BSMOR Methed 2: Block Structure-preserving Model Order Reduction for Interconnected System

Hao Yu, and Lei He *

1. EXPLICIT PARTITION AND STRUCTURED BLOCK PROJECTION

To construct the structured-projection matrix $\mathcal{V}$, we in Section II first obtain V by applying PRIMA to the entire $\mathcal{G}$, $\mathcal{C}$ matrices, and then partition V using SVD and K-means algorithm. As all operations are on the entire matrix, its computation effort could be too big for large scale circuits (although fewer iterations are needed to achieve higher accurate reduced model compared to PRIMA). To reduce the computation cost, we propose to first explicitly partition the original circuit into m interconnected blocks with $(\mathcal{G}_{i,i}, \mathcal{C}_{i,i})$, then apply PRIMA to each block and obtain the projection matrix $V_i(span(V_i) = \mathcal{K}(\mathcal{A}_{i,i}, \mathcal{R}_{i,i}, q))$ for each block, and construct the structured-projection matrix

$$\mathbf{V} = diag[V_1, V_2, ..., V_m]. \quad (1)$$

Below, we prove that the structured-projection matrix $\mathbf{V}$ constructed in this fashion guarantees q moments matching as well.

1.1 General Interconnected Block Structure

For a interconnected block structure, we assume that the input source vector for ith-block is

$$u_{ci}(s) = u_{ei}(s) + \sum_{j\in 1,...,m}^{j\neq i} X_{ij} y_{cj}(s), \quad (2)$$

where $u_{ei}(s)$ is the external sources at ith block, $X_{ij}(s) = X_{g_{ij}} + sX_{cij}$ $(X_{ij} \in R^{n_i \times n_j})$ is the branch admittance matrix that connects ith and jth block, and $y_{cj}(s)$ is the output vector at jth block. Moreover, the transfer function of ith block is

$$\begin{aligned} y_{ci}(s) &= h_i(s) u_{ci}(s) \\ h_i(s) &= b_i^T (\mathcal{G}_{i,i} + s\mathcal{C}_{i,i})^{-1} b_i. \end{aligned} \quad (3)$$

where b_i is an extended port matrix $(\in R^{n_i \times n_i})$ to connect boundary ports at ith block. Note that

$$\mathcal{G}_{i,j} = b_i^T X_{g_{i,j}} b_j, \quad \mathcal{C}_{i,j} = b_i^T X_{ci,j} b_j. \quad (4)$$

Define the block connection matrix

$$\mathbf{X} = \begin{bmatrix} 0 & X_{1,2} & \dots & X_{1,m} \\ X_{2,1} & 0 & \dots & X_{2,m} \\ . & . & . & . \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1} & X_{m,2} & \dots & 0 \end{bmatrix}; \quad (5)$$

the block transfer function

$$\mathbf{H}_0(s) = diag\{h_1(s), ..., h_m(s)\} = \mathbf{B}_0^T (\mathbf{G}_0 + s\mathbf{C}_0)^{-1} \mathbf{B}_0, \quad (6)$$

where

$$\begin{aligned} \mathbf{B}_0 &= diag\{b_1, ..., b_m\}, \\ \mathbf{G}_0 &= diag\{\mathcal{G}_{1,1}, ..., \mathcal{G}_{m,m}\}, \\ \mathbf{C}_0 &= diag\{\mathcal{C}_{1,1}, ..., \mathcal{C}_{m,m}\}; \end{aligned} \quad (7)$$

the block connection input/output vector

$$u_c(s) = [u_{c1}(s), ..., u_{cm}(s)], \quad y_c(s) = [y_{c1}(s), ..., y_{cm}(s)]; \quad (8)$$

and the block external input/output vector

$$\mathbf{F} = [F_1, ..., F_m]. \quad (9)$$

The system equation, therefore, can be rewritten as

$$\begin{aligned} u_c(s) &= \mathbf{F} u_e(s) + \mathbf{X}(s) y_c(s), \\ y_c(s) &= \mathbf{H}_0(s) u_c(s), \\ y_e(s) &= \mathbf{F}^T y_c(s). \end{aligned} \quad (10)$$

