

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/256846537>

Power Grid Simulation using Matrix Exponential Method with Rational Krylov Subspaces

Article · September 2013

DOI: 10.1109/ASICON.2013.6811937 · Source: arXiv

CITATIONS

12

READS

72

3 authors, including:



[Hao Zhuang](#)

Google Inc.

34 PUBLICATIONS 275 CITATIONS

SEE PROFILE

Power Grid Simulation using Matrix Exponential Method with Rational Krylov Subspaces

Hao Zhuang, Shih-Hung Weng, and Chung-Kuan Cheng

Department of Computer Science and Engineering, University of California, San Diego, CA, 92093, USA

hao.zhuang@cs.ucsd.edu, s2weng@ucsd.edu, ckcheng@ucsd.edu

Abstract—One well-adopted power grid simulation methodology is to factorize matrix once and perform only backward/forward substitution with a deliberately chosen step size along the simulation. Since the required simulation time is usually long for the power grid design, the costly factorization is amortized. However, such fixed step size cannot exploit larger step size for the low frequency response in the power grid to speedup the simulation. In this work, we utilize the matrix exponential method with the rational Krylov subspace approximation to enable adaptive step size in the power grid simulation. The kernel operation in our method only demands one factorization and backward/forward substitutions. Moreover, the rational Krylov subspace approximation can relax the stiffness constraint of the previous works [12][13]. The cheap computation of adaptivity in our method could exploit the long low-frequency response in a power grid and significantly accelerate the simulation. The experimental results show that our method achieves up to 18X speedup over the trapezoidal method with fixed step size.

I. INTRODUCTION

Power grid simulation is a very essential and computational heavy tasks during VLSI design. Given current stimulus and the power grid structure, designers could verify and predict the worst-case voltage noise through the simulation before signing off their design. However, with the huge size of modern design, power grid simulation is a time-consuming process. Moreover, manifesting effects from the package and the board would require longer simulation time, e.g., up to few μs , which worsens the performance of the power grid simulation. Therefore, an efficient power grid simulation is always a demand from industry.

Conventionally, the power grid simulation is based on the trapezoidal method where the major computation is to solve a linear system by either iterative approaches or direct methods. The iterative methods usually suffer from the convergence problem because of the ill-conditioned matrix from the power grid design. On the other hand, the direct methods, i.e., Cholesky or LU factorizations, are more general for solving a linear system. Despite the huge memory demanding and computational effort, with a carefully chosen step size, the power grid simulation could perform only one factorization at the beginning while the rest of operations are just backward/forward substitutions. Since a power grid design usually includes board and package models, a long simulation time is required to manifest the low-frequency response. Hence, the cost of expensive factorization can be amortized by many faster backward/forward substitutions. Such general factorization and fixed step size strategy[14][15][16][17] is widely adopted in industry.

The matrix exponential method (MEXP) for the circuit simulation has better accuracy and adaptivity because of the analytical formulation and the scaling invariant Krylov subspace approximation[12][13]. Unlike the fixed step size strategy, MEXP could dynamically adjust the step size to exploit the low-frequency response of the power grid without expensive computation. However, the step size in MEXP is limited by the stiffness of circuit. This constraint would drag the overall performance of MEXP for the power grid simulation.

In this work, we tailor MEXP using *rational Krylov subspace* for

the power grid simulation with adaptive time stepping. The rational Krylov subspace uses $(\mathbf{I} - \gamma\mathbf{A})^{-1}$ as the basis instead of \mathbf{A} used in the conventional Krylov subspaces, where \mathbf{I} is an identity matrix and γ is a predefined parameter. The rational basis limits the spectrum of a circuit system, and emphasizes small magnitude eigenvalues, which are important to exponential function, so that the exponential of a matrix can be accurately approximated. As a result, MEXP with rational Krylov subspace can enjoy benefits of the adaptivity and the accuracy of MEXP. Even though the rational Krylov subspace still needs to solve a linear system as the trapezoidal method does, MEXP can factorize the matrix only once and then constructs the rest of rational Krylov subspaces by backward/forward substitutions. Therefore, MEXP can utilize its capability of adaptivity to accelerate the simulation with the same kernel operations as the fixed step size strategy. Overall, our MEXP enables adaptive time stepping for the power grid simulation with only one LU factorization, and allows scaling large step size without compromising the accuracy. The experimental results demonstrate the effectiveness of MEXP with adaptive step size. The industrial power grid designs can be accelerated 17X on average compared to the trapezoidal method.

