Extended Krylov Subspace Method for Reduced Order Analysis of Linear Circuits with Multiple Sources

Janet M. Wang* and Tuyen V. Nguyen**
*Dept. of EECS, University of California at Berkeley, CA94706
**IBM Austin Research Laboratory, Austin, TX 78758
meilingw@uclink4.berkeley.edu tuyenn@austin.ibm.com

Abstract

This paper presents a method for general reduced order analysis of linear circuits with a large number of independent sources. This type of circuit is used to model the power grid in power supply noise analysis for example. The large size of the linear circuit model renders circuit simulation inefficient. The large number of independent sources makes conventional multi-port model reduction ineffective. In order to address these problems, this paper proposes an extended Krylov subspace method which constructs a transformation matrix based on the dynamics of the circuit as well as the source excitations, thus avoiding the multi-port problem of model reduction. The transformation matrix is then used to reduce the given circuit to a smaller circuit model, which allows for more efficient analysis.

1.Introduction

Signal integrity has become an important issue in today ‘s deep submicron design. The signal degradation comes from a variety of sources which include coupled noise from adjacent signals, reflection noise from impedance mismatches, substrate coupling noise, and power supply noise [1]. In particular, the power supply noise has become more important for new technology such as Silicon-On-Insulator (SOI). As a result, the design of the power distribution network and the associated electrical analysis has become very critical for chip design.

In general, the challenge in the design of the power grid is to ensure a stable DC level while supplying large AC currents to various switching gates. In CMOS circuit, current is drawn from the power network when logic gates make a transition. This current flows from the contact to the power grid to the ground contact of the gate. This current flow will cause voltage variation at the power and ground contacts due to parasitic resistance, capacitance, and inductance of the power grid. The electrical analysis of the entire chip to assess this voltage variation is prohibitively expensive. The common approach to make this problem manageable is to decouple the analysis of the power grid from the analysis of the logic gates. First, the logic gates are simulated assuming a perfect DC voltage for the power contact to produce the current signatures. These current signatures are used to model the logic gates as linear time dependent current sources. The power grid is then simulated with these current sources as circuit excitations to assess the voltage variation at every node of the grid.

The electrical analysis of the power grid then becomes a linear transient analysis problem. For post layout analysis, the linear circuit model of the power grid may contain millions of variables. The large size of this linear circuit model makes circuit simulation computationally expensive. On the other hand, conventional multiport model reduction techniques [2][3][4] are rendered ineffective because of the large number of ports due to the independent current sources modeling the logic gates. For early design and analysis of the power grid, the size of the circuit model can be reduced, but the circuit is analyzed more often. Therefore, it is important to improve the efficiency of the linear transient analysis of the power grid. A common technique is to employ fixed time step integration such as Backward Euler or Trapezoidal [5]. The advantage of fixed time step integration is that it requires only one LU decomposition of the circuit matrix and one backward-forward substitution (BFS) at each time step. This method definitely improves that efficiency of the transient analysis of the power grid when the voltage variation has largely high frequency content. i.e the voltage is changing constantly and rapidly. However, since a sufficiently small timestep must be chosen a priori for this method, it becomes inefficient for problems when low frequency content is more important, i.e. when the slow variation of the responses is of concern and when there are long intervals of steady state. Other approaches aim at extending model reduction techniques to address the problem of multiple sources. A symbolic algorithm is proposed in [6]. However, this approach is limited to delayed impulse excitations only. The approach proposed in [1] allows for more general piecewise linear representation of the sources. However, this approach suffers from the numerical ill conditioning problem of explicit moment matching.

In order to address these problems, this paper presents a general method for the reduced order analysis of large linear circuits with multiple sources. It is organized as follows. Section 2 presents an overview of various methods for power grid analysis to provide the background for the proposed extended Krylov subspace method. The basic idea and the mathematical details of the proposed method are then described in section 3. Practical considerations such as scaling and error control are described in section 4. Results are presented in section 5 to substantiate the accuracy and efficiency of the proposed method. Section 6 concludes the paper and suggests some areas for future work.

