Exploiting Input Information in a Model reduction Algorithm for Massively Coupled Parasitic Networks

L. Miguel Silveira
INESC ID/IST, Cadence Laboratories
Technical University of Lisbon
Lisbon, Portugal
lms@inesc-id.pt

Joel R. Phillips
Cadence Berkeley Laboratories
Cadence Design Systems
San Jose, CA 95134, U.S.A.
jrp@cadence.com

ABSTRACT

In this paper we present a model reduction algorithm that circumvents some of the issues encountered for parasitic networks with large numbers of input/output "ports". Our approach is based on the premise that for such networks, there are typically strong dependencies between the input waveforms at different network "ports". We present an approximate truncated balanced realizations procedure that, by exploiting such correlation information, produces much more compact models compared to standard algorithms such as PRIMA.

Categories & Subject Descriptors: B.7.2 Simulation.
General Terms: Algorithms.
Keywords: Model order reduction, interconnect, parasitic.

1. INTRODUCTION

Model reduction algorithms are the backbone of contemporary parasitic and interconnect modeling technologies. Such algorithms are able to efficiently reduce the size of linear interconnect models without much accuracy degradation and with substantial gains in terms of simulation time. Projection-based Krylov subspace algorithms such as PRIMA [1] and PVL [2] provide a general-purpose, rigorous framework for deriving interconnect modeling algorithms.

Our concern in this paper is with interconnect and parasitic networks having a large number of input/output connections. It is well known that the Krylov-subspace projection based reduction algorithms algorithms are impractical for networks with large numbers of input/output ports. That happens because the cost associated with model computation is directly proportional to the number of inputs, i.e. to the number of columns in the matrices defining the inputs. This is often the case for such "massively coupled" parasitic networks as occur in substrate and package modeling. For example, in the PRIMA algorithm, if only two (block) moments are to be matched at each port, and the network has 1000 ports, the resulting model will have 2000 states, and the reduced system matrices will be dense. This makes simulation in the presence of nonlinear elements impractical.

DAC 2004, June 7–11, 2004, San Diego, California, USA

A related algorithm, often regarded as an alternative for such problems, is the PACT method [3] that relies on eigenvalue analysis via iterative (Lanczos) methods. At low frequencies, PACT can lead to smaller numbers of states than PRIMA, since it does not rely on matching (block) moments. However, PACT still leads to matrices that are dense, and whose size is still bounded from below by the number of ports. At higher frequencies the number of states required can again become large.

In this paper we propose a new reduction algorithm that circumvents some of these issues. Our approach is based on the premise that there are typically strong dependencies between the waveforms at the different inputs to the interconnect network. We start from the viewpoint of truncated balanced realizations (TBR) [4] model reduction. As will shall demonstrate, TBR, as a reduction procedure, is intrinsically somewhat less sensitive to the number of inputs ports. Much more importantly, however, in the TBR framework it is possible to exploit circuit functional information that results in correlations between the waveforms incident on the parasitic network ports. By exploiting this information, an "input-correlated" TBR procedure can be derived that reduces the size of the final models produced.

2. MODEL REDUCTION ALGORITHMS

In this section we will review the most common reduction algorithms for interconnect and parasitic analysis applications. The PRIMA algorithm [1], a Krylov-subspace order reduction procedure, reduces a state-space model, written in the form

$$E\frac{dx}{dt} = Ax + Bu, \quad y = Cx \tag{1}$$

with input waveforms $u(t) \in \mathbb{R}^p$ and output waveforms $y(t)$, by means of a projection matrix V through the operations

$$\hat{E} \equiv V^T EV \quad \hat{B} \equiv V^T B \quad \hat{A} \equiv V^T AV \quad \hat{C} \equiv CV. \tag{2}$$

This leads to the reduced model

$$\hat{A}\frac{dz}{dt} = \hat{A}z + \hat{B}u, \quad y = \hat{C}x. \tag{3}$$

where $z = Vx$. In the standard approach, the V matrix is obtained from a block Krylov subspace. As previously mentioned, the difficulty with these algorithms is that the model size is proportional to the number of moments matched multiplied by the number of ports. For large port numbers (more than 20-30 or so) the algorithms leads necessarily to impractically large models.