The rest of paper is organized as follows. Section II presents the background of the power grid simulation and MEXP. Sections III and IV show the theoretical foundation of the rational Krylov subspace and our adaptive step scheme for the power grid simulation, respectively. Section V presents experimental results of several industrial power grid designs. Finally, Section VI concludes this paper.

II. PRELIMINARY

A. Power Grid Formulation

A power grid can be formulated as a system of differential equations via modified nodal analysis as below:

$$\mathbf{C}\dot{\mathbf{x}}(t) = -\mathbf{G}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad (1)$$

where matrix \mathbf{C} describes the capacitance and inductance, matrix \mathbf{G} represents the resistance and the incidence between voltages and currents, and matrix \mathbf{B} indicates locations of the independent sources. Vector $\mathbf{x}(t)$ describes the nodal voltages and branch currents at time t , and vector $\mathbf{u}(t)$ represents the corresponding supply voltage and current sources associated to different blocks. In this work, we assume those input sources are given and in the format of *piece-wise linear*.

B. Matrix Exponential Method

MEXP[12][13] is based on the analytical solution of (1). With initial solution from the DC analysis via direct[3] or iterative approaches[7], the equation of MEXP from t to $t+h$ can be expressed as

$$\mathbf{x}(t+h) = e^{\mathbf{A}h}\mathbf{x}(t) + \int_0^h e^{\mathbf{A}(h-\tau)}\mathbf{b}(t+\tau)d\tau. \quad (2)$$

where $\mathbf{A} = -\mathbf{C}^{-1}\mathbf{G}$, and $\mathbf{b}(t) = \mathbf{C}^{-1}\mathbf{B}\mathbf{u}(t)$, when \mathbf{C} is not singular. Assuming that input $\mathbf{u}(t)$ is piece-wise linear (PWL), we integrate the last term in (2) analytically, turning the solution into the sum of three terms associated with matrix exponential operators,

$$\begin{aligned} \mathbf{x}(t+h) &= e^{\mathbf{A}h}\mathbf{x}(t) \\ &+ (e^{\mathbf{A}h} - \mathbf{I})\mathbf{A}^{-1}\mathbf{b}(t) \\ &+ (e^{\mathbf{A}h} - (\mathbf{A}h + \mathbf{I}))\mathbf{A}^{-2}\frac{\mathbf{b}(t+h) - \mathbf{b}(t)}{h}. \end{aligned} \quad (3)$$

Eqn. (3) has three matrix exponential terms, which are generally referred as φ -functions of the zero, first and second order [6]. It has been shown in [1, Theorem 2.1] that one can obtain the sum of them in one shot by computing the exponential of a slightly larger $(n+p) \times (n+p)$ matrix, where n is the dimension of \mathbf{A} , and p is the order of the φ -functions ($p = 2$ in (3)). Thus, (3) can be rewritten into

$$\mathbf{x}(t+h) = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \end{bmatrix} e^{\tilde{\mathbf{A}}h} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{e}_2 \end{bmatrix}, \quad (4)$$

with

$$\begin{aligned} \tilde{\mathbf{A}} &= \begin{bmatrix} \mathbf{A} & \mathbf{W} \\ \mathbf{0} & \mathbf{J} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \frac{\mathbf{b}(t+h) - \mathbf{b}(t)}{h} & \mathbf{b}(t) \end{bmatrix} \\ \mathbf{J} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \quad (5)$$

