Model Reduction in the Time-Domain using Laguerre Polynomials and Krylov Methods *

Y. Chen, V. Balakrishnan, C-K. Koh and K. Roy
ECE, Purdue University
West Lafayette, IN 47907–1285
Email: {yc, ragu, chengkok,kaushik}@ecn.purdue.edu

Abstract

We present a new passive model reduction algorithm based on the Laguerre expansion of the time response of interconnect networks. We derive expressions for the Laguerre coefficient matrices that minimize a weighted square of the approximation error, and show how these matrices can be computed efficiently using Krylov subspace methods. We discuss the connections between our method and other methods such as PRIMA [4]. Numerical simulations show that our method can better approximate the original model as compared to PRIMA.

1 Introduction

With the increasing complexity of VLSI circuits, the models of interconnect networks have become so large that their analysis using brute-force techniques has become impractical. A popular approach to solve this problem is model reduction, where an approximate small-scale model is used in place of the original large-scale system.

A number of model-reduction methods are available in the literature. Moment matching techniques [1, 9, 4] can be used to generate passive reduced order models for interconnects. More sophisticated methods such as balance-and-truncate model reduction have also been proposed for interconnect modeling [5, 3, 6]. Krylov methods such as the Arnoldi and Lanczos iterations apply naturally in the numerical implementation of these techniques. All of these methods focus on approximating the frequency response of the original system and are hence referred to as frequency-domain methods.

A natural counterpart to frequency-domain methods are time-domain methods where the focus is on directly approximating the time response of the systemi (it can be argued that this is a more natural approach , as typically it is the time response that is of interest). Krylov methods have been applied in this context as well; see [2], where the derivatives of the circuit response are preserved to a given order for the reduced-order model. Another approach towards approximating time responses uses orthogonal polynomials. The basic idea is to project the time response of the original system onto a smaller dimensional subspace spanned by an orthogonal basis. A natural optimality criterion in this setting is the (possibly weighted) square of the approximation error. An example of such an approach can be found in [8], where Chebyshev polynomials are used to approximate the impulse response.

The algorithm that we present in this paper are closest in spirit to the approach in [8]. Motivated by a number of reasons, we propose the use of Laguerre polynomials for the orthogonal basis. Unlike with Chebyshev polynomials that are defined over the interval $[-1,1]$, Laguerre polynomials are defined over $[0,\infty]$, making them a more natural candidate for approximating continuous-time responses. In addition, we will show that a closed-form expression for the optimal coefficients can be derived very simply. These coefficients can be *numerically computed* very efficiently using Krylov methods. And the amount of computation required by our methods is the same as that required by PRIMA. We will also show that our algorithm can handle more general inputs, and can accommodate differing time-horizons for the approximation.

2 Background

2.1 Problem Description

Consider an single-input, single-output[1] RLC network, described by the state equations:

$$M\frac{dx}{dt} + Nx(t) = bu(t), \quad y(t) = Cx(t) \tag{1}$$

where $x(t) \in \mathbf{R}^n$, $u(t)$ and $y(t)$ are the state, input, and output respectively. M and N are square matrices of size n, and $b \in \mathbf{R}^n$. These matrices arise from modified nodal analysis (MNA). Typically n is very large.

*This work was supported in part by NSF under contract number CCR-9984553 and SRC under contract number 99-TJ-689.

[1] We consider single-input, single-output systems only for simplicity of exposition; the results presented herein can be extended to multi-input multi-output systems without much difficulty.

