A Variable Reduction Technique for the Analysis of Ultra Large-Scale Power Distribution Networks

Jong-Eun Koo, Kyung-Ho Lee, Young-Hoe Cheon, Joon-Ho Choi,
Moon-Hyun Yoo, Jeong-Taek Kong

CAE Team, Memory Division, Device Solution Network, Samsung Electronics Co., Ltd.
{jongeun.koo, kay.lee, yh.cheon, joonho.choi, moonyoo, jkong}@samsung.com

Abstract

In the trends of high density, high speed and low power consumption, the analysis of power distribution networks is rapidly becoming an essential step in the design of high performance ICs. Nevertheless, the analysis is a great challenge, due to the large size of the networks. In this paper, we propose a variable reduction technique for the analysis of large-scale power distribution networks. The basic procedure of the proposed technique is reducing the power distribution network to a manageable size, solving the equation of the reduced network and then recovering the solutions of the original network. With the proposed variable reduction technique, we have achieved speed improvements up to several tens of times compared to the high performance linear system solution techniques with no loss of accuracy.

1. Introduction

With reduced supply voltage, increased power density, and shrunk wires in DSM (Deep Submicron) designs, the probability of failures related to the power distribution network is increasing. IR drop due to the parasitic resistance of the power supply lines weakens the driving capability of the gates, increases the overall delay, and reduces the noise margin. The excessive high current density leads to the growth of voids and shorts at the power supply lines due to the electromigration[2][3]. Therefore, it is becoming an essential step to analyze power distribution network, fix critical areas, and achieve excellent voltage regulation across the chip[3].

One of the most critical issues in the analysis of power distribution networks is the size of the networks [2]. In the power distribution networks extracted from state-of-the-art microprocessor or high speed memory designs, there are 1 million to 100 million nodes and parasitics. Simulating such networks together with all the nonlinear gates in the chip is computationally infeasible. To cope with this problem, the simulation is done in two steps[1-7]. First, the current profile drawn by an individual gate is measured assuming a perfect supply voltage. Then

measured current profiles are modeled as independent time-varying current sources for simulating the power distribution network and voltage drops at the gates and current densities through the branches are measured. By performing these two steps, the power distribution network analysis problem reduces to solving a linear network. However, the size of the network is still intractable. Furthermore, circuit reduction techniques such as PRIMA[8] and TICER[9] cannot be applied to power distribution network analysis because not only they cannot handle circuits with independent current sources but current densities through all branches in the network should be estimated for electromigration analysis[3]. These demonstrate the demand on the capacity of the analysis tools. Although several methods such as a hierarchical approach[5] and a multi-grid technique[6] were proposed to alleviate this problem by sparsfication, the accuracy was compromised.

In this paper, we propose a variable reduction technique for the analysis of large-scale power distribution networks. Originally, variable reduction (or substitution) method is one of the solution techniques for linear systems[11]. We had inspired by the procedure of the variable reduction method, and have applied it to the power distribution network analysis. The proposed technique consists of three main steps. First, reduce a power distribution network to a manageable size and next, analyze the reduced network. Finally, recover the solutions of the original network based on the solutions of the reduced network. With the proposed variable reduction technique, it is possible to reduce any circuit whether there are independent current sources or not. Moreover, the solutions of the original network can be recovered without any loss of accuracy, which is the essential feature for electomigration analysis.

To apply the proposed technique to the analysis of realistic large-scale power distribution networks, we have considered several things in the whole procedure. In the network reduction step, we have simplified the network reduction to a set of linear transforms and defined node selection rules to increase network reduction ratio. A simple set of linear transforms for recovering the solutions of the original network accurately has been presented in the solution recovery step. Finally, we have introduced

repeated network reductions and solution recoveries to increase efficiency of the analysis.

This paper is organized as follows. The Section 2 briefly describes the overview of power distribution network analysis. In Section 3, we propose a variable reduction technique for the analysis of large-scale power distribution networks. Experimental results on real design examples show the performance of the technique in Section 4. We conclude in Section 5.

2. Overview of Power Distribution Network Analysis

A power distribution network consists of linear (power supply lines) and nonlinear (gates or transistors) portions. The analysis of such a power distribution network is converted to a problem of solving linear system, after the modeling techniques[1][2][4] as follow:

- The power grid is modeled as an R or RC network
- The nonlinear devices (gates or transistors) are modeled as independent current sources
- The decoupling capacitors are modeled as lumped capacitors
- The package is modeled as an RLC network

The power distribution network therefore becomes a linear circuit as shown in Fig. 1.
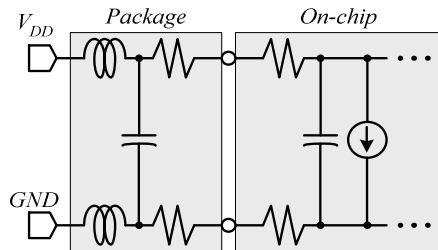

Fig. 1. Power distribution network modeling.