To keep the notations simple, we use \mathbf{v} to represent $[\mathbf{x}(t) \ \mathbf{e}_2]^\top$ in the rest of paper, respectively. Note that the kernel computation of MEXP is to derive the exponential of a matrix, i.e., $e^{\mathbf{A}\mathbf{v}}$, which is approximated by the Krylov subspace method in works [13][12]. The Krylov subspace method has better scalability mainly from its sparse matrix-vector product centric computation. However, such approximation is only better for those eigenvalues with small magnitude, which means the maximum step size of MEXP in a stiff circuit has to be constrained in order to maintain the accuracy of approximation. In the following section, we will present how the rational basis could relax the stiffness constraint.

III. MEXP WITH RATIONAL KRYLOV SUBSPACE

In [13][12], Eqn. (4) is calculated via the Krylov subspace method using Arnoldi process. The subspace is defined as

$$\mathbf{K}_m(\tilde{\mathbf{A}}, \mathbf{v}) = \text{span}\{\mathbf{v}, \tilde{\mathbf{A}}\mathbf{v}, \dots, \tilde{\mathbf{A}}^{m-1}\mathbf{v}\}, \quad (6)$$

where \mathbf{v} is an initial vector. The Arnoldi process approximates the eigenvalues with large magnitude well. But when handling a stiff circuit system, the formed matrix usually contains many eigenvalues with small magnitude. Besides, $e^{\tilde{\mathbf{A}}h}$ is mostly determined by the eigenvalues with smallest magnitudes and their corresponding invariant subspaces. In this scenario, due to the existence of eigenvalues with large magnitude in $\tilde{\mathbf{A}}$, the Arnoldi process for Eqn. (6) requires large m to capture the important eigenvalues (small magnitudes) and invariant spaces for exponential operator. Therefore, the time steps in MEXP has to be small enough to capture the important eigenvalues. This suggests us transforming the spectrum to intensify those eigenvalues with small magnitudes and corresponding invariant subspaces. We make such transformation based on the idea of *rational Krylov subspace method*[5][11]. The details are presented in the following subsections.

A. Rational Krylov Subspaces Approximation of MEXP

For the purpose of finding the eigenvalues with smallest magnitude first, we use a preconditioner $(\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-1}$, instead of using $\tilde{\mathbf{A}}$

directly. It is known as the rational Krylov subspace[5][11]. The formula for the rational Krylov subspace is

$$\mathbf{K}_m((\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-1}, \mathbf{v}) = \text{span}\{\mathbf{v}, (\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-1}\mathbf{v}, \dots, (\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-(m-1)}\mathbf{v}\}, \quad (7)$$

where γ is a predefined parameter. The Arnoldi process constructs \mathbf{V}_{m+1} and $\mathbf{H}_{m+1,m}$, and the relationship is given by

$$(\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-1}\mathbf{V}_m = \mathbf{V}_m\mathbf{H}_{m,m} + h_{m+1,m}\mathbf{v}_{m+1}\mathbf{e}_m^\top, \quad (8)$$

where \mathbf{e}_m is the m -th unit vector. Matrix $\mathbf{H}_{m,m}$ is the first $m \times m$ square matrix of an upper Hessenberg matrix of $\mathbf{H}_{m+1,m}$, and $h_{m+1,m}$ is its last entry. \mathbf{V}_m consists of $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m]$, and \mathbf{v}_{m+1} is its last vector. After re-arranging (8) and given a time step h , the matrix exponential $e^{\tilde{\mathbf{A}}h}\mathbf{v}$ can be calculated as

$$\begin{aligned} e^{\tilde{\mathbf{A}}h}\mathbf{v} &\approx \mathbf{V}_m\mathbf{V}_m^\top e^{\tilde{\mathbf{A}}h}\mathbf{v} = \|\mathbf{v}\|_2 \mathbf{V}_m\mathbf{V}_m^\top e^{\tilde{\mathbf{A}}h}\mathbf{v}_1 \\ &= \|\mathbf{v}\|_2 \mathbf{V}_m e^{\alpha\tilde{\mathbf{H}}_{m,m}}\mathbf{e}_1, \end{aligned} \quad (9)$$

