# Modeling and Simulation for On-Chip Power Grid Networks by Locally Dominant Krylov Subspace Method *

Boyuan Yan†, Sheldon X.-D. Tan†, Gengsheng Chen‡ and Lifeng Wu§
†Department of Electrical Engineering, University of California, Riverside, CA 92521
‡ASIC & System State-Key-Lab, Microelectronics Dept., Fudan Univeristy, Shanghai, China, 200433
§Cadence Design Systems Inc., San Jose, CA 95134

## ABSTRACT

Fast analysis of power grid networks has been a challenging problem for many years. The huge size renders circuit simulation inefficient and the large number of inputs further limits the application of existing Krylov-subspace macromodeling algorithms. However, strong locality has been observed that two nodes geometrically far have very small electrical impact on each other because of the exponential attenuation. However, no systematic approaches have been proposed to exploit such locality. In this paper, we propose a novel modeling and simulation scheme, which can automatically identify the dominant inputs for a given observed node in a power grid network. This enables us to build extremely compact models by projecting the system onto the locally dominant Krylov subspace corresponding to those dominant inputs only. The resulting simulation can be very fast with the compact models if we only need to view the responses of a few nodes under many different inputs. Experimental results show that the proposed method can have at least 100X speedup over SPICE-like simulations on a number of large power grid networks up to 1M nodes.

## 1. INTRODUCTION

Signal integrity on the on-chip power distribution networks has become a limiting factor for designing high performance VLSI systems in today's deep submicron technology. The challenges for designing and verifying a reliable on-chip power deliver network lie in the lack of efficient simulation tools to handle the increasing size of the network circuits as technology scales. Conventional circuit analyzers (such as SPICE) cannot meet such demanding simulation tasks and efficient simulations are highly required to reduce the increasing design productivity gap in deep submicron design regime.

Different methods have been proposed in the past to address this problem in recent years [2–4, 6, 7, 14, 15]. The existing approaches include the hierarchical and macromodeling method [15], preconditioned iterative method [3], the multi-grid method [6], random walk method [11], localized simulation approach [4] and model order reduction based approaches like EKS [14], IEKS [2] and recently proposed ETBR [7], which is based on more accurate balanced truncation method. Those proposed approaches have improved the simulation efficiency of large power grid networks.

To make the P/G simulation truly scalable, model order reduction of power grid networks is a viable approach. The Krylov subspace techniques [5,8,12,13] have been proved to be very efficient to reduce large-scale interconnect circuits. Because of the moment-matching properties of Krylov-subspace, the reduced transfer function will agree with original transfer function up to the first $q$ derivatives on an expansion around some chosen point in complex plane (usually $s = 0$). However, while suitable for analysis of large-scale circuits, Krylov-subspace methods cannot generate compact models when there are a lot of inputs, which is the case for a power grid network or clock tree networks. For Krylov-subspace algorithms, the cost associated with model computation is directly proportional to the number of inputs, i.e. to the number of columns in the transfer function matrix. For example, in the PRIMA algorithm [8], if only two (block) moments are to be matched at each port, and the network has 1000 ports, the resulting reduced model will have 2000 states.

To mitigate this problem, extended Krylov-subspace techniques, EKS/IEKS [2, 14], and recently proposed extended truncated balanced realization, ETBR [7], were proposed, which construct a transformation matrix based on the dynamics of the circuits as well as the source excitations, thus avoiding the multi-port problem of model reduction. However, the drawback is also obvious: with different source excitations, the reduction process has to be performed for each input signal.

Recently, spatial *locality* of power grid analysis has been studied in [4,9]. Typically, the global power grid is extracted as an RC network and the cells or gates are modeled as independent current sources. Recently, it has been shown that the RC models of power grid networks are still valid models at least for 90nm [9]. The RC network can be deemed as a cascaded low-pass RC filter and for the filter, the attenuation of a ramp input signal is proportional to the exponential of the distance. Because of this exponential attenuation, if two nodes are geometrically far, they have very small electrical impact on each other. As a result, the voltage response of a node is only dominated by the excitations of a small number of inputs nearby. However, those proposed methods do not provide a systematic way to identify those dominant inputs.

