European Design & Test Conference,
Paris, France, 3/95

Efficent Reduced-Order Modeling of Frequency-Dependent Coupling Inductances associated with 3-D Interconnect Structures

L. Miguel Silveira Mattan Kamon Jacob White
Research Laboratory of Electronics
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139
e-mail: {lms,matt,white@rle-vlsi.mit.edu}

Abstract

Reduced-order modeling techniques are now commonly used to efficiently simulate circuits combined with interconnect, but generating reduced-order models from realistic 3-D structures has received less attention. In this paper we describe a Krylov-subspace based method for deriving reduced-order models directly from the 3-D magnetoquasistatic analysis program FASTHENRY. This new approach is no more expensive than computing an impedance matrix at a single frequency.

1 Introduction

The dense three-dimensional packaging used in compact electronic systems may produce magnetic interactions which interfere with system performance. Such effects are difficult to simulate because they occur only as a result of an interaction between the field distribution in a complicated geometry of conductors, and the circuitry connected to those conductors. Recent work on reduced-order modeling techniques have made it possible to efficiently simulate circuits combined with interconnect [1], but generating the reduced-order models from realistic 3-D structures has received less attention. The most commonly used approach to generating reduced-order models is to use a 3-D field solver to compute impedance matrices over a range of frequencies, and then use a rational polynomial fitting algorithm [2]. This approach has been shown to produce accurate frequency-domain reduced-order models which are directly amenable to inclusion into a standard circuit simulator [3].

In order to use frequency-domain fitting as described above, it is necessary to use the field solver to compute impedance matrices at dozens of frequency points, and this is computationally expensive. It is possible to derive a more efficient approach by exploiting the fact that 3-D field solvers typically use Krylov-subspace based iterative methods. These iterative methods can provide more than just a solution at a particular frequency, they can be used to directly construct reduced-order models [4].

In this paper, we present a numerically robust and accurate approach for computing reduced-order models of magnetoquasistatic coupling in complicated 3-D structures. The approach is based on using the multipole-accelerated program FASTHENRY [5], combined with the Krylov-subspace algorithm Arnoldi [6]. We begin, in section 2, by describing the mesh-formulation approach of FASTHENRY. In section 3, the standard Padé approximation approach as well as an Arnoldi-based approach are derived. In section 4 results are presented comparing the accuracy of the two model-order reduction methods on an RLC filter and a package example. Finally, in section 5, we present conclusions and acknowledgments.

2 The Mesh Formulation Approach

The frequency dependent resistance and inductance matrices describing the terminal behavior of a set of conductors can be rapidly computed with the multipole-accelerated mesh-formulation approach as implemented in FASTHENRY [5]. To describe the approach, consider that each conductor is approximated as piecewise-straight sections. The volume of each straight section is then discretized into a collection of parallel thin filaments through which current is assumed to flow uniformly.

To derive a system of equations for the filament currents, we start by assuming the system is in sinusoidal steady-state and following the partial inductance approach in [7], the branch current phasors can be re-

lated to branch voltage phasors by

$$\boldsymbol{V}_b = (\boldsymbol{R} + j\omega\boldsymbol{L})\boldsymbol{I}_b = \boldsymbol{Z}\boldsymbol{I}_b \tag{1}$$

where $\boldsymbol{V}_b$, $\boldsymbol{I}_b \in \mathbb{C}^b$, b is the number of branches (number of current filaments), and ω is excitation frequency. The entries of the diagonal matrix $\boldsymbol{R} \in \mathbb{R}^{b\times b}$ represent the DC resistance of each current filament, and $\boldsymbol{L} \in \mathbb{R}^{b\times b}$ is the dense matrix of partial inductances.