Behavior of the circuit as shown in Fig. 1 can be expressed as the following ordinary differential equation using MNA (Modified Nodal Analysis) formulation [5][6]:

$$G \cdot x(t) + C \cdot x^{'}(t) = u(t) \qquad (1)$$

where G is the conductance matrix resulting from resistive components, C is the admittance matrix resulting from capacitive and inductive components, $x(t)$ is the vector of time-varying voltages at the nodes and currents through the inductors and voltage sources, and $u(t)$ is the vector of independent current sources. Applying Backward Euler (BE) integration[10] to this equation, we can obtain a set of linear equations:

$$\left(G + C/h\right) \cdot x(t) = u(t) + C/h \cdot x(t - h) \qquad (2)$$

which can be expressed as a form of $Ax(t)=b(t)$, where $A=G+C/h$ and $b(t)=u(t)+C/h\cdot x(t-h)$. If we keep time step h constant in (2), the coefficient matrix A is never changed and only one initial factorization is required, with forward and backward substitution procedures at each time step[5][6]. Since, for large linear equations, the factorization of the coefficient matrix is significantly more expensive than the substitution procedure, the use of a constant time step results in large savings.

When x consists of only node voltages, as in case of MNA formulation of a RC network with current sources only, the coefficient matrix can be shown to be symmetric and positive definite. Even when inductive elements and voltage sources are included in the analysis, the symmetric positive definite formulation is feasible with an additional reduction step using the MNA formulation[5]. If the coefficient matrix is symmetric positive definite and very sparse, the system can be solved very efficiently using specialized linear system solution techniques, such as Cholesky factorization (direct method)[10][11] and preconditioned conjugate gradient (iterative method) technique[7][10].

3. Variable Reduction Technique

The variable reduction method is one of the solution techniques for linear systems such as Gaussian elimination, Cramer's rule, and so on[11]. The procedure of the variable reduction method is that first, reducing the number of the variables by substituting some variables, then estimating the solutions of kept variables, and finally recovering the solutions of removed variables based on the solutions of kept variables. This procedure inspired us to develop a new simulation technique for the analysis of large-scale power distribution networks.

3.1. Network Reduction

In this step, all nodes in the network are classified into two categories of removed and kept. Then, given power distribution network is reduced to a smaller one.

Consider an N-terminal star power distribution network as shown in Fig. 2. The center of the star is node N, and N terminals are labeled 0 to N-1. A branch consisting of a conductance joins each terminal to the central node and a capacitance and current source in parallel joins to the ground. Some elements may be absent, in which case the value of corresponding element is zero. The nodal equation of the network is similar to (1), where $C\in \mathbf{R}^{(N+1)\times(N+1)}$ and $G\in \mathbf{R}^{(N+1)\times(N+1)}$ are the nodal capacitance and conductance matrices, $x\in \mathbf{R}^{N+1}$ is the

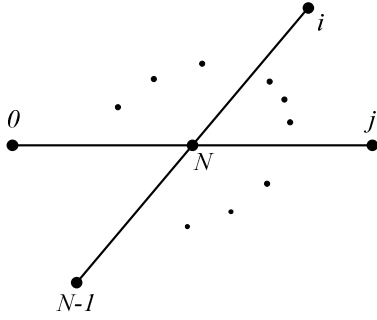vector of nodal voltages, and $u \in \mathbf{R}^{N+1}$ is the vector of current sources at the nodes.

Fig. 2. An N-terminal star network.

For the transient analysis, applying Backward Euler integration to obtain a linear equation:

$$A \cdot x(t) = b(t) \qquad (3)$$

where $A = G + C/h$ and $b(t) = u(t) + C/h \cdot x(t-h)$. For simplicity, assume that the node to be removed is the last node N. Writing (3) as a block system:

$$\begin{bmatrix} \tilde{A} & a \\ a^T & g_N + c_N/h \end{bmatrix} \begin{bmatrix} \tilde{x} \\ x_N \end{bmatrix} = \begin{bmatrix} \tilde{b} \\ b_N \end{bmatrix} \qquad (4)$$

and substitute x_N from the second block equation into the first block equation to obtain