Its transfer function is

$$\mathbf{H}(s) = \mathbf{F}^T (I - \mathbf{H}_0(s)\mathbf{X}(s))^{-1} \mathbf{H}_0(s) \mathbf{F}, \quad (11)$$

which can be realized by the following state space

$$\begin{aligned} \mathbf{G} &= \mathbf{G}_0 + \mathbf{B}_0^T \mathbf{X_g} \mathbf{B}_0, \\ \mathbf{C} &= \mathbf{C}_0 + \mathbf{B}_0^T \mathbf{X_c} \mathbf{B}_0, \\ \mathbf{B} &= \mathbf{B}_0^T \mathbf{F}. \end{aligned} \quad (12)$$

It is easy to check that: $\mathbf{G} = \mathcal{G}, \mathbf{C} = \mathcal{C}$, and $\mathbf{B} = \mathcal{B}$, i.e., the interconnected representation $(\mathbf{G}, \mathbf{C}, \mathbf{B})$ has the same state space with the original system $(\mathcal{G}, \mathcal{C}, \mathcal{B})$. Note that the rationale to use the presentation in the interconnected block structure is to obtain following relation for the Krylov subspace between the diagonal block structure and the original system.

*Hao Yu and Lei He are with Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: {hy255,lhe}@ee.ucla.edu).

THEOREM 1. *Define the state matrices for the block diagonal structure:* $\mathbf{A}_0 = (\mathbf{G}_0 + s_0\mathbf{C}_0)^{-1}\mathbf{C}_0$, $\mathbf{R}_0 = (\mathbf{G}_0 + s_0\mathbf{C}_0)^{-1}\mathbf{B}_0$, *and for the general interconnected structure:* $\mathbf{A} = (\mathbf{G} + s_0\mathbf{C})^{-1}\mathbf{C}$, $\mathbf{R} = (\mathbf{G} + s_0\mathbf{C})^{-1}\mathbf{B}$. *Then the corresponding Krylov subspaces satisfy:*

$$\mathcal{K}(\mathbf{A}, \mathbf{R}, q) \subseteq \mathcal{K}(\mathbf{A}_0, \mathbf{R}_0, q). \tag{13}$$

Proof: It is equivalent to prove that for any $v \in \mathcal{K}(\mathbf{A}, \mathbf{R}, q)$, it has $v \in \mathcal{K}(\mathbf{A}_0, \mathbf{R}_0, q)$ as well. Note that if $v \in \mathcal{K}(\mathbf{A}, \mathbf{R}, q)$, it means that there exists a q-dimensional column vector w_i:

$$\begin{aligned} vw_0 &= \mathbf{R}, \quad vw_1 = -\mathbf{A}\mathbf{R}, ..., \\ vw_i &= (-1)^i\mathbf{A}^i\mathbf{R}, \quad vw_{q-1} = (-1)^{q-1}\mathbf{A}^{q-1}\mathbf{R}. \end{aligned}$$

Sum the left-hand-side and right-hand-side respectively,

$$\begin{aligned} vw &= (I + s\mathbf{A})^{-1}\mathbf{R} = (\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B} \\ &= (\mathbf{G}_0 + s\mathbf{C}_0 + \mathbf{B}_0^T\mathbf{X}\mathbf{B}_0)^{-1}\mathbf{B}_0\mathbf{F} \end{aligned} \tag{14}$$

where $w = \sum_{i=0}^{q-1}(-1)^i w_i$.

Therefore,

$$vw' = (\mathbf{G}_0 + s\mathbf{C}_0)^{-1}\mathbf{B}_0^T, \tag{15}$$

where $w' = w(\mathbf{F} - \mathbf{X}\mathbf{B}_0 vw)^{-1}$. As a result,

$$\begin{aligned} vw_0' &= \mathbf{R}, \quad vw_1' = -\mathbf{A}_0\mathbf{R}_0, ..., \\ vw_i' &= (-1)^i\mathbf{A}_0^i\mathbf{R}_0, \quad vw_{q-1}' = (-1)^{q-1}\mathbf{A}_0^{q-1}\mathbf{R}_0 \end{aligned}$$

i.e., $v \in \mathcal{K}(\mathbf{A}_0, \mathbf{R}_0, q)$.