2.Power grid analysis

In general, power grid analysis requires a full chip simulation of both the transistors and the interconnects. This is a prohibitively expensive process. Therefore, in a typical chip design, the problem is separated into a local macro level and global interconnect level. Groups of nonlinear devices, i.e. transistors, performing a particular sub-function are grouped together with small local interconnect into a module called a circuit macro. Following this, different functional macros are physically placed and interconnected with longer metal interconnect lines, called global interconnects, into the final chip assembly. Each macro, in addition to being fed by the signal lines from other macros, is fed on the top by Vdd and Gnd buses. A small example macro is shown in Figure 1.

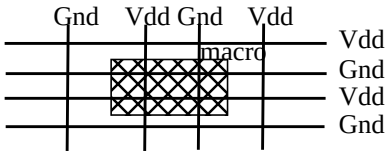


Fig. 1 *An example macro fed by Vdd and Gnd buses at four input points on the grid.*

When full-chip power grid analysis is performed, each macro is simulated with typical and expected input vectors and its typical (time-dependent) average current draw at its “pins” is calculated (assuming

DAC 2000, Los Angeles, California

a “perfect”, i.e. stable, Vdd supply). The interest of the designer is to estimate whether the switching activities of all the these macros in a typical and/or worst case condition will cause voltage droops on Vdd lines or ground bounces. In order to manage the computational cost, each macro is modeled as one or more simple linear time-dependent current sources using the results of the stand-alone simulation. The global power grid is extracted and modeled as an R(L)C network and the macro current sources are attached to the grid as shown in Fig. 2 to form a completely linear network.

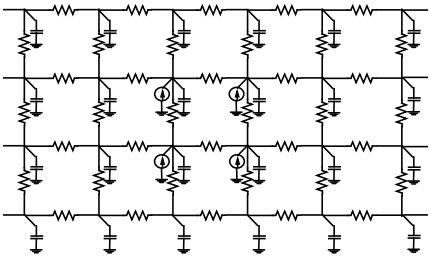


Fig. 2 *A model of the macro drawing current from the power network at four points.*

A transient analysis of this network allows the assessment of power supply noise and ground bounce. If necessary, another iteration of the macro design and analysis is performed. This process is continued until the supply noise is brought within the specified noise margin. For the purpose of electrical analysis, the circuit model of the power grid as shown above can be described by the following MNA equation

$$Gx + C\dot{x} = Bu \tag{1}$$

where C and G represent the frequency independent nxn conductance and susceptence MNA matrices, x is the vector of node voltages and inductor/source currents, u is the vector of independent sources, and B is the input selector matrix mapping the sources to the internal states. In on-chip power analysis, matrices C and G are large, and x consists of thousands or millions of state variables while the number of sources contained in u is in the range of hundreds or thousands.

2.1. Multiport reduced order modeling of the power grid

Model order reduction is another approach for improving the efficiency of linear circuit modeling and analysis. The present state of the art model reduction algorithms are typically based on Krylov subspace methods. In such a method, the following subspace is generated

$$span\{\boldsymbol{G}^{-1}\boldsymbol{B}, \boldsymbol{G}^{-1}\boldsymbol{C}\boldsymbol{G}^{-1}\boldsymbol{B}, \ldots, (\boldsymbol{G}^{-1}\boldsymbol{C})^{n}\boldsymbol{G}^{-1}\boldsymbol{B}\ldots\}$$

The matrix representing this subspace is used to transform the original system as given by Equation (1) to a smaller system of the same form [3][4]

$$\hat{G}z + \hat{C}\dot{z} = \hat{B}u \tag{2}$$

However, the bottleneck in such approaches is the number of sources in the input vector u. The longer this vector (i.e. the more unique input vectors exist), the bigger the Krylov subspace to be used because of the bigger B matrix. For each extra column in the B matrix, another backward-forward substitution (BFS) is required in the computation of the Krylov subspace. Therefore, the cost of the reduction algorithm is directly proportional to the number of independent sources (i.e. ports). As a result, model order reduction techniques are only applicable for a small number of sources in power grid analysis.