In this paper, we propose a novel modeling and simulation scheme, which can automatically identify the dominant inputs for a given observed node in a power grid network. This enables us to build extremely compact models by projecting the system onto the locally dominant Krylov subspace corresponding to those dominant inputs only. The resulting simulation can be very fast with the compact models if we only need to view the responses of a few nodes under many different inputs. Our contributions lies in the followings: (1) We employ the *relative gate array*(RGA) [1] to build locally dominant Krylov subspace for reduction in power grid networks. (2) We show how the RGA can be efficiently computed if we are only interested in observing a few nodes and how it can be applied to power grid networks with both

*This work is funded by NSF CAREER Award CCF-0448534, UC Micro #07-105 via Cadence Design System Inc, the visiting scholar fund of ASIC & System State-Key-Lab of Fudan University.

voltage and current sources. (3) We demonstrate that RGA values computed in low frequency is a good indicator for wide-band frequency range as high frequency components attenuate more in general power grid circuits.

This paper is organized as follows: In Section 2, we review the Krylov-subspace methods. In Section 3, we introduce a RGA tool to evaluate the contribution of each input in terms of one output. Our new approach is described in Section 4. Experimental results are reported in Section 5 to demonstrate the effectiveness of our proposed method. Section 6 concludes the paper.

# 2. REVIEW OF KRYLOV SUBSPACE METHODS

Using any circuit-equation formulation method such as modified nodal analysis, a lumped, linear, time-invariant circuit can be described by the following system of first-order differential equations:

$$\begin{aligned} C\dot{x}(t) &= -Gx(t) + Bu(t) \\ y(t) &= L^T x(t) \end{aligned} \tag{1}$$

where the vector $x(t) \in R^n$ represents the circuit variables, the matrix $G \in R^{n\times n}$ represents the contribution of memoryless elements, such as resistors, $C \in R^{n\times n}$ represents contribution from memory elements, such as capacitors and inductors, $L \in R^{n\times q}$ is the output matrix, $B \in R^{n\times p}$ is the input matrix, $y \in R^q$ is the output of interest, and $u \in R^p$ represents excitations from independent sources.

Typically, we have $p \ll n$. Model reduction algorithms seek to produce a smaller system

$$\begin{aligned} \tilde{C}\dot{\tilde{x}}(t) &= -\tilde{G}\tilde{x}(t) + \tilde{B}u(t) \\ \tilde{y}(t) &= \tilde{L}^T\tilde{x}(t) \end{aligned} \tag{2}$$

where $\tilde{C}, \tilde{G} \in R^{r\times r}$, $\tilde{B}, \in R^{r\times p}$. Order $r$ is much smaller than the original order $n$, i.e. $r \ll n$, but the output $y(t)$ and $\tilde{y}(t)$ are approximately equal for inputs $u(t)$ of interest. This can be achieved by constructing a matrix $V$ whose columns span a useful subspace, and projecting the original equations onto the column space of $V$

$$\tilde{C} = V^T CV, \tilde{G} = V^T GV, \tilde{B} = V^T B, \tilde{L} = V^T L \tag{3}$$

Often the transfer functions

$$H(s) = L^T(sC + G)^{-1}B \tag{4}$$

are used as a metric for approximation.

Usually, projection matrix $V$ is constructed so that the columns span a Krylov subspace. The Krylov subspace $K_m(A, R)$ generated by matrix $A$ and matrix $R$, of order $m$, is the space spanned by the set of vectors ($R$, $AR$, $A^2R$, ..., $A^{m-1}R$). For example, a typical implementation (PRIMA) is to construct $V$ by using the Arnoldi algorithm, thereby spanning a Krylov subspace with

$$A = (G + s_0C)^{-1}C \quad R = (G + s_0C)^{-1}B \tag{5}$$

Because of the moment-matching properties of Krylov-subspace, the reduced transfer function $\tilde{H}(s)$ will agree with the original $H(s)$ up to the first $m$ derivatives on an expansion around some chosen point in the complex plane (usually $s_0 = 0$)

$$H(s) = \tilde{H}(s) + O((s - s_0)^m) \tag{6}$$

However, for Krylov-subspace based algorithms, the cost associated with model computation is directly proportional to the number of inputs, i.e. to the number of columns in the transfer function matrix.

