Model Reduction of Time-Varying Linear Systems Using Approximate Multipoint Krylov-subspace Projectors

Joel R. Phillips
Cadence Design Systems, San Jose, CA 95134
jrp@cadence.com

Abstract

In this paper a method is presented for model reduction of systems described by time-varying differential-algebraic equations. This method allows automated extraction of reduced models for nonlinear RF blocks, such as mixers and filters, that have a near-linear signal path but may contain strongly nonlinear time-varying components. The models have the accuracy of a transistor-level nonlinear simulation but are very compact and so can be used in system-level simulation and design. The model reduction procedure is based on a multipoint rational approximation algorithm formed by orthogonal projection of the original time-varying linear system into an approximate Krylov subspace. The models obtained from the approximate Krylov-subspace projector can be obtained much more easily than the exact projectors but show negligible difference in accuracy.

1 Overview

The problem of automated macromodel generation is interesting from the viewpoint of system-level design because if small, accurate reduced-order models of system component blocks can be extracted, then much larger portions of a design, or more complicated systems, can be simulated or verified than if the analysis were to have to proceed at a detailed level. The prospect of generating the reduced model from a detailed analysis of component blocks is attractive because then the influence of second-order device effects or parasitic components on the overall system performance can be assessed. In this way overly conservative design specifications can be avoided.

There has been considerable recent interest, primarily in the context of simulation of electrical interconnect, in extracting low-order models of lumped (often passive) components that are time-invariant. There are many systems, however, that are not linear time-invariant (LTI) but can be accurately modeled as linear time-varying (LTV). For example, if a nonlinear circuit model is linearized around a time-varying large signal, the resulting model is linear time-varying. In particular, many RF components (e.g., mixers and filters) are designed to have a near-linear response in the signal path, but may have strongly nonlinear response to other excitations, such as the clock of a switched-capacitor filter, or a mixer's local oscillator. Such components are prime candidates for LTV model reduction. Clearly, the set of circuits that can be accurately modeled as linear-time varying is much larger than the set that can be described as LTI.

Most of the recent work on LTI model reduction has been based, implicitly or explicitly, on *projection*-based formulations.[4, 11] The reduced model is obtained from the full model by projecting the linear system into a subspace of lower dimension. The subspace chosen determines the approximation properties of the reduced model.

It is now generally accepted that in LTI systems, choosing the projection subspaces to be Krylov subspaces is effective and efficient. Efficiency arises because the Krylov subspaces are easily computed. The effectiveness of the approach is motivated by noting that projecting into a Krylov subspace corresponds to matching derivatives of the Laplace-domain transfer function (the moments). Methods based on multipoint rational approximations [11, 22, 5, 2, 9] are known to be particularly efficient.

Unfortunately, however, model reduction for time-varying systems appears to have received little attention. Balanced truncation approaches[30] have been proposed, but it is unclear how to implement these techniques effectively. Therefore, given the wealth of knowledge that has developed about model reduction using rational approximations and the projection formulations, the first task of this paper is to ask if the LTV systems can be analyzed in this framework. As is shown in Section 4, this is indeed the case. Once the projection framework has been identified as a viable candidate for model reduction, the next task is to determine appropriate subspaces and how to compute them efficiently.

Since our work is motivated by RF problems, periodically linear time-varying systems (PLTV) are of particular interest and our implementation and application examples focus on these systems. In Section 5 we show that models based on approximate multipoint Krylov spaces can be very efficiently extracted from a time-domain RF circuit simulator. The methodology, however, should generalize to arbitrary time waveforms.

2 Model Reduction for LTI Systems

In hope of gaining insight into possible approaches to time-varying systems, in this section we review the essentials of the Krylov-subspace based projection methods.

Consider the linear time-invariant multi-input, multi-output (MIMO) linear system written in differential-algebraic form[1]

$$C\dot{x}(t) = -Gx + Bu(t)$$

$$y(t) = D^T x$$

where $C, G \in R^{n\times n}; x(t) \in R^n, B \in R^{n\times n_i}, D \in R^{n\times n_o}, u(t) \in R^{n_i}, y(t) \in R^{n_o}$, n is the system order, and n_i and n_o are the number of system inputs and outputs respectively.

For simplicity assume for the moment that $C = I$ where I is the identity matrix. After performing a Laplace transformation,

[1] We omit any direct input-output feed as these terms do not enter the model reduction procedure.

the system output is $y(s) = D^T(sI+G)^{-1}Bu(s)$. The transfer function $D^T(sI+G)^{-1}B$ is a rational function in s, so it seems logical to approximate it with a rational function, such as a Padé approximant[1]. Pade approximants, and most of the other approximants used for model reduction, have the property that they match the transfer function and some of its derivatives with respect to s.