An alternative class of reduction algorithms are based on Truncated Balanced Realizations (TBR) [4]. The TBR algorithm first computes the "Gramians" X,Y from the Lyapunov equations

$$AXE^T + EXA^T = -BB^T, \quad (4)$$
$$A^TYE + E^TYA = -C^TC. \quad (5)$$

and then reduces the model by projection onto the space associated with the dominant eigenvalues of the product XY [4] Model size selection and error control in TBR is based on the eigenvalues of XY, the Hankel singular values σ_k. In the proper case, the frequency-domain error in the order k TBR model is bounded by $2\sum_{i=k+1}^{N}\sigma_k$ [5]. Note that the model selection criteria does not depend *directly* on the number of inputs, though, as we shall see, there is an indirect dependence in most problems. In principle, it is possible to have a 1000-port starting model, and obtain a good reduced model of only, say, ten states, if the A,B,C,E matrices are such that all but the the first ten Hankel singular values are small. In the next section we will examine when such a situation might occur.

In practice, solution of the Lyapunov equations (5) is computationally too intensive for large systems as encountered in the type of interconnect networks we are considering here. Therefore, a variety of approximate methods have been proposed [6, 7]. In this work we will utilize one particularly simple method, the PMTBR approach (Poor Man's TBR) [8], which is motivated by an alternative frequency-domain expression for the Gramians:

$$X = \int (j\omega E - A)^{-1}BB^T(j\omega E - A)^{-H}d\omega \quad (6)$$

The PMTBR algorithm works by constructing a matrix Z whose kth column is

$$z_k = (s_kE - A)^{-1}B \quad (7)$$

where s_k is a complex number in the right half-plane. It can be shown that for suitably chosen complex s_k, the singular value decomposition (SVD) of Z, $Z = U\Sigma V^T$, produces a matrix U whose columns approximately span the same space as the dominant eigenspaces of X. U can thus be used as a projection matrix in an approximate TBR procedure.

3. INPUT-CORRELATED TBR

3.1 Input vectors and TBR behavior

To motivate our algorithm, let us consider the impact of the input matrices on the Gramians needed by TBR. For simplicity, consider the case where $A = A^T$ and $E = I$, $B = C^T$. We only need consider one Gramian, given by

$$AX + XA^T = -BB^T. \quad (8)$$

First consider the Hankel singular values for a simple system, such as a uniform RC line, as the number of ports (i.e. columns in the B-matrix) varies. Figure 1 shows the singular values as a function of the number of inputs. Generally speaking, the order needed for good accuracy grows with the number of inputs. This is contrary to the common expectation that a few poles are sufficient for RC systems. For systems with many inputs, many states may be needed because of the high dimension of the controllable space. If low accuracy (10% or so) is acceptable, sometimes models with fairly low numbers of states can be constructed for problems with large numbers of inputs, but this is not always possible even for the restricted case of RC circuits.

Based on these observations, there does not seem to be much hope of producing high-accuracy reduced order models for networks with many ports under general conditions.

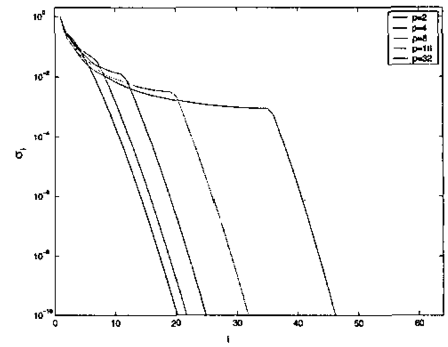

Figure 1: Hankel singular values for 100-segment RC line as function of number of inputs.

3.2 Input-correlated Algorithm

The key to a more efficient procedure lies in noting that in many practical problems, the inputs to an interconnect network are not arbitrary. Often it is necessary to retain all the input ports if the full impact of parasitic effects is to be correctly estimated [9], but there may be relations between the inputs (or outputs) at different network ports that can be exploited to generate a smaller model.