where $\tilde{\mathbf{H}}_{m,m} = \mathbf{I} - \mathbf{H}_{m,m}^{-1}$, $\alpha = \frac{h}{\gamma}$ is the adjustable parameters for control of adaptive time step size in Section IV. Note that in practice, instead of computing $(\mathbf{I} - \gamma\tilde{\mathbf{A}})^{-1}$ directly, we only need to solve $(\mathbf{C} + \gamma\mathbf{G})^{-1}\mathbf{C}\mathbf{v}$, which can be achieved by one LU factorization at beginning. Then the construction of the following subspaces is by backward/forward substitutions.

This strategy is also presented in [5][11]. Intuitively, the “shift-and-invert” operation would intensify the eigenvalues with small magnitudes and minify the eigenvalues with large magnitudes. By doing so, the Arnoldi process could capture those eigenvalues important to the exponential operator, which originally cannot be manifested with small m in the conventional Krylov subspace. We would like to point out that the error bound for Eqn. (9) does not longer depend on $\|\tilde{\mathbf{A}}h\|$ as [13]. It is only the first (smallest magnitude) eigenvalue of $\tilde{\mathbf{A}}$. We observe that large α provides less absolute error under the same dimension m . An intuitive explanation is also given by [11], the larger α combined with exponential operators, the relatively smaller portion of the eigenvalues with smallest magnitude determine the final vector. Within the assumption of piecewise linear in Eqn. (3), our method can step forward as much as possible to accelerate simulation, and still maintain the high accuracy. The sacrifice resides in the small time step when more eigenvalues determine the final vector. So we should choose a appropriate parameter γ or increase the order m to balance the accuracy and efficiency. Even though the increasing m results more backward/forward substitutions, the m is still quite small in the power grid simulation. Therefore, it does not degrade our method too much.

The formula of posterior error estimation is required for controlling adaptive step size. We use the formula derived from [11],

$$\text{err}(m, \alpha) = \frac{\|\mathbf{v}\|_2}{\gamma} h_{m+1,m} \left| (\mathbf{I} - \gamma\tilde{\mathbf{A}})\mathbf{v}_{m+1}\mathbf{e}_m^\top \mathbf{H}_{m,m}^{-1} e^{\alpha\tilde{\mathbf{H}}_{m,m}}\mathbf{e}_1 \right| \quad (10)$$

The formula provides a good approximation for the error trend with respect to m and α in our numerical experiment.

B. Block LU factorization

In practical numerical implementation, in order to avoid direct inversion of \mathbf{C} to form \mathbf{A} in Eqn. (7), the equation $(\mathbf{C} + \gamma\mathbf{G})^{-1}\mathbf{C}$ is used. Correspondingly, for Eqn. (4), we use the equations

$$(\tilde{\mathbf{C}} - \gamma\tilde{\mathbf{G}})^{-1}\tilde{\mathbf{C}} \quad (11)$$

$$\text{where } \tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tilde{\mathbf{G}} = \begin{bmatrix} -\mathbf{G} & \tilde{\mathbf{W}} \\ \mathbf{0} & \mathbf{J} \end{bmatrix}, \text{ and } \tilde{\mathbf{W}} = \begin{bmatrix} \frac{\mathbf{B}\mathbf{u}(t+h) - \mathbf{B}\mathbf{u}(t)}{h} & \mathbf{B}\mathbf{u}(t) \end{bmatrix}$$