However, due to the strong locality of RC mesh, the node we are interested in is only predominantly influenced by a small number of inputs nearby and those electrically distant inputs have little effect. We need a metric to evaluate the contribution of each input in terms of the node we are measuring and such a metric is introduced in the following section.

# 3. MEASUREMENT OF INTERACTION

## 3.1 Relative Gain Array

*Relative Gain Array* (RGA) is a matrix of interaction measures for all possible single-input single-output (SISO) pairings in an MIMO LTI system [1]. This concept has found widespread utility in process control, and as a system robustness measure. The RGA thus indicates the preferable variable pairings in a decentralized control system based on interaction considerations.

For a system $H(s)$ with $p$ inputs and $p$ outputs, there will be $p \times p$ relative gain elements

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1p} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2p} \\ \dots & \dots & & \dots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pp} \end{bmatrix} \tag{7}$$

and the relative gains between an output $y_i$ and an input $u_j$ are given by

$$\lambda_{ij} = \frac{g_{ij}\,|_u}{g_{ij}\,|_y} = \frac{(\triangle y_i/\triangle u_j)\,|_u}{(\triangle y_i/\triangle u_j)\,|_y} \tag{8}$$

where $g_{ij}$ is the gain of the respective transfer function $h_{ij}$. A simple $2 \times 2$ coupled system is shown in Fig. 1.

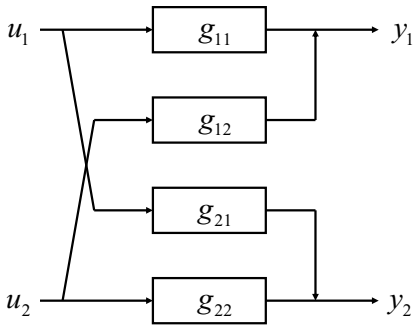


**Figure 1: A coupled $2 \times 2$ system.**

First, assume that all inputs except $u_j$ remain constant, a step change in input $u_j$ of magnitude $\triangle u_j$ will produce a change $\triangle y_i$ in output $y_i$. Thus, the gain between $u_j$ and $y_i$ when the other inputs are kept constant is given by

$$g_{ij}|_u = \frac{\triangle y_i}{\triangle u_j}|_u \tag{9}$$

which can be viewed as an open loop gain with respect to other inputs.

Second, when keeping all the outputs except $y_i$ constant, a step change in input $u_j$ of magnitude $\triangle u_j$ will result in another change in $y_i$. In this process, other outputs will also be affected due to cross-coupling. In order to keep them constant, we need to adjust other inputs correspondingly, which will also contribute to the change in $y_i$. The gain under the new set of conditions is denoted by

$$g_{ij}|_y = \frac{\triangle y_i}{\triangle u_j}|_y \tag{10}$$

which can be viewed as a closed loop gain with respect to other inputs.

Although the above gains are between the same pair of variables, they may have different values because they have been obtained under different conditions. If interaction exists, the change in $y_i$ due to a change in $u_j$ for the two cases (when other inputs and when other outputs are kept constant), will be different. The ratio,

$$\lambda_{ij} = \frac{g_{ij}|_u}{g_{ij}|_y} \tag{11}$$

defines the relative gain between the output $y_i$ and input $u_j$.

There are two extreme cases: first, if $\lambda_{ij} = 0$, $y_i$ is NOT influenced by $u_j$ at all; second, if $\lambda_{ij} = 1$, closed loop gain is equal to open loop gain, which means the interaction from other inputs is zero and $y_i$ is influenced by $u_j$ ONLY.

In fact, by taking the absolute value of each RGA element and taking the inverse for those larger than 1, the scaled elements will fall into the range of $[0, 1]$

$$\begin{array}{l} \lambda_{ij} = |\lambda_{ij}|(|\lambda_{ij}| \leq 1) \\ \lambda_{ij} = \frac{1}{|\lambda_{ij}|}(|\lambda_{ij}| > 1) \end{array} \tag{12}$$

The larger the scaled number is, the more important the corresponding input will be. For a given output $i$, the contribution of each input can be easily compared and those inputs can be arranged in a descending order in terms of their contribution. Usually, the output is only predominately influenced by a small number of inputs only.