In particular consider a probabilistic model for the network input information. Suppose a correlation matrix [10] for the input relations is known. The appropriate Gramian for this restricted problem is given by

$$AX_c + X_cA^T = -BKB^T \quad (9)$$

where K is the correlation matrix.

The key insight is, for symmetric positive definite K, the eigenvalues of X_c from Eqn. (9) decay faster than the eigenvalues of X from Eqn. (8), if the eigenvalues of K exhibit some decay. In other words, X_c is closer to a low-rank matrix than X if the inputs[1] exhibit some correlated behavior. Postulating existence of correlation is equivalent to saying that we have partial information about the relation between the inputs u. Conversely, in the perfectly uncorrelated case, the eigenvalues of K are identical, which corresponds to zero information. Standard TBR can be viewed as a "zero-input-information" version of the more general input-correlated approach. Thus, for a given truncation criterion for the singular values, using X_c for a model reduction procedure will lead to smaller models. If, *in addition*, K is a suitably representative model of the possible inputs, the model will be equally accurate. Fortunately, this is usually the case in practical problems, and can be guaranteed to occur if we are suitably conservative in the specification of the correlation matrix ($K = I$ again corresponding to the ultimate degree of safety, total ignorance). Note that, though the physical interpretation as an absolute error bound no longer applies, the eigenvalues of the Gramian can still be used for error control, as they can be given an interpretation associated with likelihood of error in a probabilistic input model.

To estimate input correlations, consider taking a set of N samples of input waveforms, u_k for input k, $k = 1 \ldots p$. The correlation matrix can be estimated as

$$K_{ij} = \frac{1}{N}\sum_{l=1}^{N} u_i^l u_j^l \quad (10)$$

[1]That is, the input waveforms $u(t)$ actually applied to the state-space model.

Algorithm 1. **Input Correlated TBR**

1. *Construct the SVD of inputs* $\mathcal{U} = V_K S_K U_K^T$
2. *Do until error satisfactory:*
3. *Draw a vector* $r \in \mathbb{R}^p$ *by taking p draws from a normal distribution, variances given by* Σ_K^2.
4. *Select a frequency point* s_i.
5. *Compute* $z_i = [s_i E - A]^{-1} B U_K r$.
6. *Form the matrix of columns* $Z = [z_1, z_2, \ldots, z_N]$.
7. *Construct the singular value decomposition of Z. If the error is satisfactory, go to Step 8. Otherwise, go to Step 3.*
8. *Construct the projection space V from the orthogonalized column span of Z, dropping columns whose associated singular values fall below a desired tolerance.*

As is the usual case, the actual correlation matrix need not be formed. Instead, we can take the SVD of the matrix $\mathcal{U}$ whose columns are the input samples u_k, i.e.

$$\mathcal{U} = V_K S_K U_K^T \qquad (11)$$

with U_K, V_K orthonormal.

Note that, in addition, from this information we can also obtain estimates of the frequency profile of the inputs. These estimates can be used to select the frequency points s_i for the PMTBR procedure.

We omit the extension to non-self-adjoint systems as this is straightforward. The final algorithm is shown as Algorithm 1.

4. EXAMPLES

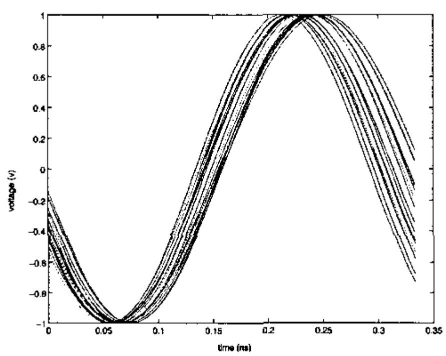

Figure 2: Set of waveform samples for one input on RC network example.

Our first example is a 32-port RC network. Using this example circuit we will illustrate the basic characteristics of the proposed reduction method.