The Arnoldi process based on Eqn. (11) actually only requires to solve \mathbf{v}_{k+1} with \mathbf{v}_k . The linear system is expressed as

$$(\tilde{\mathbf{C}} - \gamma\tilde{\mathbf{G}})\mathbf{v}_{k+1} = \tilde{\mathbf{C}}\mathbf{v}_k, \quad (12)$$

where \mathbf{v}_k and \mathbf{v}_{k+1} are k -th and $(k+1)$ -th basis in the rational Krylov subspace. If $\tilde{\mathbf{W}}$ changes with inputs during the simulation, the Arnoldi process has to factorize a matrix every time step. However, it is obvious that the majority of $\tilde{\mathbf{G}}$ stay the same for this linear system. To take advantage of this property, a block LU factorization is devised here to avoid redundant calculation. The goal is to obtain the lower triangular \mathbf{L} and the upper triangular \mathbf{U} matrices:

$$\tilde{\mathbf{C}} - \gamma\tilde{\mathbf{G}} = \mathbf{L}\mathbf{U}. \quad (13)$$

At the beginning of simulation, after LU factorization of $\mathbf{C} + \gamma\mathbf{G} = \mathbf{L}_{sub}\mathbf{U}_{sub}$, we obtain the lower triangular sub-matrix \mathbf{L}_{sub} , and upper triangular sub-matrix \mathbf{U}_{sub} . Then Eqn. (13) only needs updating via

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{sub} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{U}_{sub} & -\gamma\mathbf{L}_{sub}^{-1}\tilde{\mathbf{W}} \\ \mathbf{0} & \mathbf{I}_J \end{bmatrix}, \quad (14)$$

where $\mathbf{I}_J = \mathbf{I} - \gamma\mathbf{J}$ is an upper triangular matrix. Assume we have \mathbf{v}_k , the following equations further reduce operation \mathbf{L}_{sub}^{-1} and construct vector \mathbf{v}_{k+1} : $\mathbf{z}_1 = [\mathbf{C}_1, \mathbf{0}]\mathbf{v}_k$, $\mathbf{z}_2 = [\mathbf{0}, \mathbf{I}]\mathbf{v}_k$; $\mathbf{y}_2 = \mathbf{I}_J^{-1}\mathbf{z}_2$, $\mathbf{L}_{sub}\mathbf{U}_{sub}\mathbf{y}_1 = \mathbf{z}_1 + \gamma\tilde{\mathbf{W}}\mathbf{y}_2$. Then, we obtain $\mathbf{v}_{k+1} = [\mathbf{y}_1, \mathbf{y}_2]^T$. By doing this, it only needs one LU factorization at the beginning of simulation, and with cheap updates for the \mathbf{L} and \mathbf{U} at each time step during transient simulation.

IV. ADAPTIVE TIME STEP CONTROL

The proposed MEXP can significantly benefit from the adaptive time stepping because the rational Krylov subspace approximation relaxes the stiffness constraint as well as preserves the scaling invariant property. As a result, MEXP can effortlessly adjust the step size to different scale, during the simulation. Such adaptivity is particularly helpful in the power grid where the voltage noise includes the high- to low-frequency responses from die, package and board.

Our adaptive step scheme is to step forward as much as possible so that MEXP can quickly finish the simulation. With the insight from Eqn. (9), MEXP can adjust α to calculate results of required step sizes with only one Arnoldi process. However, even though the rational Krylov subspace could scale robustly, the step size in MEXP is restrained from input sources. As shown in Eqn. (3), MEXP has to guarantee constant slope during a stepping, and hence the maximum allowed step size h at every time instant is limited. Our scheme will first determine h from inputs at time t and construct the rational Krylov subspace from $\mathbf{x}(t)$. Then, \mathbf{x} within interval $[t, t+h]$ are calculated through the step size scaling.