## 3.2 Computation of RGA

The steady-state relative gain array of the system $H(s)$ at DC can be computed as follows

$$\Lambda(H) = H(0) \circ H(0)^{-T} \tag{13}$$

where $\circ$ denotes element-by-element multiplication (often called the Hadamard or Schur product), and $H^{-T}$ is the transpose of $H^{-1}$. We notice that we need to compute inverse of $H(0)$, which is expensive if the number of terminals, $p$ is large. But for power grid networks, we may be interested in observing a few nodes, say $q$ nodes. Then, the $H(0)$ is $q \times p$ matrix ($q \ll p$) and we can use *peudoinverse* to compute $H^{-T}$, which can be much more efficient. We give more detailed analysis in Subsection 4.3.

# 4. LOCALIZED MODELING SCHEME FOR POWER GRID ANALYSIS

## 4.1 Locality of an RC mesh

We first show that RC meshes modeling power grid networks have strong local property. The *locality* of RC mesh can be evaluated from the relative gain array. Given a $21 \times 21$ RC mesh, each node is viewed as an input and the relative gains from each node to the central node coordinate $(11, 11)$ are plotted in Fig. 2. It is easy to see the nodes closer to the central node will have more impact. In addition, the node is only predominately influenced by a small number of inputs nearby, which validates the observation in [4, 10]. As a result, RGA is a valid locality indicator to identify the most dominant inputs for a given output. With *locality*, we build a locally dominant reduction subspace (to be explained below) for a few observing nodes of interest, regardless of the size of the network, the number of inputs, and the patterns of input signals.

## 4.2 Localized compact models at DC and wide frequency range

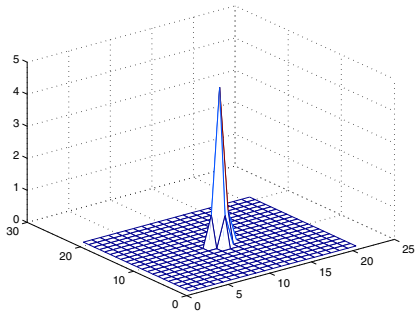


**Figure 2: Locality illustration of an RC mesh.**

A *localized* model of a power grid network can be formulated as follows

$$\begin{array}{l} C\dot{v}(t) = -Gv(t) + Bi(t) \\ y(t) = L_o^T v(t) \end{array} \tag{14}$$

where $C, G \in R^{n \times n}$ are capacitance and conductance matrices respectively, $B \in R^{n \times p}$ is the input matrix, $v(t) \in R^n$ and $i(t) \in R^p$ are the node voltage vector and input current vector.

Note that $L_o \in R^{n \times q}$ is the output matrix corresponding to $q$ nodes we are measuring. In our problem, we are only interested in a few nodes. The corresponding transfer function is

$$H(s) = L_o^T (Cs + G)^{-1} B \tag{15}$$

and the steady-state gain $H(0)$ becomes the DC gain $H_{DC}$

$$H_{DC} = L_o^T G^{-1} B \tag{16}$$

In this case, $H_{DC}$ is not a square matrix, then pseudoinverse of $H_{DC}^T$, $(H_{DC}^T)^+$ is used to compute the RGA

$$\Lambda(H) = H_{DC} \circ (H_{DC}{}^T)^+ \tag{17}$$

Now, we show that the RGA evaluated at DC ($s = 0$) is sufficient for other frequencies. The reason is that a power grid network can be deemed as a cascaded low-pass RC filter and for the low-pass filter, the attenuation of high frequency components is much faster than the attenuation of low frequency components, which means high frequency components tend to be more localized and a decision based on DC is conservative. So the results of RGA at DC are actually valid for all the frequency range for those RC networks.

In Fig. 3, we show that as the frequency increases (DC (top), $1G$(middle), $100G$(bottom)), RGA values become more locally concentrated around a few nodes for each output node.