Kirchhoff's voltage law, which implies that the sum of branch voltages around each mesh (a mesh is any loop of branches in the graph which does not enclose any other branches) in the network is represented by

$$\boldsymbol{M}\boldsymbol{V}_b = \boldsymbol{V}_s \qquad \boldsymbol{M}^T\boldsymbol{I}_m = \boldsymbol{I}_b, \tag{2}$$

where $\boldsymbol{V}_s \in \mathbb{C}^m$ is the mostly zero vector of source branch voltages, $\boldsymbol{I}_m \in \mathbb{C}^m$ is the vector of mesh currents, $\boldsymbol{M} \in \mathbb{R}^{m\times b}$ is the mesh matrix. Combining (2) and (1) yields

$$\boldsymbol{M}\boldsymbol{Z}\boldsymbol{M}^T\boldsymbol{I}_m = \boldsymbol{V}_s. \tag{3}$$

The complex admittance matrix which describes the external terminal behavior of a t-conductor system, denoted $\boldsymbol{Y}_t = \boldsymbol{Z}_t^{-1}$, can by derived from (3) by noting that

$$\boldsymbol{I}_t = \boldsymbol{Y}_t\boldsymbol{V}_t. \tag{4}$$

$\boldsymbol{I}_t$ and $\boldsymbol{V}_t$ are the terminal source currents and voltages of the t-conductor system, which are related to the mesh quantities by $\boldsymbol{I}_t = \boldsymbol{N}^T\boldsymbol{I}_m, \boldsymbol{V}_s = \boldsymbol{N}\boldsymbol{V}_t$, where $\boldsymbol{N} \in \mathbb{C}^{m\times t}$ is a terminal incidence matrix that is easily derived when the mesh equations are formulated.

Hence, to compute the i^{th} column of $\boldsymbol{Y}_t$, solve (3) with a $\boldsymbol{V}_s$ whose only nonzero entry corresponds to $\boldsymbol{V}_{t_i}$, and then extract the entries of $\boldsymbol{I}_m$ associated with the source branches.

To solve (3) by Gaussian Elimination would require $\mathcal{O}(m^3)$ operations. Instead, programs like FASTHENRY solve (3) using a multipole-accelerated GMRES iterative algorithm [6], which requires $\mathcal{O}(b)$ operations. The complexity is reduced to $\mathcal{O}(m^2)$ by using GMRES instead of Gaussian elimination, and then down to $\mathcal{O}(b)$ by using a hierarchical multipole algorithm [8].

3 Reduced-Order Modeling

3.1 State-Space Formulation

As mentioned in the introduction, to use frequency-domain fitting to generate a reduced-order model for the frequency-dependent entries of $\boldsymbol{Y}_t$, it would be necessary to construct and solve (3) for dozens of values of ω. To derive a more efficient approach, consider forming the state-space representation of (3). To that end, expand $\boldsymbol{Z}$ into $\boldsymbol{R} + s\boldsymbol{L}$ to get

$$\begin{aligned} s(\boldsymbol{M}\boldsymbol{L}\boldsymbol{M}^T)\boldsymbol{I}_m &= -(\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T)\boldsymbol{I}_m + \boldsymbol{N}\boldsymbol{V}_t \\ \boldsymbol{I}_t &= \boldsymbol{N}^T\boldsymbol{I}_m. \end{aligned} \tag{5}$$

With the representation in (5), the (i,j)-th entry of the complex admittance matrix computed using a set of terminal voltages whose only nonzero entry corresponds to $\boldsymbol{V}_{t_j}$, and written as

$$\frac{\boldsymbol{I}_{t_i}}{\boldsymbol{V}_{t_j}} = \boldsymbol{Y}_{t_{i,j}}(s) = \boldsymbol{c}^T(\boldsymbol{I} - s\boldsymbol{A})^{-1}\boldsymbol{b} \tag{6}$$

where $\boldsymbol{A} = -(\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T)^{-1}(\boldsymbol{M}\boldsymbol{L}\boldsymbol{M}^T)$ and $\boldsymbol{b} = (\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T)^{-1}\boldsymbol{N}_j$ and $\boldsymbol{c} = \boldsymbol{N}_i$, where $\boldsymbol{N}_i$ indicates the i^{th} column of $\boldsymbol{N}$. It is possible to derive extensions of all the approximations methods mentioned in this paper to directly compute approximations to the system in (5) directly, that is a system with t inputs and t outputs. In the remainder however, we will for the most part restrict our discussion to single-input single-output systems characterized by a transfer function such as (6).