In the general case a rational approximant can be obtained by

$$\hat{C}\dot{z}(t) = -\hat{G}z + \hat{B}u(t)$$

$$y(t) = \hat{D}^T z$$

where $\hat{C}, \hat{G} \in R^{r\times r}; z(t) \in R^r, \hat{B} \in R^{r\times n_i}, \hat{D} \in R^{r\times n_o}$, and $r \ll n$ if the reduction is useful. The reduced matrices are obtained from the projection matrices L, T by

$$\hat{G} = L^T GT, \hat{C} = L^T CT, \hat{B} = L^T B, \hat{D} = T^T D \qquad (1)$$

Returning to the case $C = I$ (the following results generalize), note that the kth derivative, or *moment*, of the transfer function is given by $D^T G^{-(k+1)} B$. Clearly the approximants we wish to generate are connected with powers of the matrix G^{-1} acting on B, or of G^{-T} acting on D. Connecting the moments to the projection matrices L, T is the key to the model reduction procedure. These ideas are formalized by the following definition and theorem.

Definition 1 (Krylov subspace) *The Krylov subspace $K_m(A,p)$ generated by a matrix A and vector p, of order m, is the space spanned by the set of vectors $\{p, Ap, A^2p, \ldots, A^{m-1}p\}$.*

Theorem 1 (Krylov Subspace Approximation) *If the columns of L span the order m Krylov subspace $K_m(G^{-T}, D)$ and the columns of T span the order n Krylov subspace $K_m(G^{-1}, B)$, then the reduced order transfer function $\hat{D}^T(sI+\hat{G})^{-1}\hat{B}$ matches the first $m+n$ moments of the unreduced function $D^T(sI+G)^{-1}B$.*

Proof. See [4, 11]. □

For example, in the PVL algorithm[7], which forms a Pade approximation and is thus equivalent to the AWE[18] technique because of the relation between Lanczos algorithm and Pade approximants [10], the choice of L and T is $L^T = W^T G^{-1}, T = V$, where W and V contain the biorthogonal Lanczos vectors, and in one variant of model reduction based on the Arnoldi method[24], $L^T = W^T G^{-1}, T = V$ where V is the orthonormal matrix generated by the Arnoldi process. The columns of W span $K_m(G^{-T}, D)$ and the columns of V span $K_m(G^{-1}, B)$.

In [16, 5], the projection matrices are taken directly from the Krylov basis itself, $L = T = V$. Because of the orthogonal projection[2], if the full model possesses desirable structural properties, such as stability and passivity, they are inherited by the reduced models. See Grimme[11] for further discussion of projection-based approaches.

This approach can be extended to the case of multipoint approximants where the transfer function and some of its derivatives are matched at several points in the complex plane.[11, 22, 5, 2, 9] In this case L and T must contain a basis for the union of the Krylov subspaces constructed at the different expansion points. When the expansion point is complex, real models may be efficiently obtained by exploiting the fact that for a real matrix A, if $u = (I - sA)^{-1}p$ is in the Krylov space, then so is u^*[22]. Having established approximation properties for LTI model reduction schemes, we turn to the time-varying case.

[2] An orthogonal projection results whenever $L = T$ and implies nothing about the orthogonality of L or T, though constructing L and/or T to have orthonormal columns can be a useful from a computational standpoint.

3 Time-Varying Small-Signal Analysis

We begin our analysis by describing one way in which linear time-varying systems may arise. To obtain a linear time-varying circuit description, first the differential equations describing the circuit are written (using, say, modified nodal analysis [12]), as

$$f(v^{(T)}) + \frac{d}{dt}q(v^{(T)}) = u^{(T)}(t) \qquad (2)$$

where u represents the input sources, $v^{(T)}$ describes the node voltages, f is the relation between voltages and currents, and the function q relates voltages to charges (or fluxes). b is a vector that describes the mapping from the input function u to the system internals.

We have written the voltage v and input variable u with the superscript T to indicate that they are total quantities that we will split into two parts, a large-signal part and a small signal part, in order to obtain a TVL model,

$$u^{(T)} = u^{(L)} + u, \qquad v^{(T)} = v^{(L)} + v. \qquad (3)$$

By linearizing around $v^{(L)}$, a linear time-varying system of the form

$$G(t)v + \frac{d}{dt}\left(C(t)v\right) = bu(t) \qquad (4)$$

where $G(t) = \partial f(v^{(L)})(t)/\partial v$ and $C(t) = \partial q(v^{(L)})(t)/\partial v$ are the time-varying conductance and capacitance matrices, is obtained for the small response v.

3.1 Analyzing the linear time-varying system

In what follows script ($\mathcal{A}$) will be used to denote continuous operators and upper case (A) to denote a member of the space of $n \times n$ complex matrices, $\mathbf{M}_n$. Subscripts will denote a vector or matrix variable at a particular timepoint or harmonic, and lower case unsubscripted will denote a variable over a time or frequency index. That is, for a system with N states represented with M (discrete) degrees of time-freedom (timepoints or Fourier harmonics), then $A \in \mathbf{M}_{NM}$, $A_t \in \mathbf{M}_N$, $x_t \in \mathbf{R}^N$, $x \in \mathbf{R}^{MN}$.