2.2. Input dependent explicit moment matching

In order to avoid the problems associated with multiport model reduction, an input dependent explicit moment matching [7] has been developed. In this approach, the moment representation of the voltage response is computed from the moment representation of the inputs as well as the moment representation of the system in the frequency domain. The moment representation of the output voltage responses is then used to construct the Padé approximation for computing the responses in the time domain. This approach is described in more detail below. In the frequency domain, Equation (1) can be rewritten as

$$(G + sC)X(s) = Bu(s) \tag{3}$$

Typically, the independent sources contained in u are represented by PWL functions. And in the frequency domain, these PWL sources can be modeled as a sum of delayed ramps:

$$u(s) = \frac{1}{s^2}\sum_{i=1}^{\bar{k}} r_i \exp(-\tau_i s)$$

Let $\bar{u}(s) = \sum_{i=1}^{\bar{k}} r_i \exp(-\tau_i s)$ then the moment representation of the input vector can be written as

$$\bar{u}(s) = \bar{u}_0 + \bar{u}_1 s + \bar{u}_2 s^2 + \ldots + \bar{u}_k s^k + \ldots$$

where $\bar{u}_k$ is the kth moment of $\bar{u}(s)$ and it can be calculated as

$$\bar{u}_k = \sum_{i=1}^{\bar{k}} (-1)^{k-1} \frac{r_i \tau_i^{k-1}}{(k-1)!}$$

Let $\bar{x}(s) = s^2 x(s) = (m_0 + m_1 s + m_2 s^2 + \ldots)$ be the moment representation of the unknown responses, Equation (3) can be rewritten as

$$(G + sC)(m_0 + m_1 s + m_2 s^2 + \ldots) = B(\bar{u}_0 + \bar{u}_1 s + \bar{u}_2 s^2 + \ldots)$$

This leads to a recursive relationship between the moments of $\bar{x}(s)$ and $\bar{u}(s)$: $Gm_0 = B\bar{u}_0 \qquad Gm_i + Cm_{i-1} = B\bar{u}_i$. Note, that the output response can be separated into:

$$x(s) = \left(\frac{1}{s^2}m_0 + \frac{1}{s}m_1\right) + (m_2 + sm_3 + \ldots + s^n m_n + \ldots) \tag{4}$$

The first parenthesis contains the ramp and the step responses which can be converted to a time domain ramp and a time domain step function. The next parenthesis can be approximated as a partial expansion in terms of poles and residues by an explicit Padé approximation. This approximation is obtained by mapping the

Taylor series to a rational function approximation and then solving two sets of linear equations, one for the coefficients of the denominator and one for those of the numerator [2]. The pole/residue form in the frequency domain is translated into the time domain as $\sum_i k_i e^{p_i t}$. This approach using explicit moment matching suffers the well known numerical stability problem especially when higher order moments are required. In order to overcome this problem, an extended Krylov subspace method is proposed and described in the next section.

3.Extended Krylov subspace method

The basic idea of this approach is similar to the standard Krylov subspace method in [4]. The proposed method seeks to build incrementally an orthogonal basis that spans the same subspace as represented by the finite moment representation of the responses. The computed orthogonal basis is then used to project the original system to the smaller subspace spanned by the basis. The difference is that in standard Krylov subspace method, the inputs are impulse sources for computing the transfer function of the system. These impulse sources are constant in the frequency domain and they are captured in the initial vector for the Krylov iteration. There is no further contribution at subsequent steps of the iteration. This feature allows a simple orthogonalization scheme. However, for the extended Krylov subspace method (EKS) proposed in this paper, the sources are general PWL sources with full series expansion, i.e. not a constant term, in the frequency domain. Therefore, the contribution of source moments must be taken into account at every step of the iteration. This fact presents a major complication for the orthogonalization process. This issue will be clarified later in this section.

In the EKS method, the reduced system is then solved in the time domain by standard integration algorithms. The resulting solution is then projected back to the original space to provide an approximate solution to the original system. In other words, let $V = \{\hat{r}_0, \hat{r}_1, \ldots, \hat{r}_k\}$ be an orthogonal basis of the moment subspace $\{m_0, m_1, \ldots, m_k\}$. Then, the original system described by Equation (1) can be reduced to a smaller system of a similar form as given in Equation (2) where $\hat{G} = V^T G V$, $\hat{C} = V^T C V$, and $x(t) = Vz(t)$. Equation (2) can be solved for $z(t)$ in the time domain by any standard integration algorithm. The solution of this reduced system can then be projected back to the original space to give an approximate solution as $\tilde{x}(t) = Vz(t)$. Since the extended Krylov subspace method employs a congruent transformation in the projection, the passivity of the reduced system is ensured as proven in [4].

3.1. Incremetal Orthogonalization algorithm

In order to illustrate the basic idea, the first three steps of the algorithm is described in detail.