With $A = M^{-1}N$ and $B = M^{-1}b$, the impulse response of system (1), given by $x(t) = e^{-At}B$, can be shown to satisfy

$$x(t) + A\int_0^t x(\tau)d\tau = B. \tag{2}$$

2.2 The Laguerre Polynomial Expansion

Consider the sequence of polynomials $L_i : [0,\infty) \to \mathbf{R}$ $i = 0, 1, \ldots$, given by

$$L_i(t) = e^t \frac{d^i}{dt^i}(t^i e^{-t}). \tag{3}$$

These are the *Laguerre* polynomials [7]. The Laguerre expansion suggests a natural method for approximating the function f: The kth order *Laguerre approximant* of f is simply

$$\hat{f}_k(t) = \hat{c}_0 L_1(t) + \hat{c}_1 L_1(t) + \cdots + \hat{c}_k L_k(t). \tag{4}$$

It is easy to establish that this approximation is *optimal* in the following sense: With F_k defined as

$$F_k = \left\{ f \;\middle|\; \begin{array}{c} f(t) = c_0 L_1(t) + \cdots + c_k L_k(t) \\ c_i \in \mathbf{R},\ i = 0, 1, \ldots, k \end{array} \right\},$$

we have

$$\hat{f}_k = \mathrm{argmin}_{f_k \in F_k} \int_0^\infty e^{-t}(f(t) - f_k(t))^2\, dt.$$

In other words, $\hat{f}_k(t)$ given by (4) is the *optimal* projection of $f(t)$ into the k-dimensional subspace F_k, with the error weighted by e^{-t}.

The time weighting in the Laguerre expansion has a natural engineering interpretation: Errors in the approximation of the response that are closer in time to the impulsive excitation are considered more "costly" than later errors. As the transition of a signal (from logic high to logic low or vice versa) in any useful VLSI circuitry typically occurs shortly after the input excitation, such a time weighting function enables the Laguerre approximation to capture signal transition accurately. Therefore, our argument is that for VLSI applications, the Laguerre approximation provides better approximations of signal delay and transition times.

3 Laguerre model reduction

3.1 Approximating the impulse response

Our objective is to approximate the impulse response $x(t) = e^{-At}B$ of system (1) as

$$x(t) \approx x_k(t) = \sum_{i=0}^{k} \hat{c}_i L_i(t) \tag{5}$$

Using (2), and the fact that the Laguerre polynomials satisfy the recursion

$$\int_0^t L_i(t)d\tau = -\frac{1}{i+1}L_{i+1}(t) + L_i(t), \tag{6}$$

it is easy to show that we have

$$L \begin{bmatrix} \hat{c}_k \\ \vdots \\ \hat{c}_1 \\ \hat{c}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B \end{bmatrix}, \tag{7}$$

where

$$L = \begin{bmatrix} I+A & -\frac{1}{k}A & & & & \\ & I+A & -\frac{1}{k-1}A & & & \\ & & \cdots & \cdots & & \\ & & & I+A & -\frac{1}{2}A & \\ & & & & I+A & -A \\ & & & & & I+A \end{bmatrix}. \tag{8}$$

Thus, we have the following recursions to obtain the coefficients $\hat{c}_i$:

$$\begin{array}{c} \hat{c}_0 = (I+A)^{-1}B, \\ \hat{c}_i = \frac{1}{i}(I+A)^{-1}A\hat{c}_{i-1}, \quad 1 \le i \le k. \end{array} \tag{9}$$

3.2 Model reduction

We now present a model order reduction scheme that builds on the observation that $x_k(t)$, the kth order Laguerre polynomial, serves as an optimal kth order approximant to $x(t)$. Note that with $\tilde{B} = (I+A)^{-1}B$ and $\tilde{A} = (I+A)^{-1}A$, we may write

$$x_k(t) = P_k \begin{bmatrix} \frac{1}{0!}L_0(t) \\ \frac{1}{1!}L_1(t) \\ \vdots \\ \frac{1}{k!}L_k(t) \end{bmatrix},$$

where

$$P_k = \begin{bmatrix} \tilde{B} & \tilde{A}\tilde{B} & \cdots & \tilde{A}^k\tilde{B} \end{bmatrix}. \tag{10}$$

Thus, $x_k(t)$ lies in the span of the columns of the matrix P_k for all t. This suggests the model reduction scheme of projecting the state-space on to the span of the columns of P_k. More precisely, let $P_k = Q_k R_k$ be a QR factorization, that is, $Q_k \in \mathbf{R}^{n\times(k+1)}$ with orthonormal columns, $R_k \in \mathbf{R}^{(k+1)\times(k+1)}$ is lower triangular. Then, a reduced-order model is given by

$$\hat{M}\frac{d\hat{x}}{dt} + \hat{N}\hat{x}(t) = \hat{b}u(t), \quad \hat{y}(t) = \hat{C}\hat{x}(t) \tag{11}$$

where $x(t) = Q_k\hat{x}(t)$, $\hat{M} = Q_k^T M Q_k$, $\hat{N} = Q_k^T M Q_k$, $\hat{b} = Q_k^T b$ and $\hat{C} = CQ_k$.