Based on Theorem 3, we have the following Corollary:

COROLLARY 1. *Let* V_i *satisfy* $\mathcal{K}(\mathbf{A}_{i,i}, \mathbf{R}_i, q) \subseteq span(V_i)$ *and* $\mathbf{V} = diag[V_1, V_2, ..., V_m]$. *The reduced system* $\widetilde{H}(s)$ *projected by* $\mathcal{V}$ *will match the first* q *moments in the expansion of* $\mathbf{H}(s)$ *about* s_0.

Proof: Because $\mathcal{K}(\mathbf{A}, \mathbf{R}, q) \subset \mathcal{K}(\mathbf{A}_0, \mathbf{R}_0, q) \subseteq span((\mathcal{V}))$, the q moments matching is easy to see by *Grimme's projection theorem.*

1.2 Incremental Explicit Partitioning

Similarly, if the original system has a block diagonal structure, i.e., $\mathbf{X} = 0$, the above structured-projection can match mq poles as well. In general, it needs to devise a proper initial partitioning such that the resulting model can closely approximate mq poles matching. Below, we present an *incrementally partitioning based on k-dominant-time-constant* algorithm to achieve the requirement. Note that the concept of *dominant time constant* is introduced in [1], where the system timing response is approximated by the first k dominant-time-constant, i.e., the first k most dominant eigen-values (poles) $\lambda_1, ..., \lambda_k$ of state matrix $\mathcal{A} = \mathcal{G}^{-1}\mathcal{C}$.

The outlined procedure is presented in Algorithm 2. The input is a group of m_0 natural basic blocks that are determined by the specific application such as the block current characterization for the substrate or P/G grid. Each basic block has state space $(\mathcal{G}_{i,i}, \mathcal{C}_{i,i}, b_i)$, and a projection matrix V_i is found by PRIMA such that $\mathcal{K}(\mathcal{A}_{i,i}, \mathcal{R}_{i,i}, q) = span(V_i)$. The clustering begins with the first basic block by merging its connected neighboring blocks. The state matrices of the resulted block are denoted by $(\mathcal{G}'_{i,i}, \mathcal{C}'_{i,i}, b'_i)$, and its corresponding projection matrix is V'_i. SVD is applied to V_i and V'_i to obtain their first k most dominant eigen-values (poles) $\lambda_1, ..., \lambda_k$ and $\lambda'_1, ..., \lambda'_k$. As a result, the merging operation will terminate if

$$||\frac{\sum_i^k \lambda_i/\lambda'_i}{k} - 1|| < \epsilon \tag{16}$$

where ϵ is a small value specified by the user. This termination criteria implies the condition that the k most dominant eigen-values (poles) will not change much when including more interconnected blocks together. The clustering further applies to the rest basic block till a converged partitioning V' is obtained. Consequently, a structured-projection matrix $\mathbf{V}$ is constructed to project the original system matrices $(\mathcal{G}, \mathcal{C}, \mathcal{B})$.