## 4.3 Locally dominant Krylov subspace method

For the localized model, the projection matrix $V_{domi}$ is constructed so that the columns span a Krylov subspace $K_m(A, R_{domi})$, where

$$A = G^{-1}C, R_{domi} = G^{-1}B_{domi} \tag{18}$$

In our approach, instead of all the inputs, $B_{domi}$ is only composed of a small number of *dominant* inputs corresponding to those outputs of interest. This leads to the newly proposed locally dominant Krylov subspace reduction:

$$\begin{array}{l} \tilde{C}_{loc} = V_{domi}^T C V_{domi}, \tilde{G}_{loc} = V_{domi}^T G V_{domi} \\ \tilde{B}_{loc} = V_{domi}^T B, \tilde{L}_{loc} = V_{domi}^T L_o \end{array} \tag{19}$$

where $colspan(V_{domi}) \subseteq K_m(A, R_{domi})$. If we are interested in a number of nodes, then the system is projected

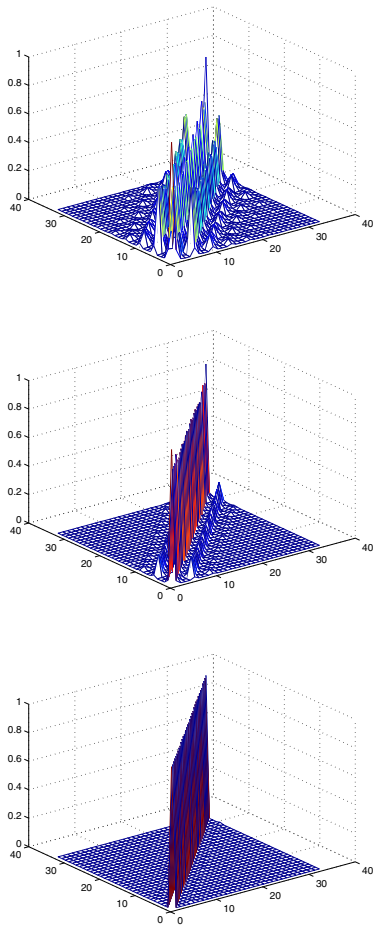

**Figure 3: RGA computed at different frequencies.**

onto a subspace which is the union of the locally dominant subspaces of those nodes. We remark that if the nodes of interests are limited to one local region, the order of reduced model may not increase even if more nodes are to be measured. The reason is that those nodes share many inputs (thus their subspaces). The proposed locally dominant Krylov subspace algorithm is shown in Fig. 4.

Our new method can be viewed by exploiting both temporal and spatial information to reduce the system complexity. Existing Krylov-subspace methods only take into consideration *frequency* (*thus temporal*) information. A reduction can be achieved because the frequency components are not viewed as equally important and only the dominant frequency subspace is preserved. However, those methods fail to consider *spatial* information and all the inputs are implicitly assumed to be equally important and fully preserved. In fact, if we focus on a few nodes or a local region, most inputs are insignificant owning to strong locality. As a result, the new approach can generate much more compact models than the existing temporal-only reduction methods.

## 4.4 Computational complexity analysis

For an RC circuit of order $n$ and with $p$ ports, it will take $O(n^\beta)$ to compute the DC moment $H_{DC}$ as matrix $G$ is very sparse in general, where, typically, $1 \leq \beta \leq 1.5$ for a $n \times n$ sparse matrix.

The computation of pseudoinverse is based on the singular value decomposition (SVD) of the matrix ${H_{DC}}^T$, which is a

THE NEW LOCALLY DOMINANT KRYLOV SUBSPACE SIMULATION ALGORITHM:
Input: $G, C, B, u(t)$, set of observation nodes
Output: transient waveform at the observation nodes

1. Solve $GM_0 = B$ for $M_0$ (DC moment)
2. Compute $H_{DC} = {L_o}^T M_0$
3. Compute RGA $\Lambda(H)$=$H_{DC} \circ (H_{DC}^T)^+$
4. Scale the RGA values to the range of $[0, 1]$ and arrange them in a descending order in terms of the contribution to each output
5. For those outputs of interests, determine the corresponding dominant input matrix $B_{domi}$ based on RGA
6. Compute localized reduced models by projection
7. Compute transient response on the reduced models for $u(t)$

**Figure 4: The locally dominant Krylov subspace simulation method for power grid network analysis.**

$p \times q$ matrix, where $q$ is the number of nodes we are interested in. Since we are only interested in a small number of nodes, we have $q \ll p$ and the cost is $O(q^2 p)$. Assume that $p \ll n$ as this is the typical case, the reduction process is still dominated by $O(n^\beta)$.

