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

L. Miguel Silveira Mattan Kamon Jacob White

lms@inesc.pt matt@rle-vlsi.mit.edu white@rle-vlsi.mit.edu

Research Laboratory of Electronics
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139

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 has 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 function fitting algorithm [2]. This ap-

proach has been shown to produce accurate frequency-domain reduced-order models which are easily included in 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 [?], combined with the Krylov-subspace based Arnoldi algorithm [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 [?]. 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 related to branch voltage phasors by

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

where $\boldsymbol{V}_b, \boldsymbol{I}_b \in \mathbb{C}^b$, b is the number of branches (number of current filaments), and ω is the 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, **and m is the number of meshes, which is typically somewhat less than b, the number of filaments**. 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

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

$\boldsymbol{I}_t$ 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 $\boldsymbol{N} \in \mathbb{R}^{m\times t}$ is a terminal incidence matrix determined by the mesh formulation.

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 **where the number of meshes, m, can exceed 10,000 when modeling frequency dependent effects**. Instead, programs like FASTHENRY solve (3) using a multipole-accelerated GMRES iterative algorithm [6], which requires $\mathcal{O}(b)$ operations. The complexity is reduced from $\mathcal{O}(m^3)$ to $\mathcal{O}(m^2)$ by using GMRES instead of Gaussian elimination, and then to $\mathcal{O}(b)$ by using a hierarchical multipole algorithm [8].

3 Reduced-Order Modeling

One approach to coupling package models with circuits is to simply include a sparse tableau version of (3) in a circuit simulator [9]. **However, since MZM^t is large and dense, this approach can be computationally intractable.** A more computationally efficient approach is to represent (3) with a reduced-order model.

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

$$\begin{aligned} s(\boldsymbol{M}\boldsymbol{L}\boldsymbol{M}^T)\boldsymbol{I}_m &= -(\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T)\boldsymbol{I}_m + \boldsymbol{N}\boldsymbol{V}_t \\ \boldsymbol{I}_t &= \boldsymbol{N}^T\boldsymbol{I}_m. \end{aligned} \tag{4}$$

With the representation in (4), the (i,j)-th entry of the complex admittance matrix **can be** computed using a set of terminal voltages whose only nonzero entry corresponds to $\boldsymbol{V}_{t_j}$, and **can be** written as

$$\frac{\boldsymbol{I}_{t_i}}{\boldsymbol{V}_{t_j}} = \boldsymbol{Y}_{t_{ij}}(s) = \boldsymbol{c}^T\left(\boldsymbol{I} - s\boldsymbol{A}\right)^{-1}\boldsymbol{b} \tag{5}$$

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 approximation methods mentioned in this paper to directly compute approximations to the system in (4) directly, that is a system with t inputs and t outputs. In the remainder however, we will restrict our discussion to single-input single-output systems characterized by a transfer function such as (5).

The standard approach to computing a reduced-order model of (5) is to use Padé approximations [10]. Padé approximates are rational functions in the Laplace transform variable s which power series expansions of the original system transfer function up to some order. More specifically, consider the admittance

$$\boldsymbol{Y}_{t_{ij}}(s) = \boldsymbol{c}^T\left(\boldsymbol{I} - s\boldsymbol{A}\right)^{-1}\boldsymbol{b} = \sum_{k=0}^{\infty} m_k s^k. \tag{6}$$

where

$$m_k = \boldsymbol{c}^T\boldsymbol{A}^k\boldsymbol{b} \tag{7}$$

is the k^{th} moment of the transfer function. A Padé approximation of q^{th} order is defined as the rational function

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

whose coefficients are selected to match the first $2q-1$ moments of the transfer function (5).