Obviously the reduced-order model described in (11) is passive. This follows from the observation that the columns

of Q_k are orthogonal. The proof is in [4].

3.3 Efficient computation

As the columns of P_k span a Krylov subspace, its QR factorization can be efficiently computed as follows:

1. Compute an LU factorization $M+N$. Use this to compute $\tilde{B} = (I+A)^{-1}B = (M+N)^{-1}b$.
2. Compute the QR factorization of P_k using the Arnoldi method, without explicitly forming $\tilde{A}$.

The overall computation required is essentially the same as that required by PRIMA [4].

4 Extensions

4.1 Accounting for the time horizon

Recall that with the standard Laguerre approximation, the error was weighted by e^{-t}; this introduces implicitly a time horizon of interest. This time horizon can be increased or decreased with a simple time-weighting technique, by replacing t by t/T. (Loosely speaking, increasing T "spreads out" the quality of approximation over larger and larger time horizons.) This corresponds to replacing A by AT and B by BT in the development thus far, that is, with

$$\tilde{B}_T = \left(\frac{1}{T}I + A\right)^{-1} B \text{ and } \tilde{A}_T = \left(\frac{1}{T}I + A\right)^{-1} A,$$

the kth order Laguerre approximant lies in the span of

$$\left[\begin{array}{cccc} \tilde{B}_T & \tilde{A}_T\tilde{B}_T & \cdots & \tilde{A}_T^k\tilde{B}_T \end{array}\right].$$

When T is very large compared to the magnitude of the eigenvalues of A, we have

$$\tilde{B}_T \approx A^{-1}B, \quad \tilde{A}_T \approx \left(I - \frac{1}{T}A^{-1}\right),$$

so that the kth order Laguerre approximant lies in

$$span\left[\begin{array}{ccccc} A^{-1}B & A^{-2}B & \cdots & A^{-K+1}B & A^{-K}B \end{array}\right],$$

which coincides with the matrix used by PRIMA to perform the model reduction.

In our experience, we have found that a value of T that is one-fifth of the reciprocal of the smallest real part of the eigenvalues of A is a good heuristic for trading off the quality of approximation of the time response and the time horizon.

4.2 Response to general inputs

Unlike with existing approximation methods such as those based on Chebyshev polynomials, we show how our approach can directly handle inputs other than impulses. Suppose the input u is a polynomial of order $k-1$. Then, we may write

$$\int_0^t u(\tau)d\tau = a_0L_0(t) + \cdots + a_kL_k(t),$$

that is, we first compute the Laguerre expansion of the integral of $u(t)$. (Note that the expansion is exact.) Then, it is readily verified that the optimal kth order Laguerre approximant of the state x is given by $x(t) \approx x_k(t) = \sum_{i=0}^{k} \hat{c}_i L_i(t)$, where

$$L\begin{bmatrix} \hat{c}_k \\ \hat{c}_{k-1} \\ \vdots \\ \hat{c}_1 \\ \hat{c}_0 \end{bmatrix} = \begin{bmatrix} a_kB \\ a_{k-1}B \\ \vdots \\ a_0B \end{bmatrix}, \tag{12}$$

with L being defined in (8). In other words, the optimal Laguerre approximant is obtained by simply changing the right-hand side of the linear equation (7). One consequence of this observation is that the reduced-order model derived in §3.2 is independent of the input when it is a polynomial of order smaller than that of the reduced-order model.

5 Numerical Results

We present a few representative circuit examples that demonstrate the performance of the Laguerre model reduction algorithm proposed in this paper. We consider three circuit examples that are either adapted or taken directly from [4, 8].