Algorithm 1 shows MEXP with adaptive step control. In order to comply with the required accuracy during the simulation, the allowed error $err(m, \alpha)$ at certain time instant t is defined as $err \leq \frac{E_{Tol}}{T}h$ where E_{Tol} is the error tolerance in the whole simulation process, T is the simulation time span, h is the step size at time t , and err is the posterior error of MEXP from Eqn. (10). Hence, when we construct the rational Krylov subspace, we will increase m until the $err(m, \alpha)$ satisfies the error tolerance. The complexity of MEXP with adaptive time stepping is mainly determined by the total number of required backward/forward substitutions during the simulation process. The number of total substitution operations is $\sum_{i=0}^N m_i$ where N is total time steps, and m_i is required dimension of the rational Krylov subspace at time step i . Compared to the trapezoidal method where the number of substitution operations depends only on the fixed step size, MEXP could use less substitution operations as long as the maximum allowed step size h is much larger than the fixed step size.

Algorithm 1: MEXP with Adaptive Step Control

Input: \mathbf{C} , \mathbf{G} , \mathbf{B} , $\mathbf{u}(t)$, τ , error tolerance E_{Tol} and simulation time T

Output: $\mathbf{x}(t)$

```

1  $t = 0$ ;  $\mathbf{x}(0) = \text{DC\_analysis}$ ;
2  $[\mathbf{L}_{sub}, \mathbf{U}_{sub}] = \text{LU}(\mathbf{C} + \gamma\mathbf{G})$ ;
3 while  $t \leq T$  do
4   Compute maximum allowed step size  $h$  from  $\mathbf{u}(t)$  and
    $\alpha = \frac{h}{\gamma}$ ;
5   Construct  $\mathbf{H}_{m,m}$ ,  $\mathbf{V}_{m,m}$ ,  $err$  by Arnoldi process and (10)
   until  $err(m, \alpha) \leq \frac{E_{Tol}}{T}h$ ;
6   Compute  $\mathbf{x}(t+h)$  by (9);
7    $t = t + h$ ;
8 end
```

Our experiments in the following section demonstrates it is usually the case for the power grid simulation.

V. EXPERIMENTAL RESULTS

In this section, we compare performance of the power grid simulation by MEXP and the trapezoidal method (TR). MEXP with adaptive step size control follows Algorithm 1. We predefine γ e.g. 10^{-10} here, and restrict the maximum allowed step size within $1ns$ to have enough time instants to plot the figure. It is possible to have more fine-grain time instants, e.g., $10ps$, with only negligible cost by adjusting α in Eqn. (9). TR is in fixed step size h in order to minimize the cost of LU factorization. Both methods only perform factorization once for transient simulation, and rest of operations is mainly backward/forward substitution. We implement both methods in MATLAB and use UMFPAK package for LU factorization. Note that even though previous works[2][9] show that using iterative approach in TR could also achieve adaptive step control, long simulation time span in power grid designs make direct method with fixed step size more desirable[14][15][17]. The experiments are performed on a Linux workstation with an Intel Core i7-920 2.67GHz CPU and 12GB memory. The power grid consists of four metal layers: M1, M3, M6 and RDL. The physical parameters of each metal layer is listed in Table I. The package is modeled as an RL series at each C4 bump, and the board is modeled as a lumped RLC network. The specification of each PDN design is listed in Table II where the size of each design ranges from 45.7K to 7.40M.

TABLE I

WIDTHS AND PITCHES OF METAL LAYERS IN THE PDN DESIGN(μm).

M1		M3		M6		RDL	
pitch	width	pitch	width	pitch	width	pitch	width
2.5	0.2	8.5	0.25	30	4	400	30

TABLE II
SPECIFICATIONS OF PDN DESIGNS

Design	Area (mm^2)	#R	#C	#L	#Nodes
D1	0.35 ²	23221	15193	15193	45.7K
D2	1.40 ²	348582	228952	228952	688K
D3	2.80 ²	1468863	965540	965540	2.90M
D4	5.00 ²	3748974	2467400	2464819	7.40M

In order to characterize a PDN design, designers can rely on the simulation result of impulse response of the PDN design. Many previous works[4][10] have proposed different PDN analysis based on the impulse response. The nature of impulse response of the PDN design, which contains low-, mid- and high-frequency components,

can significantly enjoy the adaptive step size in MEXP. We would also like to mention that the impulse response based analysis is not only for the PDN design, but also for worst-case eye opening analysis in the high speed interconnect [8].

