Hierarchical Krylov Subspace Reduced Order Modeling of Large RLC Circuits

Duo Li Dept. of Electrical Engineering University of California Riverside, CA 92521

dli@ee.ucr.edu

ABSTRACT

In this paper, we propose a new model order reduction approach for large interconnect circuits using hierarchical decomposition and Krylov subspace projection-based model order reduction. The new approach, called hiePrimor, first partitions a large interconnect circuit into a number of smaller subcircuits and then performs the projection-based model order reduction on each of subcircuits in isolation and on the top level circuit thereafter. The new approach can exploit the parallel computing to speed up the reduction process. Theoretically we show hiePrimor can have the same accuracy as the flat reduction method given the same reduction order and it can also preserves the passivity of the reduced models as well. We also show that partitioning is important for hierarchical projection-based reduction and the minimum-span objective should be required to archive best performance for hierarchical reduction. The proposed method is suitable for reducing large global interconnects like coupled bus, transmission lines, large clock nets in the post layout stage. Experimental results demonstrate that hiePrimor can be significantly faster than flat projection method like PRIMA and be order of magnitude faster than PRIMA with parallel computing without loss of accuracy.

1. INTRODUCTION

Compact modeling of passive RLC interconnect networks has been a research-intensive area in the past decade owing to increasing signal integrity effects and increasing design complexity in today's nanometer VLSI designs. Reducing the parasitic RLC circuits by approximate compact models can significantly improve the simulation and verification process in nanometer VLSI designs. As the technology moves to 45nm, the massive extracted post-layout circuits will make the reduction imperative before any meaning simulations and verifications. Hence the reduction algorithm must scale to attack very large circuit sizes in the future.

One of the most successful reduction algorithms is based on subspace projection [12, 3, 17, 6, 10, 19]. Those methods typically project the original circuit stats into the dimensioned-reduced Krylov subspace to reduce the model order. Projection-based methods were pioneered by Asymptotic Waveform Evaluation (AWE) algorithm [12] where explicit moment matching was used to compute dominant poles at low frequency. Pade via Lanczos (PVL) [3], Arnoldi Transformation method [17] improved the numerical stability of AWE, congruence transformation method [6] and PRIMA [10] can further produce passive models. At the same time, many other approaches also have been proposed, such as balanced truncation based reduction methods [11, 21, 22], local node reduction methods [15, 16] and general node reduction reduction method [13, 18]. But Krylov subspace based-reduction method remains a viable approach for Sheldon X.-D. Tan Dept. of Electrical Engineering University of California Riverside, CA 92521 stan@ee.ucr.edu

many practical interconnect reduction problems owning to its high efficiency. Existing projection based reduction methods, however, lack the general way to exploit the parallel computing capabilities, which become more popular with emerging multi-core computing architectures. We notice that Grimme has explored the parallel computation for multi-point Krylov based reduction where each Krylov subspace from each expansion point is computed in parallel [5].

But in this paper, we investigate the parallelism within the reduction operations in one expansion point for one large interconnect circuit. Hierarchical reduction of interconnects have also been studied from different perspectives. In HiPRIME algorithm [8], hierarchical reduciton has been extended in the extended Krylov subspace method (EKS) to fast compute the responses of on-chip power grid networks. The HiPRIME method reduces both system and input signals at the same time in a hierarchical way, but it does not produce a reduced model for general use. In the RecMOR method [4], Feldmann and Liu applied the combined terminal and model order reduction on the subcircuits based on the observation that partitioning may lead to many circuits with many new terminals, which will affect the efficiency for projection-based reduction methods. However, terminal reduction in general still remains a difficult problem and may not be effective for many practical problems [11, 9].

In this paper, we propose a new hierarchical Krylov subspace based reduction method. The new method combines the min-cut based partitioning and Krylov subspace method to speed up the reduction process. The proposed method is more suitable for reducing many large global interconnects like coupled bus, transmission lines and large clock nets where the number of ports are general not significant. The new method, called *hiePrimor*, first partitions a large RLC circuit into two or more levels and then perform the projection based reduction on subcircuits in a bottom-up way. Our contributions are as follows: (1) theoretically we show that if kth order block moment order is preserved in all the reduction processes for all subcircuits and top level circuit, then first k block moments will be preserved in the final reduced models; (2) we prove that the new hierarchical reduction method also preserve the passivity of the reduced models for RLC circuits at all the hierarchical level; (3) we show that the proposed method not only can exploit parallel computing to speed up the reduction process, but also can significantly improve the analysis capacity by partitioning strategy; (4) we study the impacts of partitioning on the reduction efficiency and show that partitioning is critical for the hierarchical reduction process and min-span or min-cut objective should be performed for partitioning. We apply the existing hMETIS partitioning tools [1] to perform the min-cut partitioning. Experimental results show that the proposed method can lead to significant speedup over flat projection based method like PRIMA and order of magnitudes speedup over PRIMA if parallel computing is used. Interconnect circuits with millions of nodes can be analyzed in a desktop PC using Matlab in a few minutes.