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

3.2 Arnoldi-based Approximations

An alternative approach, which robustly generates a somewhat different approximation than Padé, can be derived 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 $\mathcal{K}_k(\boldsymbol{A}, \boldsymbol{b}) = span\{\boldsymbol{b}, \boldsymbol{A}\boldsymbol{b}, \boldsymbol{A}^2\boldsymbol{b}, \cdots, \boldsymbol{A}^{k-1}\boldsymbol{b}\}$ using a modified Gramm-Schmidt process. The orthogonality between the basis vectors makes the Arnoldi algorithm a better conditioned process than direct evaluation of the moments. Note that the computation of $\boldsymbol{b}$ is inexpensive since $\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T$ is sparse. Also, because L is dense, the dominant cost of each step of an Arnoldi process is a matrix-vector product, $\boldsymbol{A}\boldsymbol{x} = -(\boldsymbol{M}\boldsymbol{R}\boldsymbol{M}^T)^{-1}(\boldsymbol{M}\boldsymbol{L}\boldsymbol{M}^T)\boldsymbol{x}$. In practice, the matrix-vector product cost dominates even when the dense part, $(\boldsymbol{M}\boldsymbol{L}\boldsymbol{M}^T)\boldsymbol{x}$, is rapidly computed with a hierarchical multipole-algorithm as in FastHenry.

The basic outline of the Arnoldi process is given in Algorithm 3.1.

After q steps, the Arnoldi algorithm returns a set of q 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 Gramm-Schmidt orthogonalization coefficients $h_{i,j}$. Following the approach in [4], these two matrices satisfy the following relationship:

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

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

Algorithm 3.1 (Arnoldi process)

arnoldi(*in:*$\boldsymbol{A}, \boldsymbol{b}, q$; *out:*$\boldsymbol{V}_q, \boldsymbol{v}_{q+1}, \boldsymbol{H}_q, h_{j+1,j}$)

{

$\boldsymbol{v}_1 = \boldsymbol{b} / \|\boldsymbol{b}\|_2$

for $(j = 1;\ j <= q;\ j++)$ {

$\boldsymbol{w} = \boldsymbol{A}\boldsymbol{v}_j$

for $(i = 1;\ i <= j-1;\ i++)$ {

$h_{i,j} = \boldsymbol{w}^T \boldsymbol{v}_i$

$\boldsymbol{w} = \boldsymbol{w} - h_{i,j}\boldsymbol{v}_i$

}

$h_{j+1,j} = \|\boldsymbol{w}\|_2$

if $(h_{j+1,j} != 0)$ {

$\boldsymbol{v}_{j+1} = \boldsymbol{w} / h_{j+1,j}$

}

}

$\boldsymbol{V}_q = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_q]$

$\boldsymbol{H}_q = (h_{i,j}),\ \ i, j = 1, \cdots, q$

}

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. We have observed empirically the extra number of poles improves the accuracy of the GMRES approximation, however, the larger number number of poles make it a more expensive to include in a circuit simulator.

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. This last observation is not surprising, due to the finite termination properties of the algorithm. 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 a valuable property of the Arnoldi-based approximation method.

4.2 Package Example

Consider the small set of package pins 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 width producing a system of size $m = 887$. This allows modeling of changes in resistance and inductance due to skin and proximity effects. Figure 3