The standard approach to derive a reduced-order model of (6) is to compute a Padé approximation [9]. To that end note that

$$\boldsymbol{Y}_{t_{ij}}(s) = \boldsymbol{c}^T(\boldsymbol{I} - s\boldsymbol{A})^{-1}\boldsymbol{b} = \sum_{k=0}^{\infty} m_k s^k. \tag{7}$$

where $m_k = \boldsymbol{c}^T\boldsymbol{A}^k\boldsymbol{b}$ is the k^{th} moment of the transfer function. A Padé approximation of q^{th} order is defined as the rational function

$$G_q^P(s) = \frac{b_{q-1}s^{q-1} + \cdots + b_1 s + b_0}{a_q s^q + a_{q-1}s^{q-1} + \cdots + a_1 s + 1} \tag{8}$$

whose coefficients are selected to match the first $2q-1$ moments of the transfer function.

Padé approximates can be computed using direct evaluation of the moments, though the approach is ill-conditioned, because the computation of the moments relies on a power iteration with the system matrix $\boldsymbol{A}$. Instead, Lanczos-style algorithms can be used [4], that are numerically more robust.

3.2 Arnoldi-based Approximations

An alternative approach, which robustly generates a somewhat different approximation, can be de-

Algorithm 1 (Arnoldi process)

```
arnoldi(input A, b, q; output
  V_q, v_{q+1}, H_q, h_{j+1,j})
{
  v_1 = b/||b||_2
  for (j = 1; j <= q; j++) {
    w = A v_j
    for (i = 1; i <= j - 1; i++) {
      h_{i,j} = w^T v_i
      w = w - h_{i,j} v_i
    }
    h_{j+1,j} = ||w||_2
    if (h_{j+1,j} != 0) {
      v_{j+1} = w/h_{j+1,j}
    }
  }
  V_q = [v_1 ... v_q]
  H_q = (h_{i,j}),  i,j = 1, ..., q
}
```

rived using an Arnoldi process as in the GMRES algorithm. The idea behind this approach is similar to that of [4], and is that of selecting an orthonormal basis for the Krylov subspace $\mathcal{K}_k(\boldsymbol{A}, \boldsymbol{b}) = span\{\boldsymbol{b}, \boldsymbol{Ab}, \boldsymbol{A}^2\boldsymbol{b}, \cdots, \boldsymbol{A}^{k-1}\boldsymbol{b}\}$. The Arnoldi algorithm is a better conditioned process than direct evaluation of the moments because it generates an orthogonal set of vectors which span $\boldsymbol{A}^k\boldsymbol{b}, k = 0, \ldots, q$. Note that the computation of $\boldsymbol{b}$ is inexpensive since $\boldsymbol{MRM}^T$ is sparse. Also, because L is dense, the dominant cost of each step of an Arnoldi process is a matrix-vector product, $\boldsymbol{Ax} = -(\boldsymbol{MRM}^T)^{-1}(\boldsymbol{MLM}^T)\boldsymbol{x}$. In practice, the matrix-vector cost dominates even when the dense part, $(\boldsymbol{MLM}^T)\boldsymbol{x}$, is rapidly computed with a hierarchical multipole-algorithm as in FASTHENRY. The basic outline of the Arnoldi process is given in Algorithm 1.

After q steps, the Arnoldi algorithm returns a set of q $\boldsymbol{A}$-orthonormal vectors, as the columns of the matrix $\boldsymbol{V}_q \in \mathbb{R}^{m\times q}$, and a $q \times q$ upper Hessenberg matrix $\boldsymbol{H}_q$ whose entries are the scalars $h_{i,j}$ generated by Algorithm 1. These two matrices satisfy the following relationship:

$$\boldsymbol{AV}_q = \boldsymbol{V}_q\boldsymbol{H}_q + h_{j+1,j}\boldsymbol{v}_{q+1}\boldsymbol{e}_q^T \tag{9}$$

where $\boldsymbol{e}_q$ is the q^{th} unit vector in $\mathbb{R}^{m\times m}$.