To simulate the situation where there is some degree of information about the relation between input waveforms, we drive the network with a set of sinusoids with fixed, but somewhat uncertain, phase relation. That is, each input is of the form shown in Figure 2: a single sinusoid, but on each input, the set from which the input is drawn has some dither introduced into its phase and frequency. This is intended to mimic the situation where signals incident on the network have some correlation for example because they originate from the same functional block (mixer, oscillator, etc.) or are time-correlated due to a common clock. The dither represents the fact that the signals themselves can be known only approximately before the reduction procedure.

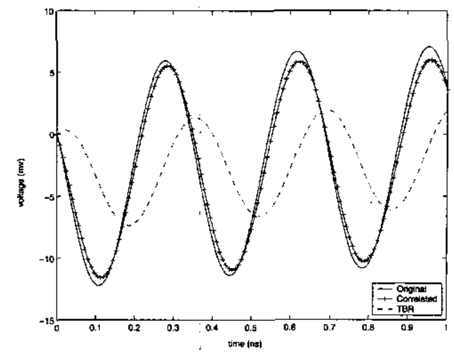

Figure 3: Simulation results for one output on RC network example. PMTBR with correlation information out-performs TBR.

Figure 3 shows results from setting the SVD tolerance to 10^{-3} in Algorithm 1, and extracting a 14-state reduced model. The results from the input-correlated TBR method are quite acceptable. For comparison, we also show the 14-state TBR model: the accuracy of this model is clearly unacceptable. For equivalent accuracy, TBR requires a model with at least 45 states. Note that PRIMA *matching only one moment*, would require a 32-state model and its accuracy would also be fairly low. For this example, PRIMA requires at least two moments for acceptable accuracy, i.e. 64 states. A PACT model incorporating poles up to only the sinusoid frequency would have over seventy states.

Of course, the drawback to the input-correlated procedure is that it is fragile. If the inputs venture far from the distribution assumed when the model was built, accuracy will deteriorate and more states will be required in the model. To illustrate this, we re-ran the same example, again using sinusoids for inputs, but completely changing the phase relation between the inputs (as opposed to the low-level dither introduced in Figure 3). Figure 4 shows the results from the same 14-state models as used previously. The TBR model is about as (in)accurate as previously. However, the accuracy of the input-correlated reduction procedure degrades noticeably. Recovering accuracy requires a model of many more states, so without some degree of information about the input correlation, there is no advantage over using TBR. However, as Figure 1 illustrates, there could still be an advantage over PRIMA.

Finally we consider application of the method to a real circuit (a data converter) with an extracted substrate network. First, for the purpose of assessing the actual error performance of the model reduction algorithm, we extracted only a small portion of the substrate network connection involving the bulk nodes of the MOS transistors. 150 ports of the substrate network were extracted using a boundary-element procedure. Both resistive and capacitive terms were retained, leading to a 150-state model. To obtain estimates of the signal correlations at the inputs of the parasitic network we simulate the nonlinear circuit *without* the substrate network and measure the values of the MOS transistor bulk current signals. We then carry those measurements as inputs to the input-correlated TBR

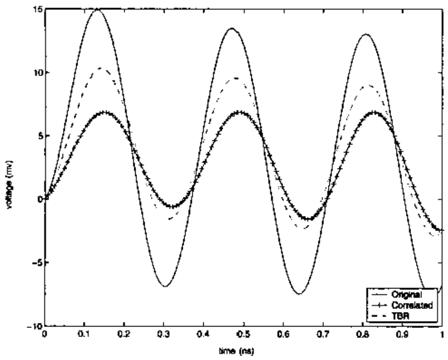

Figure 4: Simulation results for one output on RC network example, with re-randomized phase relation. PMTBR with correlation information breaks down.

procedure [2]. Using Algorithm 1, a reduced model is produced. We then compare the results of simulation with the reduced model to simulation with the full substrate model. These results are shown in Figure 5 for a representative node. In this case, fair agreement with the full model was obtained using only four states, and excellent agreement is obtained with eight states. Similar accuracy is obtained at all ports of the substrate network. We point out that this is a 20X compression from the full model. Note also that this network is, for most intents, unreducible with standard projection methods.