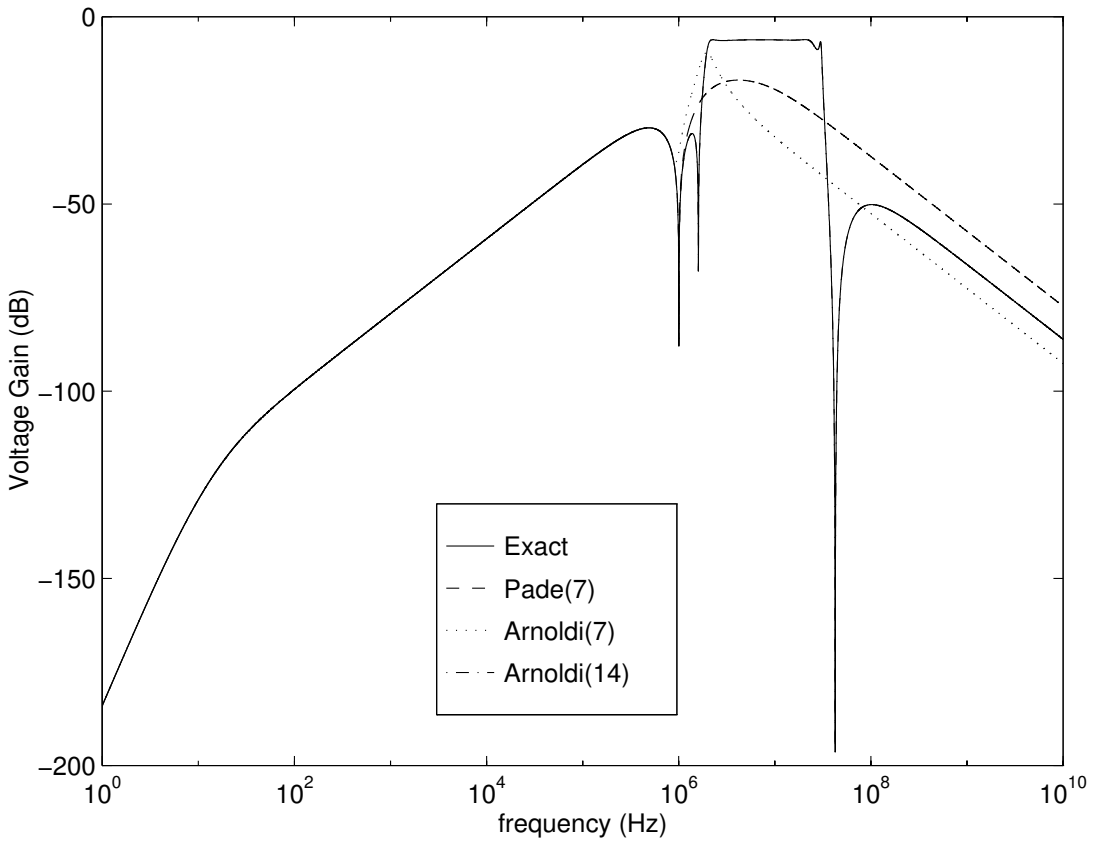

Figure 1: Bode plots for the approximations $\boldsymbol{G}_7^P(s)$, $\boldsymbol{G}_7^A(s)$ and $\boldsymbol{G}_{14}^A(s)$ to the RLC filter's transfer function.

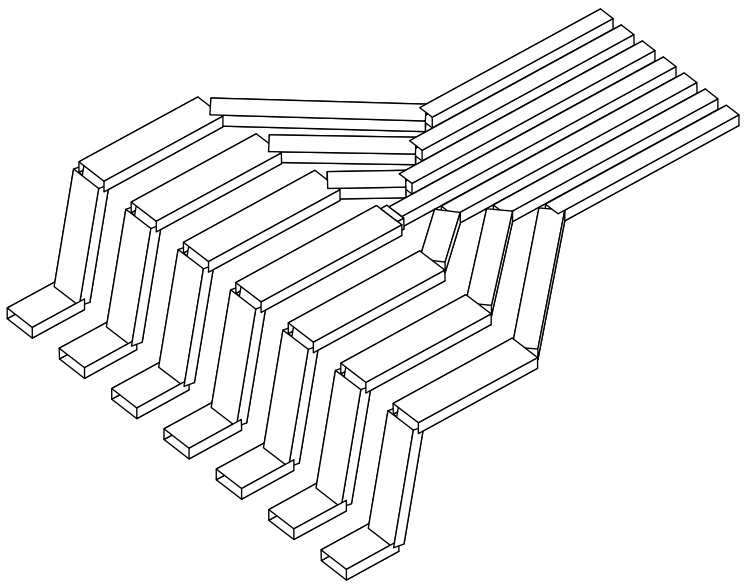

Figure 2: Seven pins of a cerquad pin package.

shows the magnitude of the error of the 8^{th} order Padé and the 8^{th} and 16^{th} order Arnoldi-based approximations to the coupled admittance transfer function between pins 1 and 2. As can be seen from the plot, all three approximation have an error well below 5%. **It is worth noting that computing the 8^{th} order Arnoldi-based model for one column of the admittance matrix requires eight matrix-vector products, while for a rational function fitting algorithm, computing one column of the admittance matrix over a range of 50 discrete frequencies would require 140 such products.**

To investigate the crosstalk effects between the package pins in Fig 2, the configuration shown in Fig. 4 is used where 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 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 8^{th} order. 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 describe an accurate approach to using the iterative method in FASTHENRY to compute reduced-order models of frequency-dependent inductance matrices associated with complicated 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,