The rest of the paper is organized as follows. We review the

^{*}This work is supported in part by NSF CAREER Award CCF-0448534, UC Micro Program #06-252 and #07-105 via Cadence Design System Inc.

Krylov subspace based model order reduction in Section 2. Then we present the main idea of new method using an illustrative example in Section 3. We show moment matching property for the new method in Section 4. In Section 5 we prove that the new method preserves the passivity of the reduced models. We discus the partitioning impacts and scheme in Section 6. We present the experimental results in Section 7, and conclude the paper in Section 8.

2. REVIEW OF SUBSPACE PROJECTION BASED MOR METHODS

In this section, we review the Krylov subspace projection-based methods, which are also used for our hierarchical projection MOR method.

Without loss of generality, a linear m-port RLC circuit can be expressed as

$$C\dot{\mathbf{x}}_n = -G\mathbf{x}_n + B\mathbf{u}_m$$
(1)
$$\mathbf{i}_m = L^T \mathbf{x}_n$$

where x_n is the vector of state variables and n is the number of state variables, m is the number of independence sources specified as ports. C, G are storage element matrix and conductance matrices respectively. B and L are position matrices for input the output ports.

Define $A = -G^{-1}C$, $A \in \mathbb{R}^{n \times n}$ and $R = G^{-1}B$, $R = [r_0, r_1, ..., r_m] \in \mathbb{R}^{n \times m}$. The transfer function matrix after Laplace transformation is $H(s) = L^T (G + sC)^{-1}B = L^T (I_n - sA)^{-1}R$ where I_n is the $n \times n$ identity matrix. The block moments of H(s) are defined as the coefficients of Taylor expansion of H(s) around s = 0:

$$H(s) = M_0 + M_1 s + M_2 s^2 + \dots$$
(2)

where $M_i \in \Re^{m \times m}$ and can be computed as $M_i = L^T A^i R$. In the sequel, we also use m_i denotes the terminal count for subcircuit *i*.

The idea of model order reduction is to find a compact system of a much smaller size than the original system. Krylov subspace based method accomplishes this by projecting the original system on a special subspace which spans the same space as the block moments of the original system. Specifically, the block Krylov subspace is defined as

$$Kr(A, R, q) = colsp[R, AR, A^{2}R, ..., A^{k-1}R, A^{k}r_{0}, A^{k}r_{1}, ..., A^{k}r_{l}]$$
(3)
$$k = \lfloor q/m \rfloor, \quad l = q - km.$$
(4)

For simplicity of expression, we assume $q = m \times k$ in the following and k is the order of block moments used in the Krylov subspace. i.e. k order block moments will be matched if Krylov subspace Kr(A, R, mk) is used. Then, projection MOR method tries to find orthogonal matrix $X \in \Re^{n \times q}$ such that colsp(X) = Kr(A, R, q). With

$$\tilde{C} = X^T C X \qquad \tilde{G} = X^T G X$$
$$\tilde{B} = X^T B \qquad \tilde{L} = X^T L$$

the reduced system of size q is found as

$$\tilde{C}\dot{\tilde{\mathbf{x}}}_n = -\tilde{G}\tilde{\mathbf{x}}_n + \tilde{B}\mathbf{u}_m$$

$$\mathbf{i}_m = \tilde{L}^T\tilde{\mathbf{x}}_n$$
(5)

The reduced transfer function become $\tilde{Y}(s) = \tilde{L}^T (\tilde{G} + s\tilde{C})^{-1}\tilde{B}$. An important result for projection-based MOR methods is that the reduced system approximates the original systems in terms of moment matching: if $Kr(A, R, q) \subseteq span(X)$, then the reduced transfer function $\tilde{Y}(s)$ and original transfer function H(s) matches the first k block moments where k = q/m. Also when L = B, the reduction process preserves passivity.

3. HIERARCHICAL PROJECTION MOR METHOD: HIEPRIMOR

We introduce our method by using an illustrative RC example circuit shown in Fig. 1. This circuit has been partitioned into three parts, the two subcircuits *I* and *II* and the top part, which connects the two subcircuits. The two subcircuits are connected via the top level circuit only.