In model reduction of LTI systems, most progress has been made in expoiting rational approximations to the frequency-domain transfer functions. Thus motivated, we adopt the formalism of Zadeh's variable transfer functions[31] that were developed to describe time-varying systems. In this formalism the response $v(t)$ can be written as an inverse Fourier transform of the product of a time-varying transfer function and the Fourier transform of $u(t)$, $u(w)$. That is,

$$v(t) = \int_{\infty}^{-\infty} h(i\omega', t)u(\omega')e^{i\omega' t}d\omega' \qquad (5)$$

To obtain the frequency-by-frequency response, we let u be a single-frequency input, $u_{\omega'} = u_\omega \delta(\omega - \omega')$, and see that

$$v(t) = h(i\omega, t)u(\omega)e^{i\omega t} \qquad (6)$$

Writing $s = i\omega$ and substituting into Equation 4, an equation for the $h(s,t)$ is obtained,

$$G(t)h(s,t) + \frac{d}{dt}\left(C(t)h(s,t)\right) + sC(t)h(s,t) = b \qquad (7)$$

Defining

$$\mathcal{K} = G(t) + \frac{d}{dt}C(t) \qquad \mathcal{C} = C(t) \qquad (8)$$

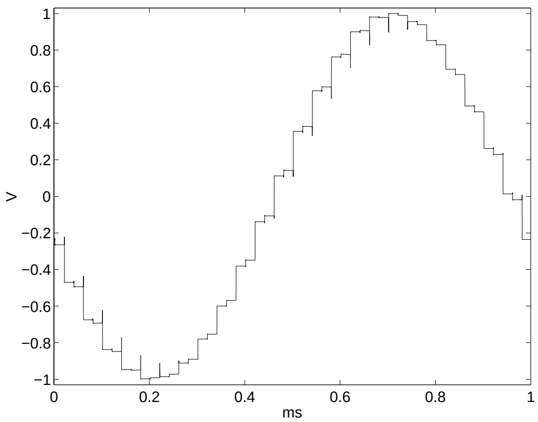

Figure 1: Response of the switched-capacitor filter to a small 1kHz sinusoid.

this may be written more compactly as

$$[\mathcal{K} + s\mathcal{C}]\, h(s,t) = b. \tag{9}$$

This expression may also be obtained from the finite-difference formulations[15, 27] in the limit as the timestep goes to zero, or from the multivariate-partial-differential-equation formalism[20, 19].

Equation 9 has a form similar to the frequency-domain expressions of LTI transfer functions. However, it involves continuous operators instead of finite-dimensional matrices. Finite-dimensional representations will be obtained in Section 4 below. In general, since $h(s,t)$ comes from a lumped linear system, it will be a rational function with an infinite number of poles [3]. For example, in a periodically time-varying system with fundamental frequency ω_0, if η is a pole (in particular, a Floquet multiplier) of the system, then $\eta + k\omega_0$, k an integer, will be a pole of $h(s,t)$. This is because signals can be converted by harmonic k in the time-varying description from a frequency $\eta - k\omega_0$ to the pole at η.

3.2 Systems concepts

The formalism above completely describes the internal system response to an arbitrary outside input. However, we have not described how the detailed differential-equation description is related to outside systems. In the LTI context this role is played by the D and B matrices.

Time-varying systems differ from LTI ones also in that there is considerable flexibility as to the choice of the possible input-output mapping functions, and thus of the transfer function itself. In the LTI case, the model ports are usually fixed in advance, and that is the end of the story. Specification of a constant matrix pair (D, B) is sufficient to describe the input-output mappings. In the time-varying case, this is no longer sufficient. We must specify D and B matrices, but it is useful to allow these matrices to vary with time.

To see this, consider a switched-capacitor filter problem. A five-pole low pass switched capacitor filter, containing 71 MOSFETs was simulated and the time-varying linear response to a 1kHz sinusoid computed. The results are shown in Figure 1. The jagged waveform shape is a result of the strong nonlinearity of the filter with respect to the clock. Incorporation of such effects is the point of time-varying modelling. An LTI system would produce a smooth sinusoid, shifted in phase and scaled in magnitude, as a response to this input. The output of the filter, however, is usually followed by some sort of function (e.g., the sample/hold of an A/D converter) that discards the filter output outside some small sample time window. To model this system at a higher level, we would need a reduced-order model that relates continuous, sinusoidal inputs to the window of output that the following circuitry (the sample and hold) needs.

In general, in the PTVL case, we must specify $(D(t), B(t))$ over a fundamental period. Fortunately in our case there is not quite as much freedom as it first appears. In circuit problems, since the output ports are fixed, we may write $D(t)$ and $B(t)$ as $D(t) = D[d_1(t), \ldots, d_{n_i}(t)], B(t) = B[b_1(t), \ldots, b_{n_o}(t)]$ for some periodic scalar functions $d_i(t), b_i(t)$. The most common choice, as in in [26, 27], and for most harmonic balance codes, is to choose the $d(t), b(t)$ to be single-tone sinusoids. In some cases, for example mixers, this is a natural choice, but in general cases, the choices of $d(t), b(t)$ are not as simple. If many harmonics are of interest in the input or output mappings, then the choice of pure tones is a poor one. In the case of the switched-capacitor filter, a natural transfer function is obtained by choosing a $d(t)$ that samples the output in a narrow interval after it has settled.