From (9), it can easily be seen that after q steps of an Arnoldi process, for $k < q-1$,

$$\boldsymbol{A}^k\boldsymbol{b} = \|\boldsymbol{b}\|_2\boldsymbol{A}^k\boldsymbol{V}_q\boldsymbol{e}_1 = \|\boldsymbol{b}\|_2\boldsymbol{V}_q\boldsymbol{H}_q^k\boldsymbol{e}_1. \tag{10}$$

With this relation, the moments can be related to $\boldsymbol{H}_q$ by

$$m_k = \boldsymbol{c}^T\boldsymbol{A}^k\boldsymbol{b} = \|\boldsymbol{b}\|_2\boldsymbol{c}^T\boldsymbol{V}_q\boldsymbol{H}_q^k\boldsymbol{e}_1 \tag{11}$$

and so the q^{th} order Arnoldi-based approximation to $\boldsymbol{Y}_{ij}$ can be written as

$$\boldsymbol{G}_q^A(s) = \|\boldsymbol{b}\|_2\boldsymbol{c}^T\boldsymbol{V}_q\left(\boldsymbol{I} - s\boldsymbol{H}_q\right)^{-1}\boldsymbol{e}_1 \tag{12}$$

corresponding to the state-space realization using the triplet $[\boldsymbol{A}_k, \boldsymbol{b}_k, \boldsymbol{c}_k] = [\boldsymbol{H}_q, \boldsymbol{e}_1, \|\boldsymbol{b}\|_2\boldsymbol{V}_q^T\boldsymbol{c}]$.

Using the eigendecomposition $\boldsymbol{H}_q = \boldsymbol{S}_q\Lambda_q\boldsymbol{S}_q^{-1}$, the expression for the approximating rational function becomes

$$\boldsymbol{G}_q^A(s) = \|\boldsymbol{b}\|_2\sum_{k=0}^{q}\frac{\mu_k\nu_k}{s - p_k} \tag{13}$$

where $\boldsymbol{\mu} = -\boldsymbol{c}^T\boldsymbol{V}_q\boldsymbol{S}_q\Lambda_q^{-1}$, $\boldsymbol{\nu} = \boldsymbol{S}_q^{-1}\boldsymbol{e}_1$, and $\boldsymbol{p} = diag(\Lambda_q^{-1})$ are the poles of the approximation.

Note that the rational function $\boldsymbol{G}_q^A(s)$ is *not* a Padé approximation as it has q poles, but only matches $q-2$ moments, since (10) is only valid for $k < q-1$. However, computing the rational function requires only q matrix-vector products, roughly half the number of matrix-vector products required to compute a q^{th} order Padé approximate which matches $2q-1$ moments. For the same computational effort required to compute the q^{th} order Padé approximant $\boldsymbol{G}_q^P(s)$ one could obtain $\boldsymbol{G}_{2q}^A(s)$, which has $2q$ poles and matches $2q-2$ moments. The extra number of poles presumably implies more accuracy in the approximation, which is obtained at no extra cost. It is not clear whether the loss of moments matched is of any concern. The moments are local properties or values of a complex function but the information sought with these approximations is more global in nature.

The Arnoldi-based approximation method, despite this "loss" of matched moments terms of matching moments is quite accurate on a global sense. Furthermore, the Arnoldi-based method, as described, has a very interesting property when used to obtain the full matrix system transfer function of a t-terminal conductor system. In that situation end one would run the Arnoldi process t times, each one with a different column of the $\boldsymbol{b}$ matrix and notice that, in each run, the derivation leading to (10) is still valid. The model then obtained via (12) would in that case consist of a full column of the complex admittance matrix, given that the full $\boldsymbol{c}$ matrix, could now be used. Therefore

each run of the Arnoldi method would produce one column of the admittance transfer function. If a Padé approximation were sought for this system, one would have to run the robust Padé-via-Lanczos algorithm t^2 times.