$$(\tilde{A} - E) \cdot \tilde{x} = \tilde{b} + F \qquad (5)$$

$$E_{ij} = \begin{cases} \dfrac{g_{iN}^2}{g_N + c_N/h}, & \text{if } i = j \\[3mm] \dfrac{g_{iN} \cdot g_{jN}}{g_N + c_N/h}, & \text{otherwise} \end{cases} \qquad (6)$$

$$F_i = \dfrac{g_{iN}}{g_N + c_N/h} \cdot b_N \qquad (7)$$

In these equations, g_N is defined as

$$g_N = \sum_{k=0}^{N-1} g_{kN} \qquad (8)$$

with some terms in the sum possibly being zero and c_N is capacitance between node N and ground. The equations (5) through (7) can be translated into a procedure for

physically modifying the circuit. To remove node N, eliminate all resistors connecting the other nodes to node N. Then insert new resistors between former neighbors of N according to the following reduction rules. If different nodes i and j had been connected to N through conductances g_{iN} and g_{jN}, insert a conductance $g_{iN}g_{jN}/(g_N+c_N/h)$ between i and j. At the same time, the values $g_{iN}^2/(g_N+c_N/h)$ and $g_{jN}^2/(g_N+c_N/h)$ should be subtracted from A_{ii} and A_{jj} respectively, those are diagonal elements of coefficient matrix A in (3) corresponding to node i and j. If b_N that is RHS (Right-Hand Side) element corresponding to node N is not zero, add $b_N g_{iN}/(g_N+c_N/h)$ to b_i, where i is a former neighbor of N.

To simplify the above elimination rules, we defined the reduction operator R as follow:

$$R = \begin{bmatrix} 1 & 0 & 0 & \cdots & \dfrac{g_{0N}}{g_N + c_N/h} \\[3mm] 0 & 1 & 0 & \cdots & \dfrac{g_{1N}}{g_N + c_N/h} \\[3mm] \vdots & \vdots & \vdots & \cdots & \vdots \\[3mm] 0 & 0 & 0 & \cdots & \dfrac{g_{(N-1)N}}{g_N + c_N/h} \end{bmatrix} \qquad (9)$$

where the number of rows is the same as the number of kept nodes and the number of columns is of total nodes. The last column of R in (9) corresponds to the removed node N. We can obtain the reduced power distribution network with a simple set of linear transforms:

$$R \cdot \begin{bmatrix} \tilde{A} \\ a^T \end{bmatrix} \cdot \tilde{x}(t) = R \cdot b(t)$$

$$\rightarrow A_2 \cdot \tilde{x}(t) = b_2(t) \qquad (10)$$

where $A_2 \in \mathbf{R}^{N \times N}$ and $b_2 \in \mathbf{R}^N$ are coefficient matrix and RHS vector of the reduced network.

To apply the proposed variable reduction technique to realistic power distribution networks, some node selection rules to determine which nodes are kept and which are removed to be defined. The major objective of the node selection rules is to remove as many nodes as possible while maintaining the way to recover the voltages of removed nodes from the voltages of kept nodes. According to this condition, we define node selection rules as follow:

Rule 1: If a certain node is selected as removed, all the neighbors of the node should be kept.

Rule 2: Among all nodes, the node of less degree should be selected as removed first.

Rule 3: If the number of neighbors of a certain node is larger than d_{MAX} the node should be kept.

Rule 1 is defined to maintain the way to recover the voltages of removed and rule 2 is to increase the number of removed nodes under the limitation of rule 1. Rule 3 is to prevent the reduced network from being much denser than original, i.e., maintain the sparsity of the network. d_{MAX} is a user-defined number to control the sparsity of the reduced network.

The network reduction step of the proposed variable reduction technique is summarized as follow:

① According to the proposed node selection rules, classify the nodes into two groups of kept and removed.
② Build the reduction operator as shown in (9)
③ With the linear transforms as shown in (10), reduce the network to get a smaller one.

3.2. Reduced Network Analysis

After the network reduction step, the linear equation obtained from the reduced network can be solved with various solution techniques such as Cholesky factorization and preconditioned conjugate gradient. For transient analysis, direct methods with a constant time step are better choice. In this case, similar to factorization of the coefficient matrix, network reduction in (10) can be performed only once initially, which is the most expensive step in the proposed variable reduction technique.

3.2. Solution Recovery

An advantage of the proposed variable reduction technique is the ability of recovering the voltages of removed node exactly. The advantage is available even on the nodes those have the connections of current sources. The voltage of removed node N can be recovered from the voltages of neighbors as follow:

$$x_N(t) = \frac{1}{g_N + c_N/h}\left[\sum_{k=0}^{N-1} g_{kN}x_k(t) + b_N(t)\right] \qquad (11)$$

where $b_N(t)$ is RHS vector element corresponding to removed node N.