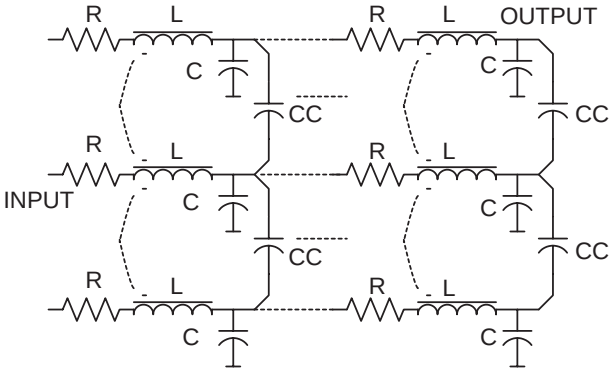

Figure 1. Circuit Example 1 [8].

Example 1: The circuit, which is shown in Figure 1, consists of three RLC lines with capacitive and inductive coupling between adjacent lines [8]. There are 20 RLC sections in each line. The circuit parameters for each section are: $R = 1\Omega$, $L = 1$nH, $C = 1$pF, the coupling capacitance $CC = 1$pF, and the inductance coupling coefficient between adjacent sections is $k = 0.2$. The middle line is the aggressor net with a step input. Figure 2 plots the exact response of the far-end node of the upper victim line. In the same plot, we also show the approximation responses of the Laguerre and PRIMA reduced order models, both of which are

of order 13. For a more direct comparison between the Laguerre model reduction method and the PRIMA method, we also plot the absolute differences (errors) between the exact response and the responses of the two reduced order models in Figure 3. The peak errors of the Laguerre model and the PRIMA model, denoted respectively by E_{pL} and E_{pP}, are 0.1104V and 0.1113V. The integrals of the errors for the Laguerre model and PRIMA model over the span of the time responses in Figure 2, denoted respectively by E_{iL} and E_{iP}, are $1.36e^{-10}$V-s (Volt-second) and $1.71e^{-10}$V-s. It is evident that the response of the the Laguerre model tracks the exact response more closely as compared to PRIMA.

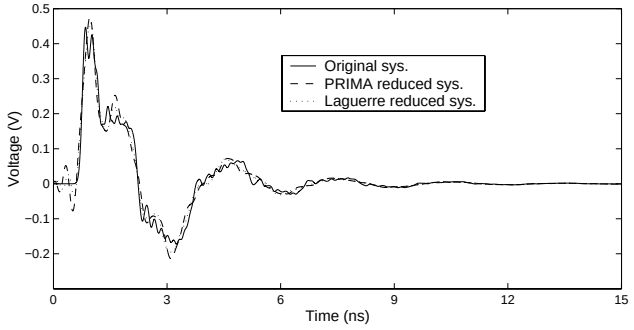

Figure 2. Time responses of Example 1 and its reduced models.

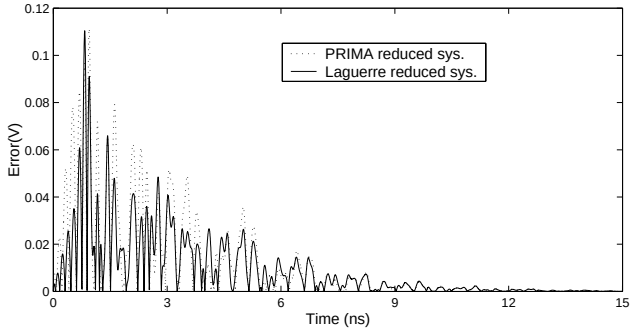

Figure 3. Approximation errors of the reduced models for Example 1.