Fig. 1. A partitioned RC circuit.

As a result, we have the partitioned MNA equations as shown in (6), where we partition the matrix into three parts and the input sources into two parts as input sources only appear in partition I and top level partition.

In general, we can write a *n*-way partitioned RLC circuit into the following general form:

$$\begin{bmatrix} G_{1} & 0 & \dots & G_{1t}^{T} \\ 0 & G_{2} & \dots & G_{2t}^{T} \\ \vdots & \vdots & \dots & \vdots \\ G_{1t} & G_{2t} & \dots & G_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{1} \\ \mathbf{x}_{2} \\ \vdots \\ \mathbf{x}_{t} \end{bmatrix}^{+} \begin{bmatrix} C_{1} & 0 & \dots & 0 \\ 0 & C_{2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & C_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{\dot{x}}_{1} \\ \mathbf{\dot{x}}_{2} \\ \vdots \\ \mathbf{\dot{x}}_{t} \end{bmatrix}$$
(7)
$$= \begin{bmatrix} B_{1} & 0 & \dots & 0 \\ 0 & B_{2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & B_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{1} \\ \mathbf{u}_{2} \\ \vdots \\ \mathbf{u}_{t} \end{bmatrix}$$

where the \mathbf{x}_i is the internal variable vector for partition *i* and \mathbf{u}_i is the external input vector for partition *i*. \mathbf{x}_t and \mathbf{u}_t are the variables and external input vectors of top level circuit. If there are no external inputs for partition *i*, then the corresponding columns in the position matrix can be removed as shown in (6). To simplify the notation, we also rewrite (7) as

$$G\mathbf{x} + C\dot{\mathbf{x}} = B\mathbf{u} \tag{8}$$

The idea of hierarchical projection-based reduction is to perform the reduction using projection MOR method for each subcircuit first assuming that the subcircuits are disconnected from the rest of the circuit. After the subcircuits are reduced, we perform the reduction on the their parent circuits of the subcircuits until we reach to the top level circuit. The benefit of doing this is that we can reduce the computation complexity by performing the reduction on the subcircuits and intermediate circuits and parallelism can be exploited to further speed up the reduction process as each subcircuit in one hierarchical level can be reduced independently.

To illustrate this idea, we still use the example in Fig. 1. To reduce the subcircuit *I*, we have the following subcircuit matrix:

$$\begin{bmatrix} G_1 & -G_1 & -1\\ -G_1 & G_1 + G_2 & 0\\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1\\ v_2\\ i_{u_1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0\\ 0 & C_1 & 0\\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v}_1\\ \dot{v}_2\\ \dot{i}_{u_1} \end{bmatrix} = \begin{bmatrix} 0 & 0\\ 0 & 1\\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1\\ \dot{i}_2 \end{bmatrix}$$
(9)

where i_2 is the current source attached to node 2, which becomes a terminal node now. The added current source is just for reduction propose. Note that the position matrix $B_1 = [0 \ 0 \ 1]^T$ for this subcircuit has been changed to

$$B_1' = \begin{bmatrix} 0 & 0\\ 0 & 1\\ 1 & 0 \end{bmatrix}$$
(10)



Fig. 2. Illustration of partitioned MNA matrix for a RC circuit.

This modification reflects the fact that the subcircuit I now has two terminal nodes: node 1 and node 2. Notice that all the internal nodes, which are inside a subcircuit and are connected to boundary node at the upper level via a device branch, become the terminal nodes of the subcircuits for the reduction propose (as the case of node 2). If a subcircuit does not have any external input (such as the subcircuit II), all the nodes incident on the boundary nodes will become the terminal nodes for the reduction of the subcircuit. As projection based-MOR method becomes less effective for increasing terminal counts, we should try to minimize the terminal counts of subcircuit. Therefore, the hierarchical reduction requires the min-span like ¹ partitioning of the circuit to as *span* is the number of terminals of all the subcircuits. In this way, we can achieve better reduction performance. We will further the discussion of the partitioning issue in the section 6.