4 Model Reduction for LTV Systems

4.1 Obtaining Discrete Rational Functions

Because, for the lumped time-varying systems considered here, the time-varying transfer functions are rational functions, it seems reasonable to believe that reduced models can be obtained from the same sorts of rational approximation paths that have been so profitable for reduction of LTI systems. Therefore, we first seek a representation of the transfer functions in terms of finite-dimensional matrices.

To obtain a discrete rational matrix function, we must discretize the operators $\mathcal{K}$ and $\mathcal{C}$.[3] Because the focus of this paper is on PTVL systems that occur in RF applications, at this point we also introduce explicit assumptions about the time-variation of the system, and explain how the input-output mappings are incorporated into the model reduction procedure. We will only discuss the SISO case as the generalization is simple.

Following[26, 27], in the example case of a backward-Euler discretization, we have

$$(K + sC)h(s) = b \tag{10}$$

with

$$K = \begin{bmatrix} \frac{C_1}{h_1} + G_1 & & & -\frac{C_M}{h_1} \\ -\frac{C_1}{h_2} & \frac{C_2}{h_2} + G_2 & & \\ & \ddots & \ddots & \\ & & -\frac{C_{M-1}}{h_M} & \frac{C_M}{h_M} + G_M \end{bmatrix} \tag{11}$$

$$C = \begin{bmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_M \end{bmatrix} \tag{12}$$

$$h(s) = \begin{bmatrix} h_1(s) & h_2(s) & \cdots & h_M(s) \end{bmatrix}^T \tag{13}$$

and

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_M \end{bmatrix}^T \tag{14}$$

[3]Of course, this step is *also* a model reduction by projection operation, though not usually an orthogonal one, as most codes use some form of collocation (BDF or pseudo-spectral discretizations). In general the model reduction may proceed directly from any representation of the operator.

where $G_j = G(t_j), C_j = C(t_j), b_j = b(t_j), h_j(s) = h(s, t_j)$. In principal, at this point, any of the algorithms developed for reduction of lumped LTI systems can be applied to the matrices and vectors defined in Equations 11-14, if the results are properly interpreted. Note that since the input functions or basis vectors v_k represent time-waveforms, the reduced input and output functions $\hat{b}$ and $\hat{d}$ will represent time-varying input and output mappings. An example will be presented in Section 7

4.2 Preserving System Structure

Recently there has been considerable interest in developing model reduction methods for passive LTI systems that preserve the system passivity.[16, 14] Unlike in the lumped RLC case, the time-varying models considered in this paper are not necessarily passive, or even stable, since they come from a linearization of a nonlinear system. Indeed, small-signal gain may be a desired property of the model. However, even if no a-priori conditions can be established for the passivity or lack thereof in the time-varying systems, we would at least like to insure that whatever nice structural properties, in particular stability or passivity, happen to be possessed by the underlying system, are not destroyed in the model reduction procedure. When an orthogonal projection is performed using orthonormal matrices as proposed in this paper, it can be shown using arguments based on the *field of values* of an operator[13] that properties such as passivity are inherited in the reduced model.

5 Approximating the Krylov space

In [27], the transfer functions from a small-signal sinusoidal input to sinudoids at harmonics of the output were obtained by the solving the finite-difference equations

$$\begin{bmatrix} \frac{C_1}{h_1}+G_1 & & & -\frac{C_M}{h_1}\cdot\alpha(s) \\ -\frac{C_1}{h_2} & \frac{C_2}{h_2}+G_2 & & \\ & \ddots & \ddots & \\ & & -\frac{C_{M-1}}{h_M} & \frac{C_M}{h_M}+G_M \end{bmatrix} \begin{bmatrix} \tilde{v}(t_1) \\ \tilde{v}(t_2) \\ \vdots \\ \tilde{v}(t_M) \end{bmatrix} = \begin{bmatrix} \tilde{b}(t_1) \\ \tilde{b}(t_2) \\ \vdots \\ \tilde{b}(t_M) \end{bmatrix} \tag{15}$$

where $\alpha(s) \equiv e^{-sT}$ and T is the fundamental period. The transfer function $h(s,t)$ is then given by $h(s,t) = e^{-st}\tilde{v}(t)$.