We perform similar comparisons of the Laguerre model reduction method and the PRIMA method for the remaining two examples. For each circuit, we compare the exact response with the responses of the reduced order models generated by the Laguerre model reduction method and the PRIMA method. The peak errors and the integral errors incurred by the two methods are summarized in Table 1. The Laguerre model reduction method has smaller peak and integral errors as compared to PRIMA. The details of the remaining two examples are as follows:

Example 2: The circuit, which is adapted from [8], is shown in Figure 4. Each line is segmented into 40 sections with parameters $R = 1\Omega$, $L = 0.01$nH, $C = 10$fF, $G = 0.001\mho$, and the coupling capacitance $CC = 1$fF. The order of the two

Table 1. Peak errors and integral errors of Laguerre and PRIMA models.

Example	1	2	3
$E_{iL} \times 10^{10}$ (V-s)	1.36	0.0316	0.132
$E_{iP} \times 10^{10}$ (V-s)	1.71	0.0404	0.159
E_{iL}/E_{iP}	0.792	0.782	0.832
E_{pL}(V)	0.1104	0.508	0.0160
E_{pP}(V)	0.1113	0.700	0.0171
E_{pL}/E_{pP}	0.992	0.726	0.936

reduced models is 15. The input is a current source. The responses of the output node of the upper line are shown in Figure 5, and the approximation errors of the Laguerre and PRIMA models are shown in Figure 6.

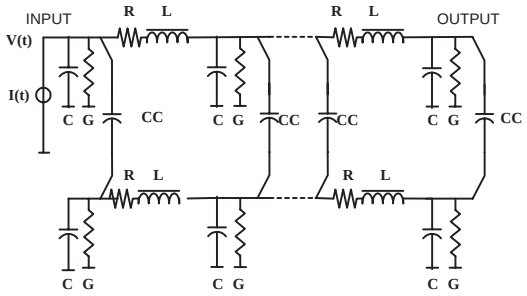

Figure 4. Circuit Example 2 [8].

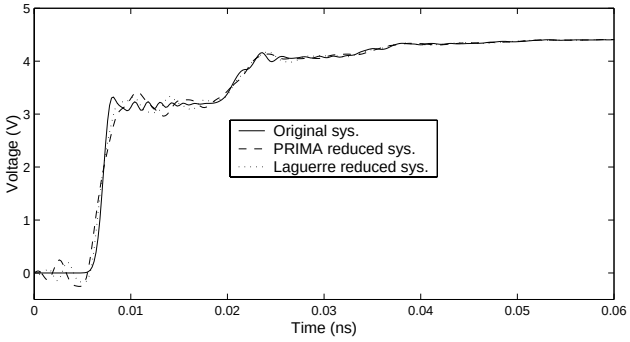

Figure 5. Time responses of Example 2 and its reduced models.

Example 3: This example is borrowed from [4]. The circuit is a 2-bit bus driven by CMOS inverters as shown in Figure 7. One of the drivers is switching while the other is quiet. The two wires, consisting of 40 coupled RLC sections each, are modeled as a four-ports system. Each wire has a resistance, inductance, and capacitance (to ground) of 10Ω, 5nH, and 15pF, respectively. The total coupling ca-

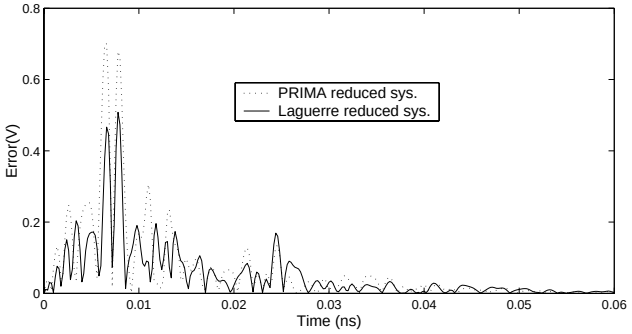

Figure 6. Approximation errors of the reduced models for Example 2.

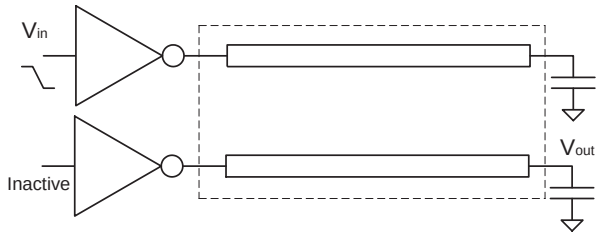

Figure 7. Circuit Example 3 [4].