After the projection matrix V_1 is computed using (9), where V_1 is the *k*th order block Krylov subspace. i.e. $V_1 \subseteq Kr(G_1^{-1}B_1, G_1^{-1}C_1, km_1)$, where m_1 is the terminal count of subcircuit 1, we can perform the reduction. But now we need to look at the subcircuit in the context of the whole circuit. From (7), for the subcircuit 1, we have

$$G_1 \mathbf{x}_1 + C_1 \dot{\mathbf{x}}_1 + G_{1t}^T \mathbf{x}_t = B_1 \mathbf{u}_1 \tag{11}$$

After the reduction, we have

$$\tilde{G}_1 \mathbf{z}_1 + \tilde{C}_1 \dot{\mathbf{z}}_1 + \tilde{G}_{1t}^T \mathbf{x}_t = \tilde{B}_1 \mathbf{u}_1$$
(12)

where $\mathbf{x} = V_1 \mathbf{z}$, $\tilde{G}_1 = V_1^T G_1 V_1$, $\tilde{C}_1 = V_1^T C_1 V_1$, $\tilde{B}_1 = V_1^T B_1$, $\tilde{G}_{1t} = V_1 G_{1t}$. Note that we use original B_1 here instead of B_1' . Since the $colsp(B_1) \subseteq colsp(B_1')$, we can use subspace defined in V_1 to perform the reduction.

We repeat the the reduction process on all the subcircuits as mentioned above. After this, we end up with the following order reduced system at the top level:

$$\begin{bmatrix} G_{1} & 0 & \dots & G_{1t}^{T} \\ 0 & \tilde{G}_{2} & \dots & \tilde{G}_{2t}^{T} \\ \vdots & \vdots & \dots & \vdots \\ \tilde{G}_{1t} & \tilde{G}_{2t} & \dots & G_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{1} \\ \mathbf{z}_{2} \\ \vdots \\ \mathbf{x}_{t} \end{bmatrix} + \begin{bmatrix} \tilde{C}_{1} & 0 & \dots & 0 \\ 0 & \tilde{C}_{2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & O_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{1} \\ \mathbf{z}_{2} \\ \vdots \\ \mathbf{z}_{t} \end{bmatrix}$$
$$= \begin{bmatrix} \tilde{B}_{1} & 0 & \dots & 0 \\ 0 & \tilde{B}_{2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & B_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{1} \\ \mathbf{u}_{2} \\ \vdots \\ \mathbf{u}_{t} \end{bmatrix}$$
(13)

In this paper, we only present the results for two level reduction as shown in (7). But the proposed method can be trivially extended to more hierarchical levels. We can also rewrite (13) as

$$G_r \tilde{\mathbf{x}}_r + C_r \tilde{\mathbf{x}}_r = B_r \mathbf{u}_r \tag{14}$$

With (13), we can continue the reduction by performing the reduction at the top level circuit using the projection-based reduction method again. After this, we have final reduced model:

$$\tilde{G}\tilde{\mathbf{x}} + \tilde{C}\dot{\tilde{\mathbf{x}}} = \tilde{B}\mathbf{u} \tag{15}$$

where $\tilde{G} = V_t^T G_r V_t$, $\tilde{C} = V_t^T C_r V_t$, $\tilde{B} = V_t^T B_r$. G_r and C_r and B_r are the circuit matrices in (14) and $V_t \subseteq Kr(G_r^{-1}B_r, G_R^{-1}C_r, q_t)$, where $q_t = km_t$ and m_t is the terminal count in top level circuit.

4. MOMENT MATCHING CONNECTION

In this section, we analyze the moment matching property of the proposed method. We show that if the kth order block moment is preserved/matched in the reductions for all the subcircuits and for the top level circuit, the final reduced model preserves the first k block moments of the original system.

Let's assume that we have an interconnected circuit system with the transfer function H(s), it consists of *n* subcircuits that connects together. Assume that we denote subcircuit *i* as (G_i, C_i, B_i) and we perform the projection based model order reduction on the subcircuit *i* only

$$(\tilde{G}_i, \tilde{C}_i, \tilde{B}_i) = (V_i^T G_i V_i, V_i^T C_i V_i, V_i^T B_i)$$
(16)

and keep all the other system unchanged. We generate the projection matrix V_i such that

$$V_i \subseteq Kr(A_i, R_i, q_i) \tag{17}$$

where $A_i = -G_i^{-1}C_i$, $R_i = G_i^{-1}B_i$ and $q_i = km_i$. Then we have the following result:

'LEMMA 1. The resulting interconnected circuit system transfer $H_1(s)$, which consists of the order reduced subcircuit $(\tilde{G}_i, \tilde{C}_i, \tilde{B}_i)$ with rest of subcircuits unchanged, matches the first k block moments of H(s).

The detailed proof of this lemma can be found at [20]. With Lemma 1, we can easily obtain the following result. For the interconnected circuit system H(s), all of its subcircuits are reduced by the projection based MOR method such that

$$(\tilde{G}_{i}, \tilde{C}_{i}, \tilde{B}_{i}) = (V_{i}^{T} G_{i} V_{i}, V_{i}^{T} C_{i} V_{i}, V_{i}^{T} B_{i}), i = 1, ..., n$$
(18)

such that $V_i \subseteq Kr(A_i, R_i, q_i)$, $q_i = km_i$ for all the subcircuits.