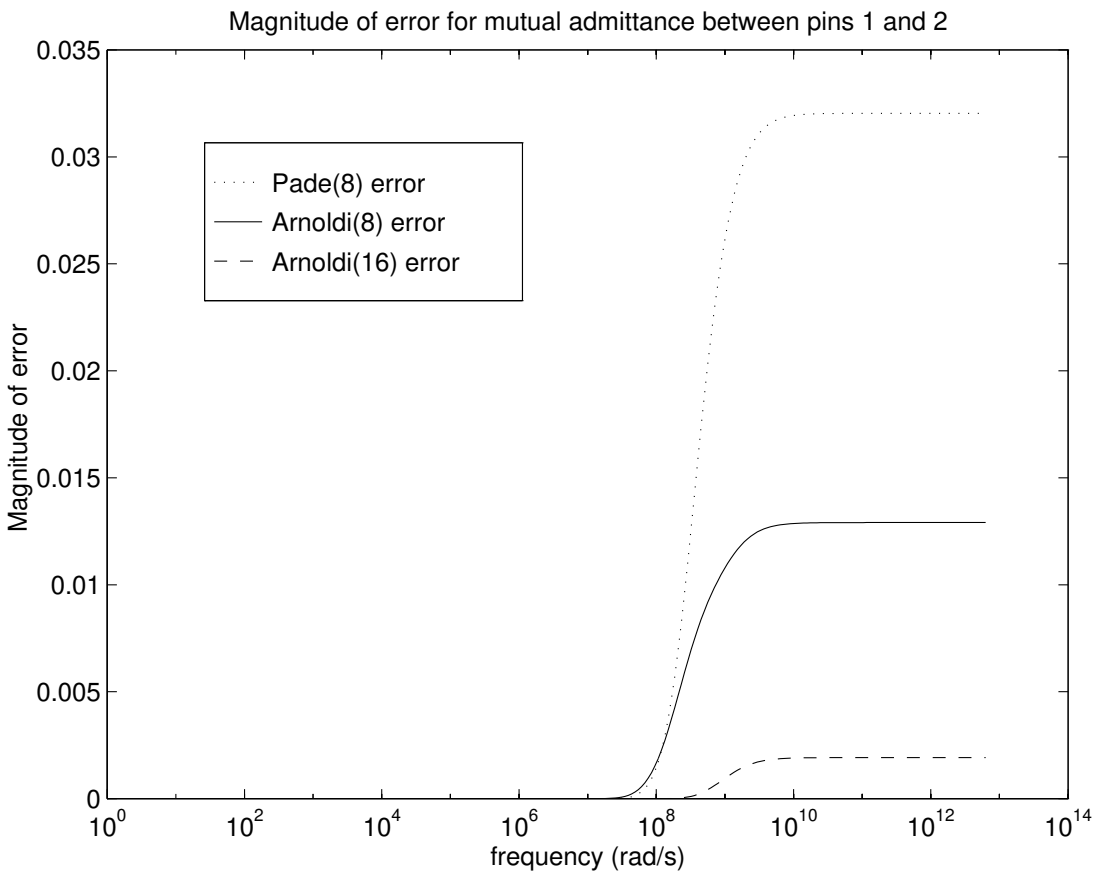

Figure 3: Magnitude of the error for the approximations $\boldsymbol{G}_8^P(s)$, $\boldsymbol{G}_8^A(s)$ and $\boldsymbol{G}_{16}^A(s)$ to the coupled admittance transfer function between pins 1 and 2.