3.3 Bounds on the approximation error

4 Experimental Results

In the preceding section, we described algorithms to compute Padé approximations of order q and Arnoldi-based models of orders q and $2q$. In this section we compare the accuracy of these three approximations first for a difficult to model RLC filter example, and then when used to obtain reduced-order models for the frequency-dependent admittance for a small set of package pins. This reduced-order model is then used to investigate crosstalk between the package pins.

4.1 Filter Example

Figure 1 shows the Bode plots of the the 7^{th} order Padé and the 7^{th} and 14^{th} order Arnoldi-based approximations to a 14^{th}-order RLC filter's transfer function. Also shown in the picture is the exact transfer function. For the low frequency range all approximations are indistinguishable. However, for higher frequencies, as is clear from the figure, the 7^{th} order Padé and the 7^{th} order Arnoldi-based approximation have comparable accuracy, while the 14^{th} order Arnoldi-based approximation, which requires the same number of matrix-vector products as the 7^{th}-order Padé, is indistinguishable from the exact transfer function and thus much more accurate. This last observation is unsurprising, in exact arithmetic the Arnoldi-based algorithm converges to the exact transfer function of an n^{th} order system in n iterations. It should be noted, however, that any 14^{th} order approximation will be significantly more expensive to use in a circuit simulator than a 7^{th} order approximation. Nevertheless the ability to compute higher orders of approximation at no extra cost remains an interesting property of the Arnoldi-based approximation method.

4.2 Package Example

Consider the small set of package pins. as shown in Figure 2. To compute the resistance and inductance matrices with FASTHENRY, the pins were discretized into three filaments along their height and four along their length producing a system of size $m = 887$. This

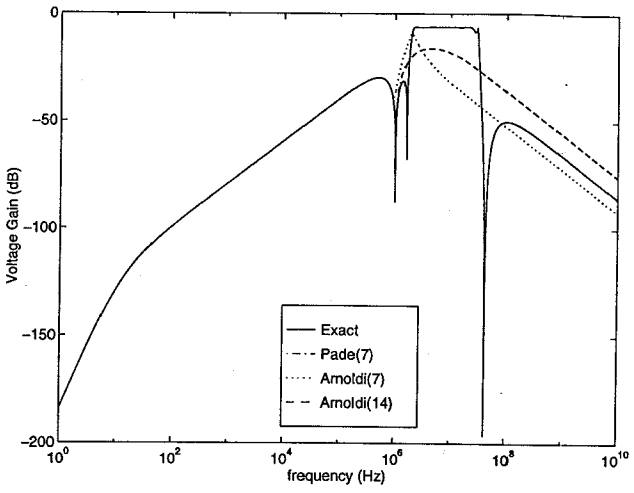

Figure 1: Bode plots for the approximations $G_7^P(s)$, $G_7^A(s)$ and $G_{14}^A(s)$ to the RLC filter's transfer function.

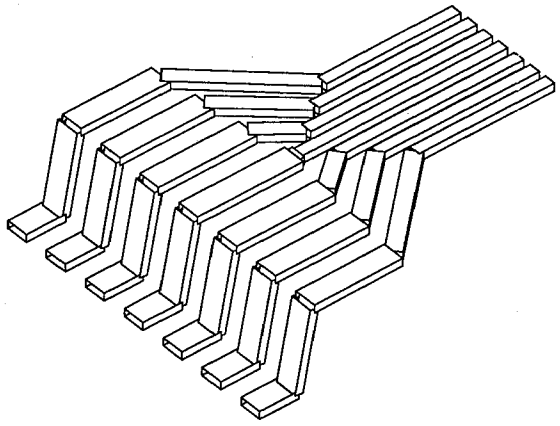

Figure 2: Seven pins of a cerquad pin package.

allows modeling of changes in resistance and inductance due to skin and proximity effects.

Figure 3 shows the Bode plots of the 5^{th} order Padé and the 5^{th} and 10^{th} order Arnoldi-based approximations to the coupled admittance transfer function between pins 1 and 2. Also shown in the picture is the exact admittance transfer function obtained by eigendecomposition of the full system. As can be seen from the plot, all three approximations seem equally accurate and are virtually indistinguishable from the full transfer function.