COROLLARY 1. The resulting interconnected circuit system transfer $\tilde{H}_2(s)$, which consists of the order reduced subcircuit $(\tilde{G}_i, \tilde{C}_i, \tilde{B}_i), i = 1, ..., n$ for all subcircuits, matches the first q block moments of H(s).

The proof of Corollary 1 can be obtained when we apply Lemma 1 n times to the interconnected circuit system H(s) such that we reduce one subcircuit at a time.

Now we are ready to present the main result regarding the proposed hierarchical model order reduction method, hiePrimor.

THEOREM 1. Given a partitioned RLC circuit defined in (7) with transfer function H(s), if we perform the projection based reduction on all the subcircuit and then top level circuit such that kth order block moment is preserved in the all reduction processes, the transfer function $\tilde{H}(s)$ of the reduced system in (15) will match the first k block moments of H(s).

¹*span* of cut net means that the number of internal nodes that a cut net connects from all the partitions.

The proof of the theory is obvious in light of Corollary 1 and the fact the top level reduction on (13) also preserves the kth order block moment. Theorem 1 also indicates that for hierarchical reduction process, we should always use the same order for all the reduction processes. For the same reduction order k, different subcircuit may have different reduced model sizes as the size of the reduced model is km_i , where m_i is the terminal count of the subcircuit *i*.

In summary, the proposed hierarchical projection based reduction method, hiePrimor, will have the same accuracy as the flat projection based method if both methods use the same reduction order (order of block moments in the Krylov subspace).

ANALYSIS OF PASSIVITY PRESERVA-5. TION

In this section, we show that the proposed hiePrimor method preserves the passivity of the reduced models.

It is clear that the passivity is preserved for the reduction of the leaf level subcircuits as it is normal projection based model order reduction [14]. The only thing left is to show that reduction on the intermediate or top level circuits indicated by (13) still preserves the passivity.

THEOREM 2. The hierarchical projection based reduction algorithm preserves the passivity of the order reduced models at intermediate and top level circuits.

Proof: For a passive circuit, its transfer function must be positive real. A network circuit with matrix transfer function H(s) is said to be positive real iff

(1)
$$H(s)$$
 is analytic, for $Re(s) > 0$

(2)
$$H^*(s) = H(s^*)$$
, for $Re(s) > 0$

(3)
$$H(s) + H(s^*)^T \ge 0$$
, for $Re(s) > 0$

where \geq means nonnegative definite (or positive semidefinite). Condition (1) and (2) are typically always satisfied for a RLC circuit as it does not have unstable poles and the system has real response. And condition (3) needs to be proved.

Let denote the transfer function of the final reduced model in (15) as $H(s) = \tilde{B}^T (\tilde{G} + s\tilde{C})^{-1}\tilde{B}$. Condition 3 is equivalent to the requirement that $\tilde{G} + s\tilde{C}$ is positive real:

$$W(s) = \tilde{G} + s\tilde{C} \ge 0 \tag{19}$$

Let's $s = \sigma + i\omega$ and $\sigma > 0$, then $W(s) + W(s^*)^T$ becomes

$$K_h(s) = \tilde{G} + \tilde{G}^T + 2\sigma\tilde{C}$$
⁽²⁰⁾

We now need to prove that $K_h(s)$ is positive real. As we know that $\tilde{G} = V_t^T G_r V_t$ and $\tilde{C} = V_t^T C_r V_t$. If we define $V_r = diag(V_1, V_2, ..., V_n, I)$, where V_i are the projection matrix for subcircuit *i*. We can rewrite (14) as

$$V_r^T G V_r \mathbf{x}_r + V_r^T C V_r \dot{\mathbf{x}}_r = V_r^T B \mathbf{u}_r$$
(21)

$$V_t^T V_r^T G V_r V_t \tilde{\mathbf{x}} + V_t^T V_r^T C V_r V_t \dot{\tilde{\mathbf{x}}} = V_t^T V_r^T B \mathbf{u}$$
(22)

As a result, $K_h(s)$ become

$$K_h(s) = V_t^T V_r^T (G + G^T + 2\sigma C) V_r V_t$$
⁽²³⁾

where G and C are the circuit matrices written in the partitioned form as defined in (7). Therefore we need to prove that $G + G^T + G^T$ $2\sigma C$ is positive real as $K_h(s)$ is obtained by congruence transformation $W^{T}(G+G^{T}+2\sigma C)W$, where $W=V_{t}V_{r}$.