At $i = 0$:

$$b_0 = B\bar{u}_0,\ m_0 = G^{-1}b_0,\ \hat{r}_0 = \alpha_0 m_0,\ \alpha_0 = norm(m_0)$$

At $i = 1$:

$$r_1 = \alpha_0 G^{-1}(B\bar{u}_1 - Cm_0) = \alpha_0 m_1 \quad (5)$$

However, during the orthonormal vector calculation, m_0 and /or m_1 are never used in order to avoid direct moment generation for the reason described in the previous section. Instead the new vector is computed as

$$r_1 = G^{-1}(\alpha_0 B\bar{u}_1 - C\hat{r}_0) \quad (6)$$

The orthogonal component of this vector with respect to the previously computed vector(s) can be generated as follows. First let $h_{1,0} = \hat{r}_0^T r_1$. Then the orthogonal component is given by

$\bar{r}_1 = r_1 - h_{1,0}\hat{r}_0$ and the new normalized basis is $\hat{r}_1 = \frac{\bar{r}_1}{norm(\bar{r}_1)}$ with $\alpha_1 = \frac{1}{norm(\bar{r}_1)}$.

Similarly, at $i = 2$:

$$r_2 = \alpha_0\alpha_1 G^{-1}(B\bar{u}_2 - Cm_1) = \alpha_0\alpha_1 m_1 \quad (7)$$

This procedure can be extended easily to $i = k$ case,

$$r_k = \left(\prod_{j=0}^{k-1}\alpha_j\right)G^{-1}(B\bar{u}_k - Cm_{k-1}) = \left(\prod_{j=0}^{k-1}\alpha_j\right)m_k \quad (8)$$

Equation(5) and Equation(6) are useful for the proof of Theorem 1 but are not used in the practical calculation. In the algorithm, r_2 is calculated as

$$r_2 = G^{-1}(\alpha_0\alpha_1 B\bar{u}_2 - C(\hat{r}_1 + \alpha_1 h_{1,0}\hat{r}_0)) \quad (9)$$

Then, let $h_{2,0} = \hat{r}_0^T r_2$, $h_{2,1} = \hat{r}_1^T r_2$.

Thus $\bar{r}_2 = r_2 - h_{2,0}\hat{r}_0 - h_{2,1}\hat{r}_1$ with $\hat{r}_2 = \frac{\bar{r}_2}{norm(\bar{r}_2)}$ and $\alpha_2 = \frac{1}{norm(\bar{r}_2)}$.

From the previous steps, it is easy to get the relationship between $\hat{r}_i$ and m_i,

$$\hat{r}_i = \alpha_i\left(\prod_{j=0}^{i-1}\alpha_j m_i - \sum_{j=0}^{i-1}h_{i,j}\hat{r}_j\right) \quad (10)$$

This orthogonalization procedure can be summarized in the following algorithm

***Algorithm 1*:**

$b_0 = B\bar{u}_0, m_0 = G^{-1}b_0, \hat{r}_0 = \alpha_0 m_0$

for i=1:number of columns in $\bar{U}$

$$r_i = G^{-1}\left(\prod_{j=0}^{i-1}\alpha_j B\bar{u}_i - C\left(\hat{r}_1 + \alpha_{i-1}\sum_{j=0}^{i-1}h_{i-1,j}\hat{r}_j\right)\right)$$

for k=1:i-1 $h_{i,k} = \hat{r}_k^T r_i$ end

$$\bar{r}_i = r_i - \sum_{j=0}^{i}h_{i,j}\hat{r}_j$$

if $norm(\bar{r}_i) < etol$ break;

else $\hat{r}_i = \frac{\bar{r}_i}{norm(\bar{r})}$, $\alpha_i = \frac{1}{norm(\bar{r})}$

end

end

Theorem 1:[8] The orthonormal vectors generated by the above algorithm span the same space as the original moment vectors: $span\{\hat{r}_0, \hat{r}_1, \ldots, \hat{r}_n\} = span\{m_0, m_1, \ldots, m_n\}$.