It is now convenient to decompose K into a lower triangular and an upper triangular piece, $K = L + R$. Using the expressions for L and R, (15) becomes

$$\left(L + \alpha(s)R\right)\ \tilde{v}\ =\ \tilde{b}(s) \tag{16}$$

If we define a small-signal modulation operator $\mathbf{\Omega}(s)$,

$$\mathbf{\Omega}(s) \equiv \begin{bmatrix} Ie^{st_1} & & & \\ 0 & Ie^{st_2} & & \\ & \ddots & \ddots & \\ & & 0 & Ie^{st_M} \end{bmatrix} \tag{17}$$

then we can make the identification

$$h(s) = \mathbf{\Omega}^H(s)\tilde{v}(s) \tag{18}$$

and more importantly,

$$K + sC \simeq \mathbf{\Omega}(s)\left[L + \alpha(s)R\right]\mathbf{\Omega}^H(s) \tag{19}$$

The left and right hand sides of Eq. 19 differ in the treatment of the small signal. The left hand side represents a spectral discretization, and the right-hand-side represents a finite-difference discretization.

Suppose we need to solve Eq. 16 for some right-hand-side $\tilde{b}$. Again following [27], consider preconditioning with the matrix L. Because L is lower triangular, with a small block bandwidth, block Gaussian elimination is very efficient at computing the inverse acting on a vector. In this procedure, once the M diagonal blocks have been factored, an operation that must be prformed exactly once, then every application of the inverse is an M step procedure, at each step needing a backsolve with the factored diagonal matrices and multiplication by the blocks off the diagonal. There is one off-diagonal block in each row for the simple backward-Euler discretization. The preconditioned system can be written as

$$\left(I + \alpha(s)L^{-1}R\right)\ \tilde{v}\ =\ L^{-1}\tilde{b}(s) \tag{20}$$

Suppose that Equation 20 is solved by a Krylov-subspace based iterative method such as GMRES[23].[4] Because the Krylov subspace of a matrix A is invariant to shifts[17, 27] of the form $A \rightarrow A + \beta I$, the same Krylov subspace may be used to solve Equation 20 at multiple frequency points. This "recycled Krylov subspace" algorithm is made even more efficient[27] by exploiting the special structural properties of $L^{-1}R$, and because the spectrum of $L^{-1}R$, being related to the the Floquet multipliers of the TVL system, is usually clustered. The recycling of the Krylov space also accelerates solution with different right-hand-sides. The final point is that systems of the form (16) may be solved very efficiently for different frequencies and right-hand-side vectors. However, the finite-difference formulation is not suited to model reduction.

In contrast, the spectral form that is amenable to model reduction is less convenient to work with (contrast Equation 10 and Equation 16). Even if we use a lower-tridiagonal preconditioner, at each different frequency point the preconditioner must be reconstructed (i.e. we must re-factor the diagonal blocks), and the resulting Krylov-subspaces are not preserved with shifts in frequency s.

To resolve this dilemna consider what would happen if the matrix V used for the projection was not a basis for the Krylov subspace of $(K + sC)^{-1}$, but instead a nearby matrix. Because of the way the reduced model is formed, only small errors would be introduced into the final model. As long as the model was not evaluated in the vicinity of a pole, the additional errors introduced into the model would be small.

This suggests that the basis for the projector in the model-reduction procedure be obtained by using the finite-difference equations. Because the basis will be a very good approximation to the Krylov-subspaces of the spectral operator, good reduced models should still be obtained. In addition, because of the recycled Krylov scheme, obtaining projectors from expansions about multiple frequency points is essentially no more expensive than single-frequency-point expansions. This leads us to the proposed model reduction algorithm, Algorithm 1. It may be considered an adaptation of the algorithm of [5] where approximate linear solves have been used instead of exact ones.

The recycled Krylov-subspace iterative solvers, and block solvers in general, are more efficient than the non-block solvers only if the sequence of right-hand-sides have Krylov spaces that are 'near' one another. To motivate why recycling might be very effective for the model reduction problem, consider the simple case where $C = I$ and the expansion point is taken at the origin. Then in the exact case the kth order projector V_k is constructed so that

$$V_k \subset K_k(K^{-1}, b) \equiv \{b, K^{-1}b, K^{-2}b, \ldots, K^{-(k-1)}b\}. \tag{21}$$

[4] These are different Krylov-subspaces than the ones used for model reduction.

Algorithm I
(*Approximate Multipoint Krylov-subspace Model Reduction*)

Set $k = 1$.
for $i = 1, \ldots, n_q$ {
  for $j = 1, \ldots, m(i)$ {
    if $j = 1$ then
      $w = b$
    else
      $w = Cv_{k-1}$
    $u = \Omega^H(s_i)\left[L + \alpha(s_i)R\right]^{-1}\Omega(s_i)w$
    for $l = 1, \ldots, k-1$ {
      $u = u - v_l^T u$
    }
    $v_k = u/||u||$
    $k = k + 1$
  }
}
$\hat{K} = V^T K V$
$\hat{C} = V^T C V$
$\hat{D} = V^T D$
$\hat{B} = V^T B$

Suppose that instead of the preconditioned procedure proposed above, each new vector in the model reduction was obtained by an inner Krylov iteration with the matrix K. Then, intuitively, since each new right-hand-side u_i in the model reduction procedure is drawn from a Krylov space of (K, p, m) for some m, it seems reasonable to expect that the next term in the space of $K_i(K^{-1}, b)$ would be related to the $K_m(K, b)$ and that some efficiency could be obtained thereby. In fact Van der Vorst[29] has shown that the vector $K^{-q}b$ can be constructed from nearly the same Krylov sequence $\{b, Kb, \ldots K^m b\}$, $m > q$, that is used to construct $K^{-1}b$. Thus is is reasonable to expect that the entire V_k can be constructed from the space $\{b, Kb, \ldots K^m b\}$, where m only slightly exceeds k.