The partitioned G matrix is not in the so-called passive form using MNA formation from a RLC circuit as shown below: [10]

$$G_p = \begin{bmatrix} N & E^T \\ -E & 0 \end{bmatrix}$$
(24)

where $G_p \ge$ is nonnegative definite as $N \ge 0$ is nonnegative definite. Notice that each subcircuit conductance matrix G_i is in the passive form as we need to perform passive reduction using projection methods for each subcircuit. In this case, one can easily prove that G_p can be obtained by permuting same set of rows and columns simultaneously. Such permutation can be represented by the permutation matrix *P*:

$$G_p = P^T G P; (25)$$

This is true for C matrix. Finally, we have

$$K_h(s) = V_t^T V_r^T P^T (G_p + G_p^T + 2\sigma C_p) P V_r V_t$$
(26)

Since $G_p + G$ and C_p are negative definite for RLC circuits, $K_h(s)$ is positive real. QED

6. **CIRCUIT PARTITIONING**

Partitioning plays an important rule for the performance of the proposed reduction method. The reason is that the nodes that are inside a subcircuit and are incident on the boundary nodes at the top level will become the terminal nodes for subcircuits. The sizes of the reduced models grow linearly with the terminal number of the original circuits in the projection based reduction framework as the size of the reduced model is km_i , where k is the block moment order and m_i is terminal count for subcircuit *i*. To have smaller sizes of the reduced matrices (thus smaller nonzero elements in the matrices) which will be stamped into the higher level circuit matrix for further reduction, we need reduce terminal count of subcircuits as much as possible. This calls for the minimum-span like partitioning to achieve this.

Also the size of subcircuits cannot be too small compared with the number of terminals to have meaningful reduction on subcircuits. As a result, the proposed method is more suitable for very large RLC networks like bus, coupled transmission lines and clock nets with loosely coupled subcircuits.

After partitioning, the subcircuit terminals generated by partitioning will be driven by current sources in general, which requires the subcircuit has DC path for all of its nodes. If this is not the case, we have to introduce voltage sources at the terminal for the reduction purpose (to make the subcircuit G_i non-singular). This will add more interface terminals to the original circuits. As a result, we should minimize the capacitive cut. But the proposed method does not have any restrictions on types of boundary nodes.

To meet the partitioning requirement, we apply hMETIS partition tool suite [1], which employs the hierarchical partitioning strategy and is the best min-cut partitioning tools available. hMETIS tool suite is very efficient for partitioning very large networks. With hMETIS, the hiePrimor is able to reduce very large interconnect cirucits with millions of nodes in a PC using Matlab.

For very densely coupled circuits, the proposed method can still be applied. First, for the capacitively coupled circuits, it is well known that the coupling is more localized, which means the coupling can easily reduced without loss of much accuacry. For inductive coupling, which has long range effects due to partial inductance formation, many methods exist to reduce the coupling using various windon-based truncation techniques [7, 2, 23]. After this we apply partition the less coupled circuits for hierarchical reduction.

EXPERIMENTAL RESULTS 7.

In this section, we report the experiment results of hiePrimor on some interconnect circuits. We compare it with PRIMA [10] with and without parallel computing settings. We implement the hiePrimor method and original PRIMA using Matlab 7.0 and Python on a Intel Xeon 3.0GHz dual CPU workstation with 2GB memory under Linux environment.

Test Ckts	#Nodes	#Sub	#Ports	PRIMA (s)	hiePrimor (s)	Speedup
Ckt1	25k	2	8	5	4	1.25
Ckt2	50k	4	16	16	9	1.78
Ckt3	100k	8	16	32	13	2.46
Ckt4	200k	8	16	69	27	2.56
Ckt5	500k	16	24	248	60	4.13
Ckt6	800k	16	24	401	99	4.05
Ckt7	1 <i>M</i>	16	32	863	154	5.60
Ckt8	1.5M	16	20	_	176	_

TABLE I REDUCTION TIME COMPARISON OF PRIMA AND HIEPRIMOR ($k = 4, q = n \times k$)

We use sparse matrix structures in Matlab. Python is used for a parser converting Spice format netlist into Matlab format. Our Matlab implemented program ran fast because Python already did most of the I/O operations and Matlab could concentrate on only matrix (vector) operations.

Our test circuits are created based on a bus circuit structure, where each circuit has capacitively-coupled bus lines with different length and each of them are modeled as RC ladder-like circuits. To partition the testing circuit in SPICE format, we transform the netlist into the one that hMETIS can read and then partition the circuits into several small spice format circuits of equal size with the min-cut objective.