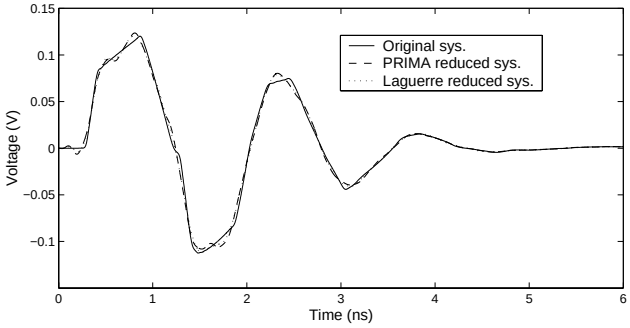

Figure 8. Time responses of Example 3 and its reduced models.

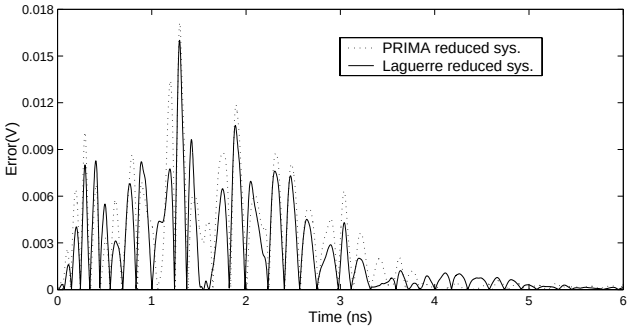

Figure 9. Approximation errors of the reduced models for Example 3.

pacitance between the two lines is 7.5pF. The order of the reduced models is 10. The input to the upper wire is a ramp input with a rise time of 1ns. The waveforms of the output of the lower wire are shown in Figure 8 and the approximation errors are shown in Figure 9.

6 Conclusion

We have presented a new passive model reduction algorithm for large-scale systems based on the Laguerre expansion of the time response. Three distinguishing features of our method are: (i) We use Laguerre expansion to approximate the continuous-time response of the original state equation over $[0,\infty)$. (ii) The Laguerre approximation minimizes a weighted square of the approximation error. (iii) We derive a closed-form expression for the optimal Laguerre coefficient matrices, whose orthonormal bases can be numerically computed very efficiently via Krylov methods. Moreover, we introduce a time-weighting technique to further reduce the error over a time horizon of interest. We also explore the connection between the proposed Laguerre approximation and the PRIMA method. Numerical simulations show that our approach can obtain better results than the PRIMA method.

References

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

[2] P. K. Gunupudi and M. S. Nakhla. Model-reduction of nonlinear circuit using krylov-space techniques. In *Proc. Design Automation Conf*, pages 13–16, June 1999.

[3] J. Li, F. Wang, and J. white. An efficient lyapunov equation-cased approach for generating reduced order model of interconnect. In *Proc. Design Automation Conf*, pages 1–6, June 1999.

[4] A. Odabasioglu, M. Celik, and L. Pileggi. Prima: Passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 17(8):645–654, Aug. 1998.

[5] P. Rabiei and M. Pedram. Model order reduction of large circuit using balanced truncation. In *Proc. Asia South Pacific Design Automation Conf.*, volume 1, pages 237–240, Jan. 1999.

[6] Q. Su, V. Balakrishnan, and C.-K. Koh. Efficient balance-and-truncate model reduction for large-scale system. In *Proc. American Control Conf*, June 2001.

[7] G. Szegö. *Orthogonal Polynomials*. American Mathematical Society, 1939.

[8] J. M. Wang, Q. Yu, and E. S. Kuh. Passive model order reduction algorithm based on chebyshev expansion of impulse response of interconnect networks. In *Proc. Design Automation Conf*, pages 520–525, June 2000.

[9] Q. Yu, J. Wang, and E. Kuh. Passive multipoint moment matching model order reduction algorithm on multiport distributed interconnect networks. *IEEE Trans. on Circuits and Systems*, 46(1):140–160, Jan. 1999.