3.2. Multipoint Extension

In the last subsection, we discuss the extended Krylov subspace method with input sources' waveform modeled by Taylor series expansion at $s = 0$. This is sufficiently accurate when there are only slowly changing input sources. However, in order to capture higher frequency content of the input waveforms, moments at other frequencies should also be considered. Let $s = \sigma_0 + \sigma$ where σ_0 is a frequency shift. Then Equation (3) can be written as

$$((G + \sigma_0 C) + \sigma C)\left(m_0 + m_1\sigma + m_2\sigma^2 + \ldots\right) = B\left(\bar{u}_0 + \bar{u}_1\sigma + \bar{u}_2\sigma^2 + \ldots\right)$$

The recursive relationship between adjacent moments holds for this case. However, the input source waveforms modeled by a series of ramp functions $r_i \exp(-\tau_i s)$ are now modeled by $\exp(-\tau_i\sigma_0) r_i \exp(-\tau_i\sigma)$ at shifted frequency point σ_0. The moments of the input sources are now scaled $\exp(-\tau_i\sigma_0)$. The congruent transformation matrix based on the implicit moment derivation at different shifted frequency points can then be computed in a similar manner. The procedure can be best described by the subsequent algorithm:

Algorithm 2: Inputs: expansion points σ_k and expansion order at each expansion point

$N = 0$ (N --- number of existing orthonormal vectors)

for k=1: number of expansion points

$b_0 = Bu_0(\sigma_k)$, $m_0(\sigma_k) = G^{-1} b_0$, $\hat{r}_0(\sigma_k) = \alpha_0(\sigma_k) m_0(\sigma_k)$

$N = N + 1$

for i=1:order at every expansion points

$$r_i(\sigma_k) = G^{-1}\left(\prod_{j=0}^{i-1} \alpha_j(\sigma_k) B\bar{u}_i(\sigma_k) - C\left(\hat{r}_1(\sigma_k) + \alpha_{i-1}(\sigma_k) \sum_{j=0}^{i-1} h_{i-1,j}\hat{r}_j(\sigma_k)\right)\right)$$

for l=1:N $h_{i,l} = \hat{r}_l^T r_i(\sigma_k)$ end

$$\bar{r}_i(\sigma_k) = r_i(\sigma_k) - \sum_{j=0}^{N} h_{i,j}\hat{r}_j(\sigma_k)$$

if $norm(\bar{r}_i(\sigma_k)) < etol$ break;

else

$$\hat{r}_i(\sigma_k) = \frac{\bar{r}_i(\sigma_k)}{norm(\bar{r}_i(\sigma_k))} \qquad \alpha_i(\sigma_k) = \frac{1}{norm(\bar{r}_i(\sigma_k))}$$

end

$N = N + 1$

end

end

The transform matrix for m shifted frequency points has the form

$$V = \left\{\hat{r}_0^{(\sigma_0)}, \hat{r}_1^{(\sigma_0)}, \ldots, \hat{r}_{k(\sigma_0)}^{(\sigma_0)}, \hat{r}_0^{(\sigma_1)}, \hat{r}_1^{(\sigma_1)}, \ldots, \hat{r}_{k(\sigma_1)}^{(\sigma_1)}, \ldots\ldots, \hat{r}_0^{(\sigma_m)}, \hat{r}_1^{(\sigma_m)}, \ldots\right\}$$

where $\sigma_0, \sigma_1, \ldots\ldots, \sigma_m$ are m shifted frequency expansion points.

3.3. Solution of the system

Let $V = \{\hat{r}_0, \hat{r}_1, \ldots, \hat{r}_k\}$, n is the order or column number of matrix V. If we use V as transform on the original system, we get $\tilde{x}(s) = V\hat{x}(s)$ then $\hat{G} = V^T G V$, $\hat{C} = V^T C V$, $\hat{B} = V^T B$. $\hat{G}$ has dimension n ,the number of column vectors in V.

If

$$z(s) = \frac{1}{s^2}\hat{x}(s) \tag{11}$$

then we get

$$(\hat{G} + s\hat{C})z(s) = \hat{B}u(s) \tag{12}$$

This system can be solved in the time domain by standard integration algorithms such as Trapezoidal or Backward Euler. Different from the typical Krylov subspace based method, the resulting vector does not have physical meaning unless it is projected back to the original space to provide the approximate solution

$$x(t) = Vz(t) \tag{13}$$

we get all the node voltage or branch current we want.