We first show that hiePrimor and PRIMA give almost the same accuracy for the given block moment order k (our claim in Section 4). We set the reduction order q as $q = n \times k$, where n is the number of ports. Fig. 3 shows the frequency responses of Y(1,1) and Y(1,2) from the reduced models by hiePrimor and PRIMA. Fig. 4 shows the differences between hiePrimor and PRIMA. In all the test circuits, the accuracy of PRIMA and the hiePrimor are the almost the same numerically, although their results may be a little bit different from the exact one.



Fig. 3. Accuracy comparison of PRIMA and hiePrimor in Ckt1 when k = 4



Fig. 4. Difference between PRIMA and hie Primor in Ckt1 when $k\!=\!4$



Fig. 5. Accuracy comparison of hiePrimor in Ckt1 when k = 8

If we increase the value of k, we can obtain the more accurate models. Fig. 5 shows the comparison results for *Ckt 1* for k = 8. We can see the results are much better than the previous case when k = 4. Notice that the results from hiePrimor and PRIMA are still almost the same this time and their sizes after reduction are the same too.

Next, we compare hiePrimor with PRIMA in a single CPU setting in terms of reduction times. Table I shows the circuit statistics and comparison results of PRIMA and the hiePrimor. #Node is the number of nodes, #Sub is the number of subcircuits, #Ports is number of ports (terminals) of the circuit and '-' means out of memory or could not end in a reasonable time. We set reduction (block moment) order k = 4 in all the test circuits so that each circuit has the same reduced order (size) after reduction. It may be not accurate enough for k = 4 in all the circuits. But given that fact that hiePrimor gives almost the same accuracy as PRIMA, k = 4 is sufficient for us to compare reduction CPU times for them. The last column is the speedup of hiePrimor over PRIMA. We can see that hiePrimor can roughly run $5 \times$ faster than PRIMA for large scale circuits. It also shows that the hiePrimor has a better performance when the size of the circuits become larger. Note that such speed up is gained without sacrificing accuracy loss.

Typically, the more partition number, the more speedup gained. But we also need to consider the cost of combining all the lower level subcircuits into higher level. Also as we get more partitions, the ratio of terminal node and internal nodes may get smaller, which may hurt the reduction efficiency as the subcircuits may not be effectively reduced. So number of partitions need to be properly selected practically based on the actual situation. Table II shows the relationship between partition number and reduction time.

Another observation is that hiePrimor becomes more efficient than PRIMA when the number of ports increases. We use different number of ports for the same circuit (Ckt7) using both hiePrimor and PRIMA with the same reduction order. With smaller number of ports, hiePimor become more faster than PRIMA. Table III shows the reduction time comparison of PRIMA and hiePrimor for the same large circuit (Ckt7) with different number of ports. One reason is that ports are dispersed into subcircuits after partitioning and model order at the top level is already much smaller than original, while for PRIMA, its time complexity is highly related to the number of ports for given the same block moments k.

Further, we compare the two methods in the artificial parallel computing settings. It is relatively easy to parallelize our method because each subcircuit obtained by partitioning is stored in an independent spice format file. In parallel computing setting, the running time of hiePrimor is only the sum of the maximum subcircuit level reduction time among all the subcircuits and the top-level reduction time (for two level reduction). The results in Table IV show that we can have one order of magnitude or more speedup in this case. With more levels, we expect more speedup as more parallelism can be exploited.

TABLE II REDUCTION TIME FOR DIFFERENT NUMBER OF PORTS (k = 4, $q = n \times k$)

Test Ckts	#Parts = 2	#Parts = 4	#Parts = 8	#Parts = 16
Ckt5	116	100	71	60
Ckt6	374	251	128	99
Ckt7	383	298	204	154
Ckt8	675	394	257	176

TABLE III
REDUCTION TIME FOR DIFFERENT NUMBER OF PARTITIONS
(CKT7, $k = 4$, $a = n \times k$)

#Ports	PRIMA	hiePrimor	Speedup
8	189	56	3.38
16	339	96	3.53
32	863	154	5.60

TABLE IV REDUCTION TIME COMPARISON OF PRIMA AND HIEPRIMOR WITH PARALLEL COMPUTING SETTING ($k = 4, q = n \times k$)

Test Ckts	Max Sub (s)	Top (s)	Sum (s)	Speedup
Ckt1	2	0	2	2.50
Ckt2	3	1	4	4.00
Ckt3	3	1	4	8.00
Ckt4	5	1	6	11.50
Ckt5	6	1	7	35.43
Ckt6	10	1	11	36.46
Ckt7	17	3	20	43.15
Ckt8	14	1	15	_