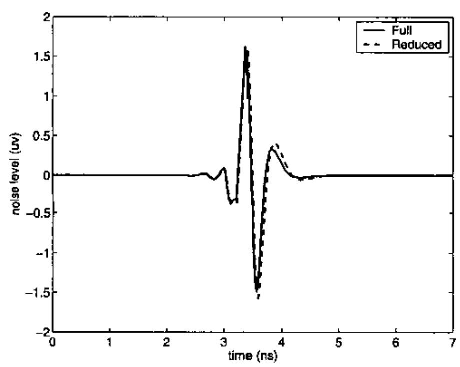

Figure 5: Simulation results for data converter example, 150 port substrate models, full vs. 4-state reduced model.

To illustrate the capabilities of the algorithm on larger networks, we also applied the proposed technique to a larger section of the extracted substrate network, this time comprising 1000 substrate ports. Figure 6 shows the error estimate data obtained from the singular value analysis in Algorithm 1. In this case a model size of 30 states is sufficient to achieve high accuracy. This represents a compression of over 30X in model size and, because of the superlinear complexity associated with factorizing dense matrix blocks, considerably more savings in time required for linear system solution in simulation.

[2]Note that, should the substrate network result in such *large* changes to the circuit operation that these estimates were completely unrepresentative, we would have to iterate this procedure to obtain a self-consistent estimate. This would probably indicate that the circuit ceased to function as designed.

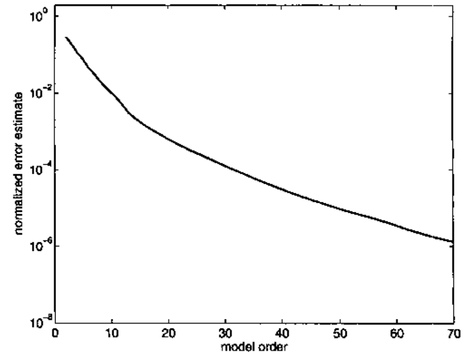

Figure 6: Error estimate based on singular value analysis of Z-matrix from input-correlated TBR, for 1000-port substrate network with inputs from data converter example.

5. CONCLUSIONS

In this work we demonstrated that exploiting input information, such as from nominal circuit function, can help reduce the size of parasitic models obtained from projection-like procedures. This is particularly relevant for problems with a large number of inputs which are known not to reduce efficiently under such methods. We introduced an input-correlated TBR-like procedure to perform the computation of the reduced model. When there is strong correlation between input waveforms on different input ports, large reductions in model size can be achieved. In many practical settings this is a common situation since spatial and temporal dependencies dictated by the circuit topology and functionality will tend to highly correlate the signals seen at the ports of the interconnect networks.

6. REFERENCES

[1] A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aided Design*, 17(8):645–654, August 1998.

[2] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 14(5):639–649, May 1995.

[3] K. J. Kerns and A. T. Yang. Stable and efficient reduction of large, multiport rc networks by pole analysis via congruence transformations. In *IEEE/ACM Design Automation Conference*, pages 280 – 285, Las Vegas, NV, U.S.A., June 1996.

[4] Bruce Moore. Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction. *IEEE Transactions on Automatic Control*, AC-26(1):17–32, February 1981.

[5] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L_∞ error bounds. *International Journal of Control*, 36:1115–1193, 1984.

[6] Jing-Rebecca Li, F. Wang, and J. White. An efficient lyapunov equation-based approach for generating reduced-order models of interconnect. In 36^{th} *ACM/IEEE Design Automation Conference*, pages 1–6, New Orleans, Louisiana, June 1999.

[7] I. M. Jaimoukha and E. M. Kasenally. Oblique projection methods for large scale model reduction. *SIAM J. Matrix Anal. Appl.*, 16:602–627, 1995.

[8] Joel R. Phillips and L. Miguel Silveira. Poor man's tbr: A simple model reduction scheme. In *DATE'2004 - Design, Automation and Test in Europe, Exhibition and Conference*, Paris, France, February 2004.

[9] Joel R. Phillips and L. Miguel Silveira. Simulation approaches for strongly coupled interconnect systems. In *International Conference on Computer Aided-Design*, pages 430–437, November 2001.

[10] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw Hill, New York, 1991.