The impulse response can be derived from the simulation result of a step input from 0V to 1V with a small transition time. Hence, we inject a step input to each PDN design and compare the performance of MEXP and TR. The transition time of the step input and the simulation time span is $10ps$ and $1\mu s$ for observing both high- and low-frequency responses. Table III shows the simulation runtime of MEXP and TR where the fixed step size is set as $10ps$ to comply with the transition time. In the table, “DC”, “LU” and “Time” indicate the runtime for DC analysis, LU factorization and the overall simulation, respectively. DC analysis is also via the LU factorization. We can also adopt other techniques[14][15][16] to improve the performance of DC analysis for both methods. It is noted that these cases are very stiff and with singular matrix C . We do not use the method[12][13] on the benchmarks, because that it cannot handle the singular C in these industrial PDN design without regularization. It is worth pointing out that even after regularization[13], the stiffness still causes large m series for matrix exponential evaluation. For example, we construct a simple RC mesh network with 2500 nodes. The extreme values of this circuit are $C_{min} = 5.04 \times 10^{-19}$, $C_{max} = 1.00 \times 10^{-15}$, $G_{min} = 1.09 \times 10^{-2}$, and $G_{max} = 1.00 \times 10^2$. The corresponding maximum eigenvalue of $-C^{-1}G$ is -1.88×10^9 and minimum eigenvalue is -3.98×10^{17} . The stiffness is $\frac{Re(\lambda_{min})}{Re(\lambda_{max})} = 2.12 \times 10^8$. During simulation of $1ns$ time span, with a fixed step size $10ps$, MEXP based on [12][13] costs average and peak dimensions of Krylov subspace $m_{avg} = 115$, and $m_{peak}=264$, respectively. Our MEXP uses rational Krylov subspaces, which only need $m_{avg}=3.11$, $m_{peak}=10$ and lead to 224X speedup in total runtime.

In these test cases, our MEXP has significant speedup over TR because it can adaptively exploit much large step size to simulate the design whereas TR can only march with $10ps$ time step for whole $1\mu s$ time span. The average speedup is 17X. Fig. 1 shows the simulation result of design D1 at a node on M1. As we can see, the result by MEXP and TR are very close to the result of HSPICE, which is as our reference result here. The errors of MEXP and TR to HSPICE are 7.33×10^{-4} and 7.47×10^{-4} . This figure also demonstrates that a PDN design has low-, mid- and high-frequency response so that long simulation time span is necessary, meanwhile, small time steps are required during the $20ns$ in the beginning.

TABLE III
SIMULATION RUNTIME OF PDN DESIGNS

Design	DC(s)	TR ($h = 10ps$)		Our MEXP ($\gamma = 10^{-10}$)		
		LU(s)	Total	LU(s)	Total	Speedup
D1	0.710	0.670	44.9m	0.680	2.86m	15.7
D2	12.2	15.6	15.4h	15.5	54.6m	16.9
D3	69.6	91.6	76.9h	93.3	4.30h	17.9
D4	219	294	204h	299	11.3h	18.1

VI. CONCLUSION

For large scale power grid simulation, we propose an MEXP framework using two methods rational Krylov subspace approximation and adaptive time stepping technique. The former method can relax stiffness constraint of [12][13]. The later one helps adaptively exploit low-, mid-, and high-frequency property in simulation of industrial PDN designs. In the time-consuming impulse response simulation, the proposed method achieve more than 15 times speedup on average over the widely-adopted fixed-step trapezoidal method.