For our model reduction implementation, a recycled version of the GMRES algorithm[8] is used as a 'black box' to solve the equations $L + \alpha(s_i)v_i = p_i$. As a result these nice properties of the Krylov spaces are automatically exploited without having to worry about the details of the preconditioning, shifting, etc. Recent work on 'thickly restarted' methods[25] for eigenvalue computation exploit similar ideas.

6 Applications

To test the model reduction procedure, the proposed algorithms were implemented in a time-domain RF circuit simulator. The large-signal periodic-steady-state is calculated using a shooting method[26]. The time-varying linear system was discretized using variable-timestep second-order backward-difference formulas.

The first example considered was the switched-capacitor filter previously discussed, running at a clock frequency of 25kHz. This example generated 58 equations in the circuit simulator, and 453 timesteps were needed to describe the steady-state waveform. The same timesteps generated in the solution of the periodic-steady-state problem were used to discretize the time-varying linear operator. For the model reduction procedure, the input function $b(t)$ (see

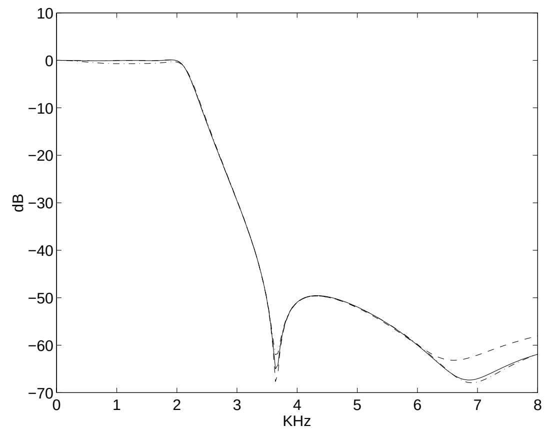

Figure 2: The time-sampled switched-capacitor filter transfer function.

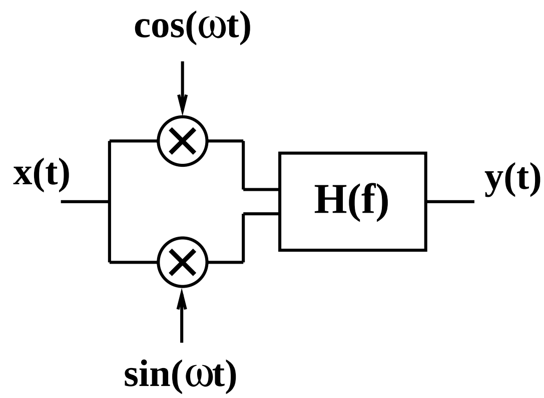

Figure 3: The receiver macromodel. The frequency of the sin and cosine elements is the mixer LO frequency, 780MHz.

Equation 10) was constant, corresponding to the continuous sinusoidal input present at the filter input. To specify an output function, we took a sample function that was constant over a 200ns period $1us$ before the clock edge at the start of the cycle. Essentially, the final model is a real LTI system that represents a tranfer function between the continuous analog input and the sampled digital output. The amplitude of the transfer function, as a function of input frequency, of the reduced model is shown in Figure 2.

Two nine-state models are shown in Figure 2. The model shown with a dashed line was generated by matching nine real moments at the origin. The dash-dot line, virtually identical to the actual transfer function, was generated from matching 3 real moments at the origin, and one moment at 200kHz, 400kHz, and 800kHz on the imaginary axis. As these expansion points were off the real axis, each complex moment in the Krylov space generates two states in the final real model, corresponding to the Krylov vector and its complex conjugate. The multi-point approximation is seen to be a better match.

The second example is the complex image-rejection receiver studied in [28]. This receiver is a complicated circuit with several functional component blocks (a low-noise amplifier, a splitting network, two double-balanced mixers, and two broad-band Hilbert transform output filters). The entire circuit has 167 bipolar transistors and generates 986 equations in the circuit simulator. 200

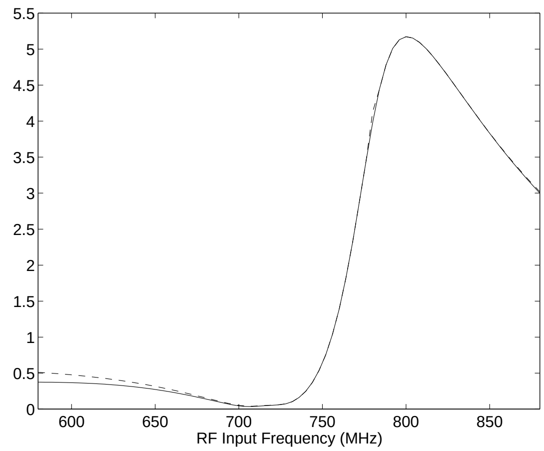

Figure 4: Transfer function of the mixer (solid) and the 15-state mixer model(dash), from RF input to mixer output.

timesteps were needed for the time-domain analysis, so that the matrix K has a rank of almost 200,000.

