.

Nevertheless, because only $\mathbf{v}_n$ is the interested state variable, and compared to direct Arnoldi method, the *second-order Arnoldi method* based orthonormalization is observed more cost efficient in both flop counts and memory requirements. More importantly, the RCL^{-1} in second-order form leads to structure-preserving model reduction as discussed in Section **??**. Therefore, we present in details of MOR for (20) in the second-order form using second-order Arnoldi orthonormalization. We call this approach as sVMOR.

3.2 Second-order Arnoldi Method

Assuming an impulse input $\mathbf{I}_0$, and expanding at $s = s_0 + \sigma$, (19) can be rewritten

$$\begin{aligned}x(\sigma) = (I - \sigma\mathcal{A})\mathbf{i}(\sigma), \mathcal{A} = \begin{bmatrix}A & B\\ I & 0\end{bmatrix}\\ A = K^{-1}D, B = K^{-1}C, \mathbf{i} = [\mathbf{E}_i\mathbf{I}_0 \quad 0]^T.\end{aligned} \tag{22}$$

It results in following moments sequence $\{\mathbf{m}_0, \mathbf{m}_1, ..., \mathbf{m}_{n-1}\}$, where

$$\begin{aligned}\mathbf{m}_0 = \mathbf{i}, \mathbf{m}_1 = A\mathbf{m}_0\\ \mathbf{m}_j = A\mathbf{m}_{j-1} + B\mathbf{m}_{j-2}, j \geq 2\end{aligned} \tag{23}$$

is called *second-order Krylov space*, i.e., $\mathcal{K}(A, B; \mathbf{i}) = colspan\{\mathbf{m}_0, \mathbf{m}_1, ..., \mathbf{m}_{n-1}\}$. Its orthonormal basis $\mathbf{q}_i$ can be constructed via a second-order Arnoldi procedure [**?**].

4. BLOCK STRUCTURE-PRESERVING MODEL REDUCTION

5. NUMERICAL EXPERIMENT

6. CONCLUSION

7. REFERENCES

[1] Z. He, M. Celik, and L. Pillegi, "SPIE: Sparse partial inductance extraction," in *Proc. Design Automation Conf. (DAC)*, pp. 137–140, 1997.

[2] A. Devgan, H. Ji, and W. Dai, "How to efficiently capture on-chip inductance effects: introducing a new circuit element K," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 150–155, 2000.

[3] H. Yu and L. He, "Vector potential equivalent circuit based on PEEC inversion," in *Proc. Design Automation Conf. (DAC)*, pp. 781–723, 2003.

[4] A. Pacelli, "A local circuit topology for inductive parasitics," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 208–214, 2002.

[5] A. Odabasioglu, M. Celik, and L. Pileggi, "PRIMA: Passive reduced-order interconnect macro-modeling algorithm," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 645–654, 1998.

[6] H. Zheng and L. T. Pileggi, "Robust and passive moel order reduction for circuits containing susceptance elements," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2002.