The recovery step is simplified to a set of linear transforms with pre-defined reduction operator and recovery operator defined as

$$C^T = \begin{bmatrix} 0 & 0 & \cdots & \dfrac{b_N(t)}{g_N + c_N/h} \end{bmatrix} \qquad (12)$$

where $C \in \mathbf{R}^{N+1}$, the elements corresponding to kept nodes are zeros and to removed are the same as the value of the last element in (12). The exact voltages of the removed nodes can be recovered with a set linear transforms as follow:

$$\begin{bmatrix} \tilde{x}(t) \\ x_N(t) \end{bmatrix} = R^T \cdot \tilde{x}(t) + C \qquad (13)$$

3.2. Multi-level Reductions and Recoveries

It can be easily found out that the reduced network in (5) and (10) is complete itself, i.e., it is also a power distribution network. Thus, the proposed variable reduction technique presented in 3.1 and 3.3 can readily applied to the reduced network once again. In this way, repeated network reductions and solution recoveries can be applied to obtain much smaller power distribution networks. First, reduce the network multiple times and factor the coefficient matrix of the network as shown in Fig. 3.
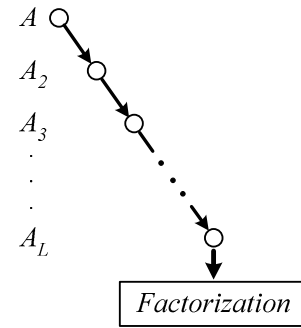
Fig. 3. Multi-level network reduction.

After the network reduction and factorization step that is the most expensive in the proposed technique, RHS vector reduction, substitutions, and then solution recoveries are performed respectively at each time point as shown in Fig. 4.
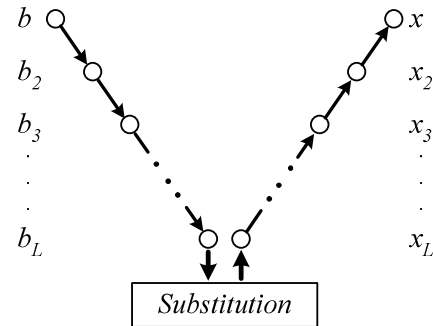
Fig. 4. Successive solutions at each time point.

Table 1. Network reduction results for large examples.

chip	original networks		reduced networks		reduction ratios	
	# of nodes [million]	# of resistors [million]	# of nodes [million]	# of resistors [million]	nodes [%]	resistors [%]
1	4.8	5.4	0.34	0.8	92.9	85.2
2	14.6	22.3	2.6	6.3	82.2	71.7
3	29.5	38.8	4.9	15.1	83.4	61.1

Table 2. CPU times and memory usages for large examples.

chip	without variable reduction		with variable reduction		improvements	
	CPU time [s]	memory [GB]	CPU time [s]	memory [GB]	speed	memory
1	375	1.92	178	1.16	×2.1	0.76G↓
2	3690	6.5	1076	7.34	×3.4	0.84G↑
3	144000	25.9	2843	15.8	×50.6	10.1G↓

4. Experimental Results

The proposed variable reduction technique is implemented using C language and integrated into an in-house power distribution network analysis tool. An efficient direct linear solver based on Cholesky factorization is used to analyze reduced networks. All the analyses are run on SUN Blade2000 workstation.

The efficiency of the proposed variable reduction technique is illustrated by applying it to the analysis of large-scale power distribution networks as shown in Table 1. Chip 1 and 3 in Table 1 are designs of high performance RISC microprocessors and chip 2 is a high speed memory design. For each design, we obtain a power distribution network model as shown in Fig. 1 and simulate the power distribution networks using both of the direct method based on Cholesky factorization and the proposed variable reduction technique.

Fig. 5 shows the result of repeated network reductions for chip 1.
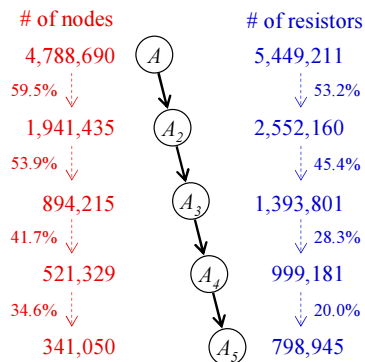

Fig.5. Result of repeated network reductions (chip 1)

With the repeated network reductions of the proposed variable reduction technique, the example networks are reduced to a manageable size as shown in Table 1. Reduction ratios of nodes and resistors are over 80% and 60% respectively. In the experiments, we set d_{MAX} defined in node selection rules as 4 and repetition level as 4 or 5. Although the numbers are not optimum ones, these are sufficient to obtain manageable size networks.