Algorithm 1 BSMOR Partitioning 2

1.Incremental partitioning
(1.1) *input*: $\mathcal{G}_{i,i}, \mathcal{C}_{i,i}, b_i$, $(i = 1, 2, ..., m_0)$;
I=1,i=0;
while $\overline{\lambda} < \epsilon$ **do**
(1.1) *form*: $\mathcal{A}_I^{i,i} = (\mathcal{G}_{I,I}^i + s_0\mathcal{C}_{I,I}^i)^{-1}\mathcal{C}_{I,I}^i$, $\mathcal{R}_I^{i,i} = (\mathcal{G}_{I,I}^i + s_0\mathcal{C}_{I,I}^i)^{-1}b_I^i$;
(1.2) *construct*: V_I^i such that $span(V_I^i) = \mathcal{K}(\mathcal{A}_I^{i,i}, \mathcal{R}_I^{i,i}, q)$;
(1.3) *calculate*: $[\lambda_1^i, ..., \lambda_k^i] = SVD(V^i, k)$;
(1.4) *cluster*: $(\mathcal{G}_{I,I}^{i+1}, \mathcal{C}_{I,I}^{i+1}, b_I^{i+1}) = merge[(\mathcal{G}_{I,I}^i, \mathcal{C}_{I,I}^i, b_I^i), (\mathcal{G}_{I+1,I+1}^i, \mathcal{C}_{I+1,I+1}^i, b_{I+1}^i)]$;
(1.5) *calculate*: $\overline{\lambda} = ||\frac{\sum_j^k \lambda_j^{i+1}/\lambda_j^i}{k} - 1||$;
end while
(1.6) *construct*: $\mathbf{V} = diag[V'_{1(n_1\times q)}, \ldots, V'_{m(n_m\times q)}]$;
3.Structured block projection
(2.1) *input*: $\mathbf{V}$;
(2.2) *project*: $\widetilde{\mathcal{G}} = \mathbf{V}^T\mathcal{G}\mathbf{V}$, $\widetilde{\mathcal{C}} = \mathbf{V}^T\mathcal{C}\mathbf{V}$, $\widetilde{\mathcal{B}} = \mathbf{V}^T\mathcal{B}$;
(2.3) *output*: $\widetilde{\mathcal{G}}, \widetilde{\mathcal{C}}, \widetilde{\mathcal{B}}$.

Note that if

$$\mathcal{G} + s\mathcal{C} = \mathbf{Y}'(s) + \mathbf{X}'(s)\mathbf{B}_0, \tag{17}$$

where

$$\begin{aligned} \mathbf{Y}'(s) &= diag\{(\mathcal{G}_{1,1} + s\mathcal{C}_{1,1}), ..., (\mathcal{G}_{m,m} + s\mathcal{C}_{m,m})\} \\ \mathbf{X}'(s) &= \mathbf{B}_0\mathbf{X}(s), \end{aligned} \tag{18}$$

we have

$$\begin{aligned} \widetilde{\mathcal{G}} + s\widetilde{\mathcal{C}} &= \mathcal{V}^T\mathbf{Y}'\mathcal{V} + \mathcal{V}^T\mathbf{X}'\mathcal{V}\mathcal{V}^T\mathbf{B}_0\mathcal{V} \\ &= \widetilde{\mathbf{Y}}'(s) + \widetilde{\mathbf{X}}'(s)\widetilde{\mathbf{B}}_0, \end{aligned} \tag{19}$$

where $\widetilde{\mathbf{Y}}'$ and $\widetilde{\mathbf{X}}'$ both $\in R^{mq\times mq}$.

2. TWO LEVEL PARTITIONED ANALYSIS OF REDUCED MODEL

Because the reduction preserves the structure, it preserves: (i) the sparsity of state matrices; In contrast, the projected matrices by PRIMA are fully dense. (ii) the reciprocity of the network which enables efficient realization of the network; In contrast, PRIMA does not preserve this property.

(*iii*) the block structure; It means that the reduced block can be distinguished by a subset of ports specified during BSMOR. Due to the preserved block structure , a *bordered-block diagonal* (BBD) partitioned solution can be applied to solve each block individually in both frequency and time domain.

Note that the runtime and memory requirement to solve a linear system are primarily determined by the size, sparsity, and structure of the matrix. Using partitioning, the large coupled network is divided into sub-blocks with manageable size and solved in blocks individually [2]. Moreover, partitioning can also be employed when network consists of repetitive identical subnetworks so that only one equation needs to be stored.