Theorem 2: [8]The reduced order model preserves the moments of the original system. That is: define $V = \{\hat{r}_0, \hat{r}_1, \hat{r}_2, \ldots\}$ as the transform, $X = \{m_0, m_1, m_2, \ldots\}$ as the moment matrix for the original system's state variable vector x, $\hat{X} = \{\hat{m}_0, \hat{m}_1, \hat{m}_2, \ldots\}$ as the moment matrix for $\hat{x}$ which is the state variable vector for the reduced order model and $\tilde{X} = \{\tilde{m}_0, \tilde{m}_1, \tilde{m}_2, \ldots\}$ as the moment matrix for $\tilde{x}$ which is the approximation of the original system variable vector via our Krylov subspace based method. Then

$$m_i = \tilde{m}_i \tag{14}$$

4.Practical considerations

4.1. Scaling

The recursive computation of the source moments can be numerically unstable. The reason is that the delay term τ_i in delayed ramp functions such as $r_i \exp(-\tau_i s)$ can be large (much greater than 1). This may cause $r_i \exp(-\tau_i s)$ to decrease too fast to ensure the accuracy of the moment computation. Recall that $r_i \exp(-\tau_i s)$ can be expanded in terms of Taylor series expansion as

$$r_i \exp(-\tau_i s) = \sum_{l=1}^{\infty} (-1)^{l-1} \frac{r_i \tau_i^{l-1}}{(l-1)!}$$

This series expansion is generally valid only for values of τ_i less than 1. For values of τ_i greater than 1, the numerical values of the moments can quickly explode. Therefore, it is necessary to scale the temporal variable to avoid this numerical problem. In order to ensure that all the delays are of values less than 1, the scaling factor for time can be chosen to be the total simulation time or the largest end point of all the delayed functions. Another issue to be considered is the scaling of the ramp function values. Note that when only

the time is scaled to be smaller, the values of the slopes r_i are made bigger. This effect makes the waveforms sharper (more high frequency content) and may require more moments to accurately capture the characteristics of the waveforms. This in turn will impact the efficiency of the method. Consequently, it is necessary to scale both the time and the value of the ramp function. In practice, we use two scaling factors: timing scaling factor and value scaling factor. Assume that the timing scaling factor is β_1 and the value scaling factor is β_2. Then the slopes of the input ramps can be written as

$$r_i = \frac{\beta_1}{\beta_2}\bar{r}_i$$

where $\bar{r}_i$ is the scaled slope which will be used in subsequent computation. Let $\tau_i s = \bar{\tau}_i \bar{s}$ where $\bar{s} = \frac{1}{\beta_1} s$ and $\bar{\tau}_i = \beta_1 \tau_i$. We may choose β_1 as $1/(total\text{-}simulation\text{-}time)$ and $\beta_1 = \beta_2$,thus $r_i = \bar{r}_i$.

4.2. Error control

In this subsection, a simple error criterion is derived to estimate the accuracy of the reduced order system. The error is based on the residual of the moments at a frequency which has not been used as an expansion point. In other words, let

$$V = \left\{ \hat{r}_0^{(\sigma_0)}, \hat{r}_1^{(\sigma_0)}, \ldots, \hat{r}_{k(\sigma_0)}^{(\sigma_0)} \right\}$$

be the transformation matrix obtained at the expansion point σ_0. The issue here is how to estimate the error incurred by using the reduced order model obtained by this transformation matrix. First the reduced order system can be described by

$$(\hat{G} + s\hat{C})z(s) = V^T B u(s) \qquad (15)$$

where $\hat{G} = V^T G V$, $\hat{C} = V^T C V$, and $x(s) = Vz(s)$. The goal now is to estimate the error at another frequency point, e.g. σ_1. Let $m_{Bu}^{(0)} = Bu_0^{(\sigma_1)}$ be the zeroth moment of the source at σ_1. This moment can be computed using the procedure described in subsection 2.2. Then the moment of the reduced system at σ_1 can be computed as

$$z_0^{(\sigma_1)} = (\hat{G} + \sigma_1 \hat{C})^{-1} V^T m_{Bu}^{(0)}$$

and the approximate moment of the response of the original system can be computed as $\tilde{x}_0^{(\sigma_1)} = V z_0^{(\sigma_1)}$. This approximate moment can be compared with exact moment

$$x_0^{(\sigma_1)} = (G + \sigma_1 C)^{-1} m_{Bu}^{(0)}$$

to give an estimate of the error. However, this computation requires another expensive LU factorization of $G + \sigma_1 C$. As an alternative, the approximate source moment is computed as

$$\tilde{m}_{Bu}^{(0)} = (G + \sigma_0 C)\tilde{x}_0^{(\sigma_1)}$$

and the residual error can be defined as $\left\| m_{Bu}^{(0)} - \tilde{m}_{Bu}^{(0)} \right\|$. This procedure provides a simple and less expensive way to estimate the error. Note also that the high oder moments are ignored in the computation of the error. The rationale is that with proper scaling as discussed in subsection 3.1 the zeroth moment of the sources will contain most of the information at each expansion point.