The localized model will take about $O(nr^2 + rn^\alpha + n^\beta)$ to reduce using Krylov-subspace method, where $r$ is the reduced order and $1 \leq \alpha \leq 1.2$ for sparse matrices. $n^\beta$ is the same cost as the computation of the DC moment. The transient simulation of the reduced system takes about $O(r^3 + r^2 m)$ where $m$ is the number of time steps in time domain.

Since the reduced order $r$ is a very small number, the total cost is still dominated by $O(n^\beta)$, which is one DC solution of the original network. The reduced models can be used for many inputs without further solving the network again.

## 4.5 Partitions of input signals for RGA computation

To efficient compute RGA values for an RC mesh network, we find that we will be better off if we can compute RGA for current source inputs only (without voltage source inputs) as the two types of the signals are quite different in terms of magnitude.

Given a power grid network, as shown in Fig. 5, there are a small number of voltage sources and a huge number of current sources. The voltages supplies are DC with a constant value and the current sources are pulse currents generated when the gates are switching.

Assume that we are interested in a particular node, the voltage responses at that node can be decomposed into two parts owning to superposition:

$$V = V_{voltage} + V_{current} \tag{20}$$

where the first part is the responses of DC voltage sources, which is *static*, and the second part is the responses of independent current sources, which is *dynamic*. And the power grid network can be decomposed into two parts correspondingly, which are shown in Fig. 6 and Fig. 7.

$V_{voltage}$ is static and it can be easily obtained via one DC analysis as shown in Fig. 6 because the number of voltage inputs is small and the conductance matrix is sparse. To compute the RGA, we only use the dynamic part shown in Fig. 7, which can give better indication of the RGA among all the current inputs.

# 5. EXPERIMENTAL RESULTS

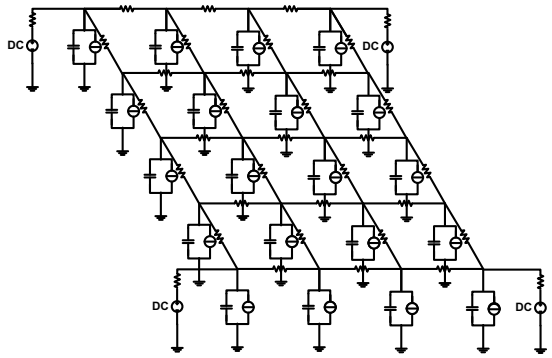


**Figure 5: Power grid model**

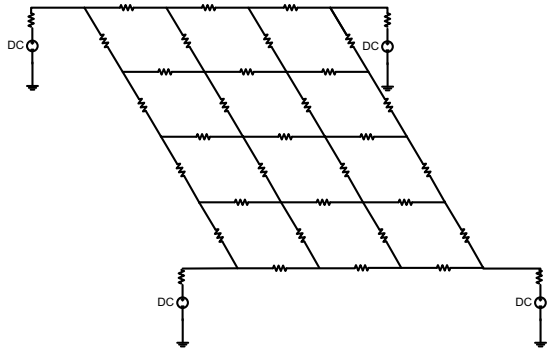


**Figure 6: Power grid model: static part**

**Table 1: Test circuits**

| Test Ckts | #Nodes | #Sources |
|---|---|---|
| Ckt1 | 1,160 | 40 |
| Ckt2 | 10,000 | 100 |
| Ckt3 | 10,000 | 1,000 |
| Ckt4 | 102,400 | 1,024 |
| Ckt5 | 250,000 | 2,500 |
| Ckt6 | 500,000 | 5,000 |
| Ckt7 | 1,000,000 | 10,000 |

The proposed method has been implemented in Matlab 7.0 and tested on an Intel quad-core workstation with 16GB memory. All the test circuits are randomly generated RC power grid networks up to one million nodes ($R$ on the order of $ohm$ and $C$ on the order of $pf$), as shown in Table 1. Efficient matrix computations benefit from sparse matrix structure and a parser implemented by Python. We compare our method with the existing Krylov-subspace method PRIMA.