Consider the system equation for the reduced model

$$\widetilde{\mathbf{Y}}x = \widetilde{\mathbf{b}} \tag{20}$$

In frequency domain at a frequency point s

$$\begin{aligned} \widetilde{\mathbf{Y}} &= \widetilde{\mathcal{G}} + s\widetilde{\mathcal{C}} = \widetilde{\mathbf{Y}}'(s) + \widetilde{\mathbf{X}}'(s)\widetilde{\mathbf{B}}_0 \\ \widetilde{\mathbf{b}} &= \widetilde{\mathcal{B}}u_e(s). \end{aligned} \tag{21}$$

Note that in time domain at a time instant t with time step h

$$\begin{aligned} \widetilde{\mathbf{Y}} &= \widetilde{\mathcal{G}} + \frac{1}{h}\widetilde{\mathcal{C}} = \widetilde{\mathbf{Y}}'(h) + \widetilde{\mathbf{X}}'(h)\widetilde{\mathbf{B}}_0 \\ \widetilde{\mathbf{b}} &= \frac{1}{h}\widetilde{\mathcal{C}}x(t-h) + \widetilde{\mathcal{B}}u_e(t). \end{aligned} \tag{22}$$

The state vector can be solved by Sherman-Morrison-Woodbury formula [3]

$$\begin{aligned} x &= P^{(0)} - Pq \\ P^{(0)} &= (\widetilde{\mathbf{Y}}')^{-1}\widetilde{\mathbf{b}}, \quad P = (\widetilde{\mathbf{Y}}')^{-1}\widetilde{\mathbf{X}}' \\ q &= (I + \widetilde{\mathbf{B}}_0 P)^{-1}\widetilde{\mathbf{B}}_0 P^{(0)}. \end{aligned} \tag{23}$$

To avoid explicit inversion, LU-factorization needs to applied $\widetilde{\mathbf{Y}}'$ and $I + \widetilde{\mathbf{B}}_0(\widetilde{\mathbf{Y}}')^{-1}\widetilde{\mathbf{X}}'$. A two level analysis is proposed in [4] and outlined in Algorithm 3. Note that as $\widetilde{\mathbf{Y}}$ shows the block diagonal form, each reduced block matrix is fist solved individually with LU/Cholesky factorization and substitution (1.1-1.4) at the bottom level, the results from each reduced block are then used further to solve the coupling block (2.1-2.4) at the top level, and the final x_k of each reduced block is updated (3.1).

Typically, LU factorization requires $n^3/3$ multiplications and back/forward substitution requires $n^2/2$ multiplications. The computational cost of Algorithm 3 is therefore, ... Moreover, this algorithm can be extended in a hierarchical fashion if the multi-level partitioning is used initially.

Algorithm 2 Two Level Partitioned Analysis

1.Solve bottom level individually
for every k *in* m **do**
(1.1) *input*: $\widetilde{\mathbf{b}}_k, \widetilde{\mathbf{X}}_k^{\prime(i)}, \widetilde{\mathbf{Y}}_k'$;
(1.2) *factor*: LU/Cholesky factor $\widetilde{\mathbf{Y}}_k'$;
$\widetilde{\mathbf{X}}_k^{\prime(0)} = \widetilde{\mathbf{b}}_k$;
for every i *in* q **do**
(1.3) *solve*: back-substitution $\widetilde{\mathbf{Y}}_k' P_k^{(i)} = \widetilde{\mathbf{X}}_k^{\prime(i)}$;
end for
end for
2.Solve top level
(2.1) *input*: $\widetilde{\mathbf{B}}_0, P, P^{(0)}$;
(2.2) *factor*: LU/Cholesky factor $I + \widetilde{\mathbf{B}}_0 P$;
(2.3) *solve*: back-substitution $(I + \widetilde{\mathbf{B}}_0 P)q = \widetilde{\mathbf{B}}_0 P^{(0)}$;
3.Update bottom level individually
(3.1) *output*: $x = P^{(0)} - Pq$.

3. REFERENCES

[1] L. Vandenberghe, S. Boyd, and A. E. Gamal, "Optimizing dominant time constant in rc circuits," *IEEE Trans. on CAD*, pp. 110–125, 1998.

[2] R.A.Rohrer, "Circuit partitioning simplified," *IEEE Trans. on CAS*, pp. 2–5, 1988.

[3] J. Sherman and W.J.Morrison, "Adjustment of the an inverse matrix corresponding to changes in the elements of original matrix," in *Annu. Math. Statist.*, p. 621, 1949.

[4] G.Guardabassi and A. Sangiovanni-Vincentelli, "A two level algorithm for tearing," *IEEE Trans. on CAS*, pp. 783–791, 1976.