8. CONCLUSION

In this paper, we have proposed a new projection based model order order reduction method. The new method, called, hiePrimor, applies divide-and-conquer strategy to reduce the reduction complexity and speed up the reduction process. It applies Krylov subspace projection based reduction on a partitioned circuit where subcircuits are reduced in a bottom-up way until top level circuit is reduced. Compared to the existing flat-projection-based reduction approaches, hiePrimor can give the same accuracy given the same reduction order. It also preserves the passivity of the reduced models as well. We also present a partitioning method and show that partitioning is important and min-span objective is required to archive best performance for hierarchical reduction. Experimental results demonstrate that hiePrimor can be significantly faster than PRIMA given a good partitioning and be much faster if parallel computing is used without loss of accuracy.

9. **REFERENCES**

- [1] "http://www.cs.umn.edu/ metis."
- [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)*, 2000, pp. 150–155.
- [3] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Pade approximation via the Lanczos process," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 14, no. 5, pp. 639–649, May 1995.
- [4] P. Feldmann and F. Liu, "Sparse and efficient reduced order modeling of linear subcircuits with large number of terminals," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2004, pp. 88–92.
- [5] E. J. Grimme, Krylov projection methods for model reduction (Ph.D. Thesis). University of Illinois at Urbana-Champaign, 1997.
- [6] K. J. Kerns and A. T. Yang, "Stable and efficient reduction of large, multiport rc network by pole analysis via congruence transformations," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, no. 7, pp. 734–744, July 1998.
- [7] B. Krauter and L. Pileggi, "Generating sparse partial inductance matrices with guaranteed stability," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 1995, pp. 45–52.
- [8] Y. Lee, Y. Cao, T. Chen, J. Wang, and C. Chen, "HiPRIME: Hierarchical and passivity preserved interconnect macromodeling engine for rlkc power delivery," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 6, pp. 797–806, 2005.
- [9] P. Liu, H. Li, L. Jin, W. Wu, S. X.-D. Tan, and J. yang, "Fast thermal simulation for runtime temperature tracking and management," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 12, pp. 2882–2893, Dec. 2006.
- [10] A. Odabasioglu, M. Celik, and L. Pileggi, "PRIMA: Passive reduced-order interconnect macromodeling algorithm," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 645–654, 1998.
- [11] J. R. Phillips, L. Daniel, and L. M. Silveira, "Guaranteed passive balanced transformation for model order reduction," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 8, pp. 1027–1041, 2003.
- [12] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 352–366, April 1990.
- [13] Z. Qin and C. Cheng, "Realizable parasitic reduction using generalized Y-Δ transformation," in Proc. Design Automation Conf. (DAC), 2003, pp. 220–225.
- [14] B. Salimbahrami and B. Lohmann, "Krylov subspace methods in linear model order reduction: introduction and invariance properties," Institute of Automation, University of Bremen, Tech. Rep., Aug. 2002.
- [15] B. N. Sheehan, "TICER: Realizable reduction of extracted RC circuits," in Proc. Int. Conf. on Computer Aided Design (ICCAD), 1999, pp. 200–203.
- [16] —, "Branch merge reduction of RLCM networks," in Proc. Int. Conf. on Computer Aided Design (ICCAD), 2003, pp. 658–664.
- [17] M. Silveira, M. Kamon, I. Elfadel, and J. White, "A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 1996, pp. 288–294.
- [18] S. X.-D. Tan, "A general s-domain hierarchical network reduction algorithm," in Proc. Int. Conf. on Computer Aided Design (ICCAD), 2003, pp. 650–657.
- [19] S. X.-D. Tan and L. He, Advanced Model Order Reduction Techniques in VLSI Design. Cambridge University Press, 2007.
- [20] A. Vandendorpe and P. V. Dooren, "On model reduction of interconnected systems," in *Proceedings International Symposium Math. Th. Netw. Syst.*, 2004.
- [21] N. Wang and V. Balakrishnan, "Fast balanced stochastic truncation via a quadratic extension of the alternating direction implicit iteration," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2005, pp. 801–805.
- [22] B. Yan, S. X.-D. Tan, P. Liu, and B. McGaughy, "SBPOR: second-order balanced truncation for passive model order reduction of rlc circuits," in *Proc. Design Automation Conf. (DAC)*, 2007, pp. 158–161.
- [23] G. Zhong, C. Koh, and K. Roy, "On-chip interconnect modeling by wire duplication," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2002, pp. 341–346.