5.Experimental results

The extended Krylov subspace (EKS) method presented in this paper has been implemented in a prototype power grid simulator. A standard sparse matrix package is used for the solution of the linear system of equations. The power grid is modeled as linear RC(L) networks with independent current sources representing the switching activities of the circuit macros. The accuracy and efficiency of the power grid simulator is compared with circuit simulation. The simulation results for a number of test circuits of various sizes are summarized in Table 1. As shown in this table, the speed up of the EKS method over circuit simulation is about one order of magnitude (~ 10X-25X). This speed up is less impressive than the speed up for conventional model reduction techniques (up to ~1000X). However, the difference is that in conventional model reduction, one is interested in the responses at a small number of ports (usually less than 20). For example, one is interested in the pin to pin delay in timing analysis or clock tree analysis and not in the responses at the intermediate nodes. Therefore, conventional model reduction can abstract away a large number of internal nodes and only need to compute the responses at only a small number of ports. However, in power grid analysis, one is generally interested in the transient responses at all the nodes of the power grid. The number of responses can be reduced to a subset of nodes for monitoring the supply noise. In any case, this number of nodes is usually much larger than the number of ports in conventional model reduction. As a result, there is no luxury in abstracting away a large number of variables. In oder to demonstrate the accuracy of the EKS method, a number of representative waveforms from some of the test circuits above are shown. For example, the voltage at a ground node of circuit 1 is shown in Fig. 3.

Table 1:

name	#of nodes	#of sources	Cir sim	ESK
Circuit 1	16	9	0.13s	0.01s
Circuit 2	100	81	0.31s	0.02s
Circuit 3	2000	200	362s	21s
Circuit 4	10400	824	952s	67s
Circuit 5	20309	2475	3678s	140s
Circuit 6	80200	4102	N/A	420s

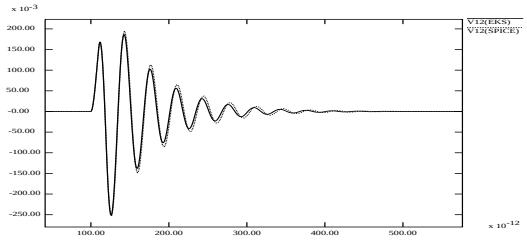


Fig. 3 *The voltage response at a ground node of circuit 1*

This circuit is the model of a section of a power grid over one circuit macro. The graph shows the inductive effect of the package. Another example waveform for circuit#3 is shown in Fig.4.

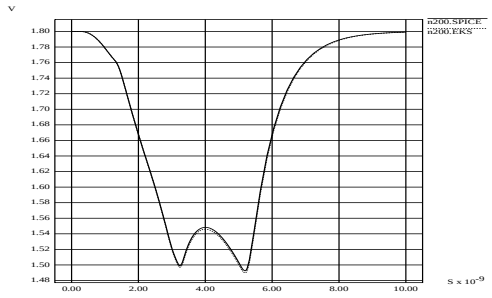
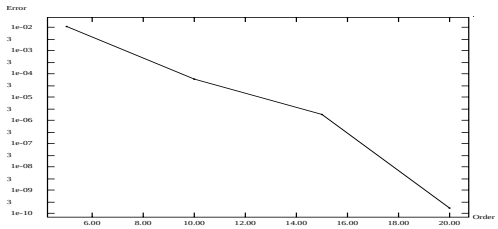
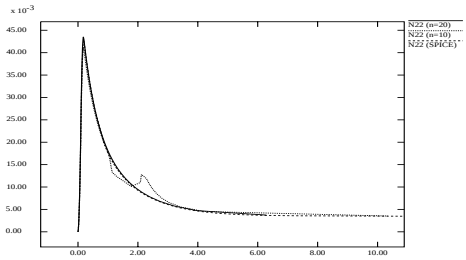


Fig. 4 *Waveform at node 200 (The voltage of the power is 1.8 v)*

This circuit represents a larger section of the power grid. The waveform shows the effect of significant switching on the power rail some time after the initial power up. .