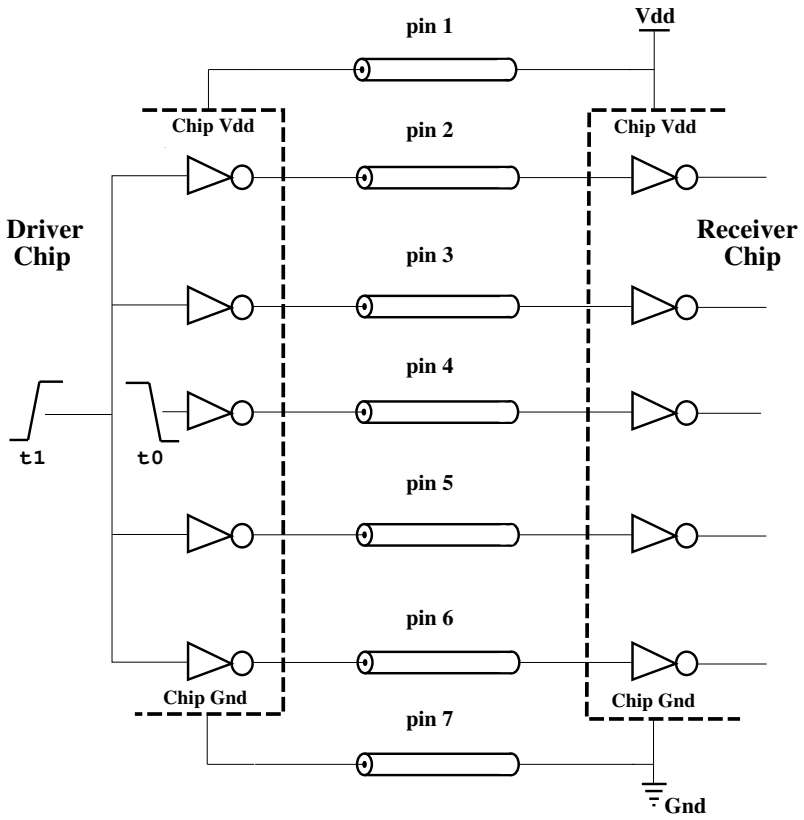

Figure 4: General configuration for the connection between receiver and driver chips. All the circuit elements inside the same chip share that chip's power and ground.

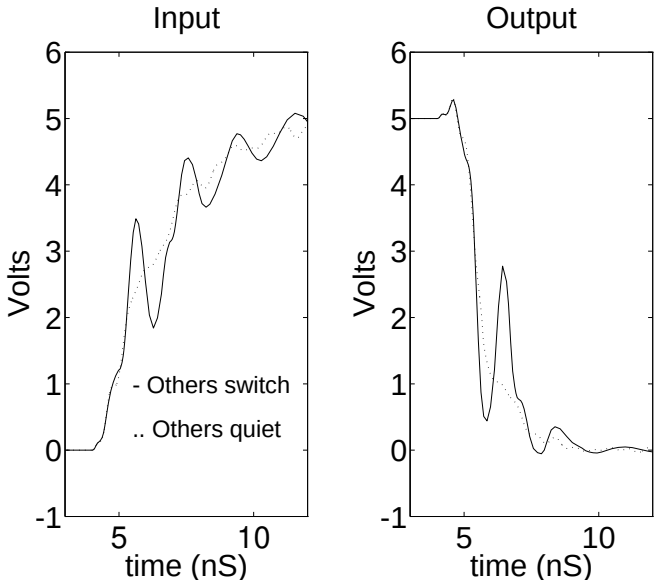

Figure 5: Results of the timing simulation for the output of the receiver gate connected to pin 4 when the adjacent pins switch 1ns after pin 4.

and if matrix-vector product costs dominate then the Arnoldi-based algorithm produces a better approximation for a given amount of work.

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* 44^{st} *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 Euro-DAC*, 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 nonsymmetric 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] Albert E. Ruehli. Equivalent Circuit Models for Three-Dimensional Multiconductor Systems. *IEEE Transactions on Microwave Theory and Techniques*, MTT-22(3):216–221, March 1974.

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

[11] Y. Saad. Variations on Arnoldi's Method for Computing Eigenelements of Large Unsymmetric Matrices. *Linear Algebra and its Applications*, 34:269–295, December 1980.

[12] J. Cullum and R. A. Willoughby. A Practical Procedure for Computing Eigenvalues of Large Sparse Nonsymmetric Matrices. In J. Cullum and R. Willoughby, editors, *Large Scale Eigenvalue Problems: Proceedings of the IBM Europe Institute Workshop on Large Scale Eigenvalue Problems*, pages 193–240. North-Holland, Amsterdam, 1986.