CPU times and memory usages for the example circuits are shown in Table 2. For the smaller network of chip 1, the factor of speed improvement is only about 2 times. For the larger network of chip 3, however, the factor is about 50 and savings in memory usage is over 10GBytes. These are dramatic improvements on both CPU time and memory usage. The reason why memory usage for chip 2 is increased with the proposed technique is that we analyze the original network using not a direct solver but iterative one. Simulating the power distribution networks of chip 2 and 3 without the proposed variable reduction technique imposes the use of an iterative solver based on preconditioned conjugate gradient technique due to memory limitation. On the other hand, the variable reduction technique utilizes a direct method based on Cholesky factorization resulting in more speed improvements for transient analysis. In case of chip 2, the CPU time for the V-cycle as shown in Fig. 4 is only about 18 seconds while the total CPU time is 1076 seconds.

After the voltages of all nodes in the original network are recovered as shown in Fig. 6, electormigration analysis can be performed with negligible effort by estimating the current densities through all branches.

5. Conclusion

In this paper, we present an efficient variable reduction technique for the analysis of large-scale power distribution networks. By reducing the network to a smaller one, solving the reduced network and then recovering the solutions of the original network, we are able to reduce both the analysis time and the memory usage dramatically with no loss of accuracy. Experimental results on realistic design examples show speed improvements up to several tens of times and savings of memory usage over 10GBytes compared to state-of-the-art linear system solution techniques.
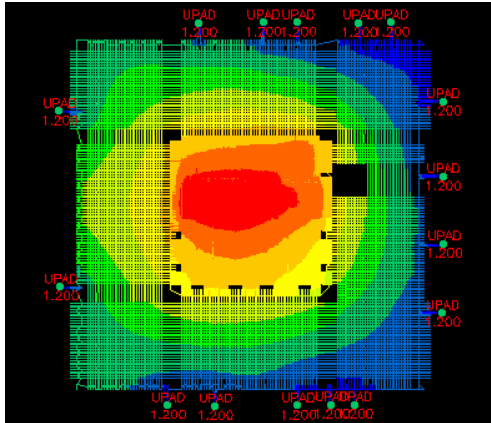

Fig. 6. Result of solution recovery (chip 3).

References

[1] H. H. Chen and J. S. Neely, "Interconnect and Circuit Modeling Technologies for Full-Chip Power Supply Noise Analysis," *IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part B, Vol. 21, No. 3*, pp. 209-215, Aug., 1998

[2] R. Chaudhry, R. Panda, T. Edwards, and D. Blaauw, "Design and Analysis of Power Distribution Networks with Accurate RLC Models," *in Proceedings of Design Automation Conference*, pp. 738-743, 1988.

[3] G. Steele, D. Overhauser, S. Rochel, and S. Z. Hussain, "Full-Chip Verification Methods for DSM Power Distribution Systems," *in Proceedings of Design Automation Conference*, pp. 744-749, 1988.

[4] R. Panda, D. Blaauw, R. Chaudhry, V. Zolotov, B. Young, and R.Ramaraju, "Model and Analysis for Combined Package and On-Chip Power Grid Simulation," *International Symposium on Low Power Electronics and Design*, pp. 179-184, 2000.

[5] M. Zhao, R. Panda, S. Sapatnekar, T. Edwards, R. Chaudhry, D. Blaauw, "Hierarchical Analysis of Power Distribution Networks," *in Proceedings of Design Automation Conference*, pp. 150-155, 2000

[6] S. R. Nassif and J. N. Kozhaya, "Multi-Grid Methods for Power Grid Simulation," *IEEE International Symposium on Circuits and Systems*, Volume: 5, pp. 457-460, 2000.

[7] T. Chen and C. Chen, "Efficient Large-Scale Power Grid Analysis Based on Preconditioned Krylov-Subspace Iterative Methods," *in Proceedings of Design Automation Conference*, pp. 559-562, 2001.

[8] A. Odabasioglu, M. Celik, and L.T. Pileggi, "PRIMA: Passive Reduced-Order Interconnect Macromodeling Algorithm," *IEEE/ACM Inter-national Conference on Computer-Aided Design*, pp. 58-65, 1997

[9] B. N. Sheehan, "TICER: Realizable Reduction of Extracted RC Circuits," *in Proceedings of IEEE International conference of CAD*, pp. 200-203, 1999.

[10] W. H. Press et. al., *Numerical Recipes in C.* Cambridge University Press, 1988.

[11] Gilbert Strang, *Linear Algebra and Its Applications*, International Thomson Publishing; 1988

IEEE
COMPUTER
SOCIETY