(a)



(b)

Fig. 5 *(a) Absolute errors of various low order approximations of 100 RC chain; (b)The time domain waveform comparison for order 10 and order 20 versus circuit simulation*

In order to illustrate the effect of error control, a 100 RC chain with a number of current sources inserted along the chain. This is not a grid topology of typical power grid. It is used to check the neccessity of error control. For this particular circuit, it requires an order 20 to give accurate results. The absolute errors at different orders of approximation and the time domain waveform comparison for order 10 and 20 versus circuit simulation result are shown in Fig.5. This example shows that the efficiency of the method is somewhat dependent on the topology as well as the values of the elements in the linear circuit model of the power grid.

6.Conclusion and future work

An extended Krylov subspace (EKS) method for general reduced order analysis of linear circuits with a large number of independent sources has been presented. This method includes an incremental orthogonalization procedure to compute an orthogonal basis of the response moment subspace. It was proved that the reduced system preserves passivity as well as the moments of the responses. The EKS method allows for an arbitrary number of sources with general PWL representation and avoids the numerical problem of explicit moment matching. Experimental results show that the proposed method achieves a moderate speed up over circuit simulation with good accuracy for global low frequency variations of the power distribution network. For high frequency variations, high order approximations may be required. Note also that the achieved speed up is less impressive than the speed up for conventional model reduction techniques. The reason is that in power grid analysis, one is usually interested in computing the responses at all the nodes or at least a large number of nodes in the power distribution network. As a result, there is no luxury in abstracting away a large number of internal nodes in order to compute efficiently the responses at a small number of ports as done in conventional model reduction.

For future work, a number of improvements can be investigated. The computational cost of the EKS method is dominated by the initial LU factorization of the MNA matrix and the orthogonalization process. A more efficient linear solution method such as path tracing can be employed. Moreover, currently a general MNA formulation is used to model the power grid. It is possible to use the nodal formulation to produce a positive definite matrix for the resulting linear system with some loss of accuracy. This formulation allows the more efficient Cholesky decomposition for the solution of the linear system. For orthogonalization, a full back orthogonalization with respect to all previous vectors is currently implemented. A Lanczos type method which allows partial back orthogonalization is under investigation.

Acknowledgments

The authors would like to thank Eli Chiprout and Sani Nassif for their helpful discussion during the course of this work.

References

[1] H. H. Chen and J. S. Neely, “Interconnect and circuit modeling techniques for full chip power supply noise analysis,” *IEEE Trans. Compon. Packag. Manuf. Technol. B, Adv. Packag.*, Vol.21, pp. 209-15, Aug. 1998.

[2] L. Pillage and R. A. Rohrer, “Asymptotic Waveform Evaluation for timing analysis”, *IEEE Trans. on Computer-aided design*, vol. 9, pp. 352-366, April, 1990.

[3] P. Feldmann and R. W. Freund, “Efficient Linear Circuit Analysis by Padé Approximation via the Lanczos Process,” *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 639-649, May 1995.

[4] A. Odabasioglu, M. Celik, and L. T. Pileggi, “PRIMA: passive reduced order interconnect macromodeling algorithm,” *IEEE Trans. Computer-Aided Design*, vol. 17, pp. 645-654, August 1998.

[5] A. Dharchoudhury, R. Panda, D. Blaauw, and R. Vaidyanathan, “Design and Analysis of Power Distribution Networks in PowerPc Microprocessors,” *Proc. DAC*, pp. 738-743, 1998.

[6] J. C. Shah, A. A. Younis, S. S. Sapatnekar, and M. M. Hassoun, “An algorithm for simulating power/ground networks using Pade approximants and its symbolic implementation ,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, Vol.45, pp. 1372-82, Oct. 1998

[7] E. Chiprout and T. Nguyen, "Power analysis of large interconnect grids withmultiple sources using model reduction", *European Conference on Circuit Theory and Design*, Stressa, Italy, Sept. 1999.

[8] J.Wang, T, Nuyen, “Extended Krylov Subspace Mehtod for power grid analysis”,ERL Memorandum 2000.