The following simulation results are from Ckt3, a power grid network with 10000 nodes and 1000 current sources. However, similar results can be observed from any test circuit in Table 1. Fig. 8 shows the transient responses at a particular node of interest. The input waveforms are randomly generated with typical pattern in a power grid analysis. A localized model of order 3 can approximate original model very well by matching three moments of the most dominant input determined by RGA. In contrast, the performances of

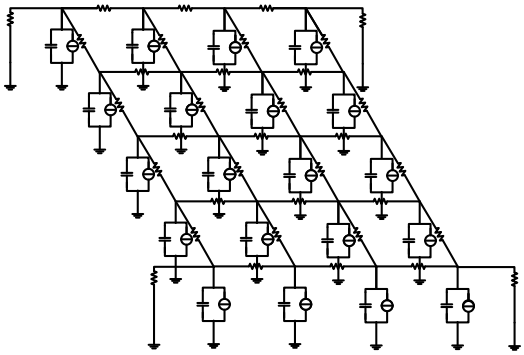

**Figure 7: Power grid model: dynamic part**

PRIMA reduced models of order 3 and 600 are also shown in the same figure.

Notice that PRIMA-like reduction algorithm is not suitable for power grid analysis, we just show how inefficient those reduction methods can be if we need a good accuracy. This is evidenced by the fact that PRIMA reduced model of order 600 still has a noticeable error. Therefore, our new approach can lead to much more compact and accurate models, which are suitable for simulations.

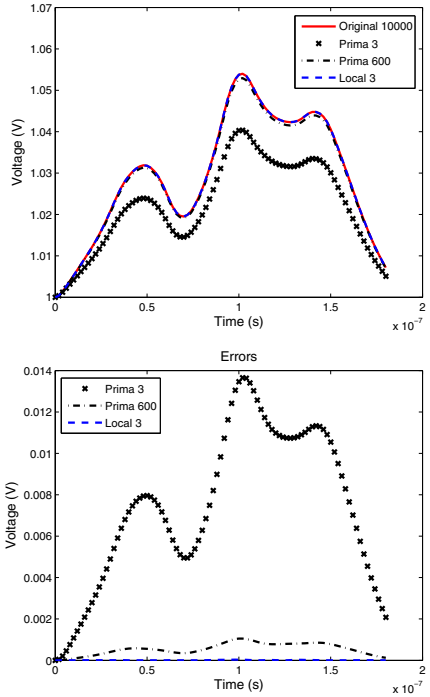


**Figure 8: The simulation results with typical power grid inputs.**

It is known that, for a circuit with a large number of inputs, extended Krylov subspace (EKS) method [14], was also proposed to generate compact models by constructing a transformation matrix based on the dynamics of the circuit as well as the source excitations. However, for those methods, the modeling process depends on the input signal. As a result, once the pattern of input signals has been changed,

the model needs to be rebuilt. More important, they become less useful when input information is unavailable *a priori*.

Our localized modeling scheme does not rely on the pattern of input signals. To illustrate this point, input signals of different styles are applied. For example, sawtooth waveforms are used and the simulation results are shown in Fig. 9. Note that, the localized model is more accurate than PRIMA reduced model of order 1000.

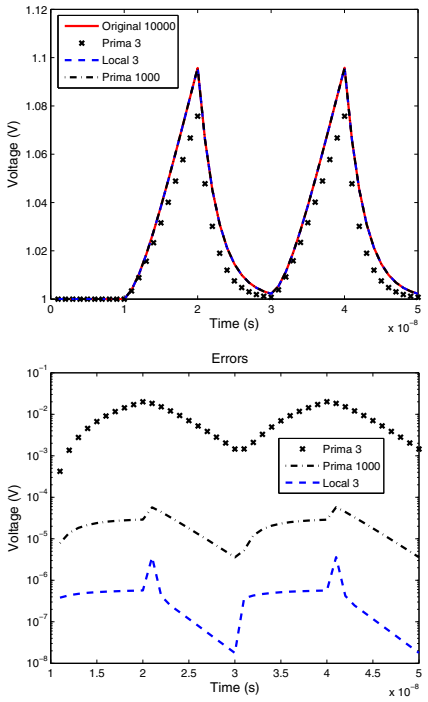


**Figure 9: The simulation results with sawtooth waveforms on Ckt3.**

The simulation times between the *SPICE*-like method and the localized method (*Loc method*) are compared in Table 2. In all test examples, a localized model of order 3 is good enough to approximate the original circuits. Since the size of the reduced model is a constant, the larger the original circuit is, the more significant the speedup will be.