To investigate the crosstalk effects between the package pins in Fig 2, the configuration shown in Fig. 4 is used were it was assumed that the five middle lines carry output signals from the chip and the two outer pins carry power and ground. The signals are driven and received with CMOS inverters which are

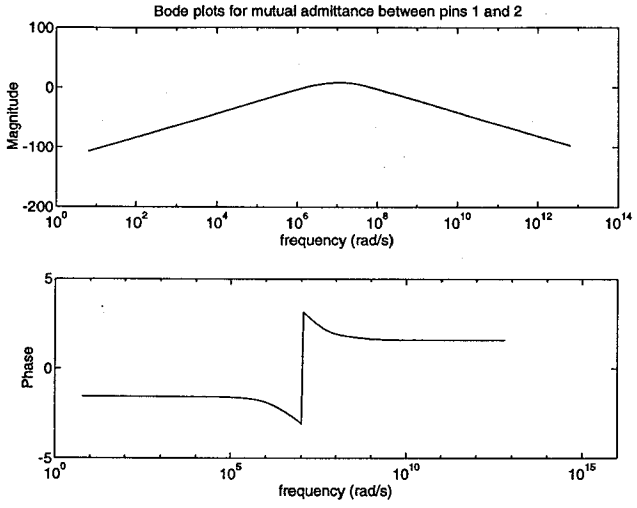

Figure 3: Bode plots for the approximations $G_5^P(s)$, $G_5^A(s)$ and $G_{10}^A(s)$ to the coupled admittance transfer function between pins 1 and 2.

capable of driving a large current to compensate for the impedance of the package pins. The capacitance is assumed to be 8pF and the interconnect from the end of pin to the receiver is modelled with a capacitance of 5pF. A $0.1\mu F$ decoupling capacitor is connected between the driver's power and ground to minimize supply fluctuations. The frequency dependence of each element in the admittance matrix is modeled via Arnoldi-based approximations of 5^{th} order whose error is below 5%. These models are then incorporated into SPICE3 as a frequency-dependent voltage-controlled current source VCCS. As a sample time domain simulation, imagine that at time $t_0 = 4$ns the signal on pin 4 of Fig.4 is to switch from high to low and pins 2, 3, 5, and 6 are to switch from low to high but that due to delay on chip, pins 2, 3, 5, and 6 switch at $t_1 = 5$ns. In this case, significant current will suddenly pass through the late pins while pin 4 is in transition. Due to crosstalk, this large transient of current has significant effects on the input of the receiver on pin 4, as shown in Fig. 5. Note that the input does not rise monotonically. Fig. 5 also shows that the bump in the waveform is carried through to the output of receiver, as a large glitch.

5 Conclusions

In this paper we described an accurate approach for using the iterative method in the FASTHENRY program to compute reduced-order models of frequency-dependent inductance matrices associated with com-

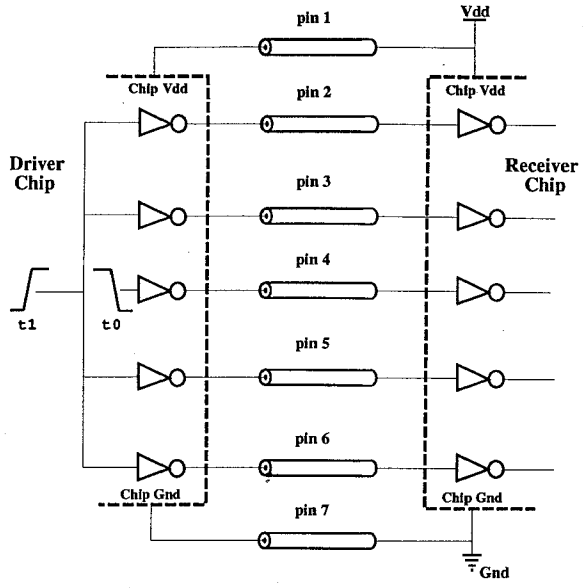

Figure 4: General configuration for the connection between received and driver chips. All the circuit elements inside the same chip share that chip's power and ground.