A fifteenth order real-valued time-varying model was generated to represent the receiver conversion path from RF to output. Since a model from multiple sidebands to mixer output was desired, the adjoint matrix K^T was used to generate the model reduction. In this case the time-varying elements appear before a LTI filter in the final model,[5] as is shown schematically in Figure 3. The mixing elements shift the input from the RF frequency by 780MHz, the mixer LO frequency. Following these elements is a multi-input LTI filter whose response is shown in Figure 4. The lower-sideband rejection characteristic of this mixer is evident in the model. If the model were used in a suppressed-carrier DSP simulator, the mixing elements would simply be omitted.

Table 1 shows the statistics for the computational costs for the reduced model extraction and evaluation. 200 frequency points were considered in the filter example and fifty in the mixer example. In both examples the reduced model took less time to extract than the single frequency sweep, and the evaluation was vastly more efficient (in fact the overhead in the code was sufficient that it was difficult to determine exactly how much time was consumed in the actual model evaluation). Note the efficiency in particular of the reduction of the switched-capacitor example. The time-varying model has a rank of 26,274, yet the reduced model was generated in only 7 CPU seconds.

We did not compare the effectiveness of the exact moment-matching procedure. This approach requires a new complex matrix factorization for multipoint approximants and is very inefficient if many separate frequency points are considered. However, the efficiency of the recycled Krylov algorithm is evident. In the case of the filter, for a order-nine real model only 18 applications of L^{-1} were required in the recycled GMRES procedure that obtained the moments. The remaining matrix vector products (or backsolves) were needed to perform the projection or for the initial preconditioning steps. Even in the mixer example, where GMRES had much more difficulty converging, 124 backsolves were needed for the reduction, about 8 per model order, which is still good.

[5] Basing the model reduction on K is more convenient if we wish to place time-varying elements after the LTI component.

Circuit		MOR/Recycle	MOR/Std	Pointwise
SCF	MVP	45	99	410
	Reduce	7s	12s	–
	Solve	1s	1s	1.5m
Receiver	MVP	166	1980	280
	Reduce	4m	48m	–
	Solve	2s	2s	8.5m

Table 1: Comparison of time-varying model reduction procedures and pointwise frequency sweeps. The frequency sweeps were accelerated by the recycled GMRES algorithm. MVP refers to equivalent real-real matrix vector solves with the matrix L. "Reduce" is the model reduction time in seconds or minutes, and "solve" the CPU time required to obtain the frequency response.

7 Conclusions and Future Work

In this paper we have demonstrated extraction of simple and compact macromodels from nonlinear, time-varying transistor-level circuit descriptions. As the switched-capacitor and receiver examples demonstrate, these models are capable of representing very complicated underlying dynamics. The class of systems that can be represented as linear time-varying is very broad and so the approach proposed here is potentially a powerful analysis and abstraction tool. The use of the multi-point Krylov-subspace based rational approximant with the inner approximate recycle GMRES solver appears to be a particularly synergistic combination and is primarily responsible for the efficiency of the method.

In fact, we have observed that an advantage of the multi-point approximants is that they seem to be more robust in their ability to use approximant solves of very low accuracy, when coupled with the recycled GMRES solver. When the requested solution tolerance from the recycled GMRES solver is set very loosely, it will fairly quickly cease to generate new vectors for use in the model reduction projection space. Moving to another frequency expansion point introduces new information into the recycled solver, and new directions are generated for the model reduction projection.

There are several immediate extensions of the ideas of this paper. The formalism and algorithms can be trivially extended to the case of quasi-periodic small-signal analysis. Similarly the methodology of this paper can be directly applied to obtain reduced models of (poly)cyclostationary noise transfer functions[21], as has essentially already been shown in [6].

References

[1] George A. Baker Jr. and Peter Graves-Morris. *Padé Approximants Part I: Basic Theory*. Encyclopedia of Mathematics and its Applications. Addison-Wesley Publishing Company, Reading, MA, First edition, 1981.

[2] Eli Chiprout and Michel S. Nakhla. Analysis of interconnect networks using complex frequency hopping (CFH). *IEEE Trans. CAD*, 14:186–200, February 1995.

[3] Henry D'Angelo. *Linear Time-Varying Systems*. Allyn and Bacon, Boston, 1970.