**Table 2: CPU times (in seconds) comparison for transient simulation ($r = 3$)**

| Test Ckts | Loc method (s) | SPICE (s) | Speedup |
|---|---|---|---|
| Ckt1 | 0.01 | 0.18 | 18.0 |
| Ckt2 | 0.02 | 3.07 | 153.5 |
| Ckt3 | 0.17 | 3.68 | 21.7 |
| Ckt4 | 0.18 | 53.74 | 298.6 |
| Ckt5 | 0.43 | 152.18 | 353.9 |
| Ckt6 | 0.87 | 281.00 | 323.0 |
| Ckt7 | 1.76 | 605.98 | 344.3 |

## 6. CONCLUSION

In this paper, the strong locality of power grid network was explored to speed up the simulation of a few nodes in a large power grid network. We have proposed a locally dominant Krylov subspace method to obtain extremely compact models before the simulation by projecting the system onto the Krylov subspace corresponding to those dominant inputs only. The method enables the efficient reduction of a linear system with massive ports without considering the pattern of input waveforms. The resulting simulation can be very fast with the compact models if we only need to view the responses of a few nodes under many different input waveforms. Experimental results show that the proposed method can have at least 100X speedup over SPICE-like simulations on a number of large power grid networks up to 1M nodes.

## 7. REFERENCES


[1] E. H. Bristol, "On a new measure of interaction for multivariable process control," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 132–134, 1966.

[2] Y. Cao, Y. Lee, T. Chen, and C. C. Chen, "HiPRIME: hierarchical and passivity reserved interconnect macromodeling engine for RLKC power delivery," in *Proc. Design Automation Conf. (DAC)*, 2002, pp. 379–384.

[3] T. Chen and C. C. Chen, "Efficient large-scale power grid analysis based on preconditioned Krylov-subspace iterative method," in *Proc. Design Automation Conf. (DAC)*, 2001, pp. 559–562.

[4] E. Chiprout, "Fast flip-chip power grid analysis via locality and grid shells," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, Nov. 2004, pp. 485–488.

[5] 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.

[6] J. N. Kozhaya, S. R. Nassif, , and F. N. Najm, "A multigrid-like technique for power grid analysis," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 10, pp. 1148–1160, Oct. 2002.

[7] D. Li, S. X.-D. Tan, and B. McGaughy, "ETBR: Extended truncated balanced realization method for on-chip power grid network analysis," in *Proc. European Design and Test Conf. (DATE)*, 2008, pp. 432–437.

[8] 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.

[9] S. Pant and E. Chiprout, "Power grid physics and implications for CAD," in *Proc. Design Automation Conf. (DAC)*, 2006, pp. 199–204.

[10] ——, "Power grid physics and implications for cad," in *Proc. Design Automation Conf. (DAC)*, 2006, pp. 24–28.

[11] H. F. Qian, S. R. Nassif, and S. S. Sapatnekar, "Random walks in a supply network," in *Proc. Design Automation Conf. (DAC)*, 2003, pp. 93–98.

[12] 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.

[13] S. X.-D. Tan and L. He, *Advanced Model Order Reduction Techniques in VLSI Design*. Cambridge University Press, 2007.

[14] J. M. Wang and T. V. Nguyen, "Extended Krylov subspace method for reduced order analysis of linear circuit with multiple sources," in *Proc. Design Automation Conf. (DAC)*, 2000, pp. 247–252.

[15] M. Zhao, R. V. Panda, S. S. Sapatnekar, and D. Blaauw, "Hierarchical analysis of power distribution networks," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 2, pp. 159–168, Feb. 2002.