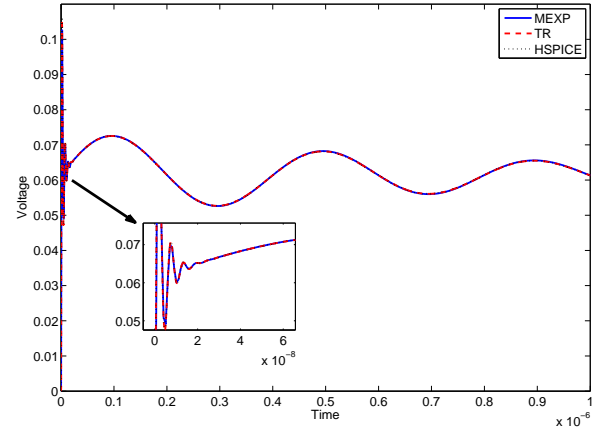


Fig. 1. Result of D1

VII. ACKNOWLEDGEMENTS

This work was supported by NSF CCF-1017864.

REFERENCES

- [1] A. H. Al-Mohy and N. J. Higham. Computing the action of the matrix exponential, with an application to exponential integrators. *SIAM Journal on Scientific Computing*, 33(2):488–511, 2011.
- [2] T.-H. Chen and C. C.-P. Chen. Efficient large-scale power grid analysis based on preconditioned krylov-subspace iterative methods. In *Proc. Design Automation Conference*, pages 559–562, 2001.
- [3] T. A. Davis. *Direct Method for Sparse Linear Systems*. SIAM, 2006.
- [4] P. Du, X. Hu, S.-H. Weng, A. Shayan, X. Chen, A. Ege Engin, and C. K. Cheng. Worst-case noise prediction with non-zero current transition times for early power distribution system verification. In *Intl. Symposium on Quality Electronic Design*, pages 624–631, 2010.
- [5] I. Moret and P. Novati. Rd-rational approximations of the matrix exponential. *BIT Numerical Mathematics*, 44(3):595–615, 2004.
- [6] J. Niesen and W. M. Wright. A Krylov subspace algorithm for evaluating the φ -functions appearing in exponential integrators. *ACM Trans. Math. Software*, in press.
- [7] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, 2003.
- [8] R. Shi, W. Yu, Y. Zhu, C. K. Cheng, and E. S. Kuh. Efficient and accurate eye diagram prediction for high speed signaling. In *Proc. Intl. Conf. Computer-Aided Design*, pages 655–661, 2008.
- [9] H. Su, E. Acar, and S. R. Nassif. Power grid reduction based on algebraic multigrid principles. In *Proc. Design Automation Conference*, pages 109–112, 2003.
- [10] W. Tian, X.-T. Ling, and R.-W. Liu. Novel methods for circuit worst-case tolerance analysis. *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, 43(4):272–278, 1996.
- [11] J. van den Eshof and M. Hochbruck. Preconditioning lanczos approximations to the matrix exponential. *SIAM Journal on Scientific Computing*, 27(4):1438–1457, 2006.
- [12] S.-H. Weng, Q. Chen, and C. K. Cheng. Circuit simulation using matrix exponential method. In *Intl. Conf. ASIC*, pages 369–372, 2011.
- [13] S.-H. Weng, Q. Chen, and C. K. Cheng. Time-domain analysis of large-scale circuits by matrix exponential method with adaptive control. *IEEE Trans. Computer-Aided Design*, 31(8):1180–1193, 2012.
- [14] X. Xiong and J. Wang. Parallel forward and back substitution for efficient power grid simulation. In *Proc. Intl. Conf. Computer-Aided Design*, pages 660–663, 2012.
- [15] J. Yang, Z. Li, Y. Cai, and Q. Zhou. Powerrush: Efficient transient simulation for power grid analysis. In *Proc. Intl. Conf. Computer-Aided Design*, pages 653–659, 2012.
- [16] T. Yu, M. Wong, et al. Pgt-solver: an efficient solver for power grid transient analysis. In *Proc. Intl. Conf. Computer-Aided Design*, pages 647–652, 2012.
- [17] M. Zhao, R. V. Panda, S. S. Sapatnekar, and D. Blaauw. Hierarchical analysis of power distribution networks. *IEEE Trans. Computer-Aided Design*, 21(2):159–168, 2002.