[4] C. de Villemagne and R. E. Skelton. Model reduction using a projection formulation. *Int. J. Control*, 46:2141–2169, 1987.

[5] Ibrahim Elfadel and David D. Ling. A block rational Arnoldi algorithm for multipoint passive model-order reduction of

multiport rlc networks. In *International Conference on Computer Aided-Design*, pages 66–71, San Jose, California, November 1997.

[6] P. Feldmann and R. Freund. Circuit noise evaluation by Padé approximation based model-reduction techniques. In *International Conference on Computer Aided-Design*, pages 132–138, San Jose, California, November 1997.

[7] Peter Feldmann and Roland W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Trans. CAD*, 14:639–649, May 1995.

[8] D. Feng. unpublished.

[9] K. Gallivan, E. Grimme, and P. Van Dooren. Multi-point padé approximants of large-scale systems via a two-sided rational krylov algorithm. In 33^{rd} *IEEE Conference on Decision and Control*, Lake Buena Vista, FL, December 1994.

[10] William B. Gragg. Matrix interpretations and applications of the continued fraction algorithm. *Rocky Mountain Journal of Mathematics*, 4(2):213–225, 1974.

[11] Eric J. Grimme. *Krylov projection methods for model reduction*. PhD thesis, University of Illinois at Urbana-Champaign, 1997.

[12] Chung-Wen Ho, Albert E. Ruehli, and Pierce A. Brennan. The modified nodal approach to network analysis. *IEEE Transactions on Circuits and Systems*, 22(6):504–509, June 1975.

[13] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1995.

[14] K. J. Kerns, I. L. Wemple, and A. T. Yang. Stable and efficient reduction of substrate model networks using congruence transforms. In *IEEE/ACM International Conference on Computer Aided Design*, pages 207 – 214, San Jose, CA, November 1995.

[15] H. Tanimoto M. Okumura, T. Itakura, and T. Sugawara. Numerical noise analysis for nonlinear circuits with a periodic large signal excitation including cyclostationary noise sources. *IEEE Trans. Circuits and Systems - I*, 40:581–590, 1993.

[16] Altan Odabasioglu, Mustafa Celik, and Lawrence Pileggi. Prima: Passive reduced-order interconnect macromodeling algorithm. In *International Conference on Computer Aided-Design*, San Jose, California, November 1997.

[17] B. N. Parlett. *The Symmetric Eigenvalue Problem*. Prentice Hall, 1980.

[18] Lawrence T. Pillage and Ronald A. Rohrer. Asymptotic Waveform Evaluation for Timing Analysis. *IEEE Transactions on Computer-Aided Design*, 9(4):352–366, April 1990.

[19] J. Roychowdhury. Efficient methods for simulating highly nonlinear multirate circuits. In *Proceedings of the* 34^{th} *Design Automation Conference*, pages 269–274, Anaheim, CA, June 1997.

[20] J. Roychowdhury. Mpde methods for efficient analysis of wireless systems. In *Proceedings of the 1998 Custom Integrated Circuits Conference*, pages 451–454, Santa Clara, CA, Mayy 1998.

[21] J. Roychowdhury, D. Long, and P. Feldmann. Cyclostationary noise analysis of large RF circuits with multitone excitations. *IEEE Journal Solid-State Circuits*, 33:324–336, March 1998.

[22] A. Ruhe. The rational Krylov algorithm for nonsymmetric eigenvalue problems III: complex shifts for real matrices. *BIT*, 34:165–176, 1994.

[23] Youcef Saad and Martin Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, July 1986.

[24] L. Miguel Silveira, Mattan Kamon, and Jacob K. White. Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-d interconnect structures. In *Proceedings of the* 32^{nd} *Design Automation Conference*, pages 376–380, San Francisco, CA, June 1995.

[25] A. Stathopolous, Y. Saad, and K. Wu. Dynamic thick restarting of the Davidson, and the implicitly restarted Arnoldi methods. *SIAM Journal on Scientific Computing*, 19:227–245, 1998.

[26] R. Telichevesky, J. White, and K. Kundert. Efficient steady-state analysis based on matrix-free krylov-subspace methods. In *Proceedings of 32rd Design Automation Conference*, June 1995.

[27] R. Telichevesky, J. White, and K. Kundert. Efficient AC and noise analysis of two-tone RF circuits. In *Proceedings of 33rd Design Automation Conference*, June 1996.

[28] R. Telichevesky, J. White, and K. Kundert. Receiver characterization using periodic small-signal analysis. In *International Custom Integrated Circuits Conference*, May 1996.

[29] H. A. van der Vorst. An iterative solution method for solving $f(A)x = b$ using Krylov subspace information obtained for the symmetric matrix A. *J. Comp. App. Math.*, 18:249–263, 1987.

[30] E. I. Verriest and T. Kailath. On generalized balanced realizations. *IEEE Transactions on Automatic Control*, 28:833–844, 1983.

[31] L. Zadeh. Frequency analysis of variable networks. *Proc. I.R.E.*, pages 291–299, March 1950.