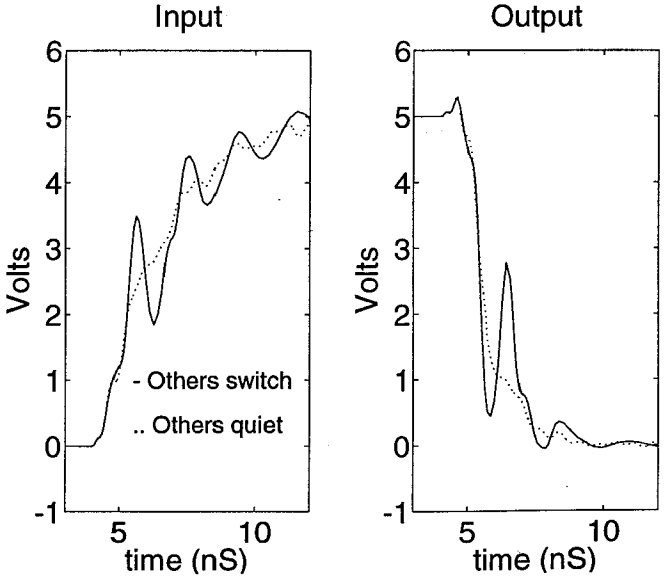

Figure 5: Results of the timing simulation of a receiver pin in the presence of changes on other adjacent pins. Pin 4's receiver when four adjacent pins switch 1ns after pin 4.

plicated 3-D structures. The key advantage of this method is that it is no more expensive than computing the inductance matrix at a single frequency. We also compared two approaches to the model-order reduction, the reformulated Padé-based approach using the Lanczos algorithm (PVL) and an Arnoldi-based approach using an algorithm based on the Arnoldi process. We showed that the Arnoldi-based algorithm can have advantages over PVL in certain applications. In particular, in the Arnoldi-based algorithm, each set of iterations produces an entire column of the inductance matrix rather than a single entry, and if matrix-vector product costs dominate then the Arnoldi-based algorithm produces a better approximation for a given amount of work. We should note that block generalizations of these algorithm can be used to directly obtain matrix transfer function for multi-terminal systems.

Acknowledgments

The authors wish to acknowledge the very helpful discussions with Dr. Peter Feldmann and Dr. Roland Freund of the A.T. & T research center.

This work was supported by the Defense Advanced Research Projects Agency under Contract N00014-91-J-1698, the Semiconductor Research Corporation under Contract SJ-558, the National Science Foundation contract (9117724-MIP), an NSF Graduate Research Fellowship, and grants from IBM and DEC.

References

[1] Lawrence T. Pillage and Ronald A. Rohrer. Asymptotic Waveform Evaluation for Timing Analysis. *IEEE Trans. CAD*, 9(4):352–366, April 1990.

[2] L. Miguel Silveira, Mattan Kamon, and Jacob K. White. Algorithms for Coupled Transient Simulation of Circuits and Complicated 3-D Packaging. In *Proceedings of the 44st Electronics Components and Technology Conference*, pages 962–970, Washington, DC, May 1994.

[3] Luís Miguel Silveira. *Model Order Reduction Techniques for Circuit Simulation.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, May 1994.

[4] Peter Feldmann and Roland W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. In *Proceeding of the EuroDAC*, September 1994.

[5] M. Kamon, M. J. Tsuk, and J. White. Fasthenry, a multipole-accelerated 3-d inductance extraction program. In *Proceedings of the ACM/IEEE Design Automation Conference*, Dallas, June 1993.

[6] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7:856–869, July 1986.

[7] A. E. Ruehli. Inductance calculations in a complex integrated circuit environment. *IBM J. Res. Develop.*, 16:470–481, September 1972.

[8] L. Greengard. *The Rapid Evaluation of Potential Fields in Particle Systems.* M.I.T. Press, Cambridge, Massachusetts, 1988.

[9] George A. Baker Jr. *Essentials of Padé Approximants.* Academic Press, New York, NY, First edition, 1975.