Krylov Subspace Techniques for Reduced-Order Modeling of Nonlinear Dynamical Systems

Zhaojun Bai
Department of Computer Science
and Department of Mathematics
University of California
Davis, CA 95616, USA

Daniel Skoogh
Swedish Defense Research Agency
Division of System Technology
Department of Autonomous Systems
172 90 Stockholm, Sweden

Abstract

Means of applying Krylov subspace techniques for adaptively extracting accurate reduced-order models of large-scale nonlinear dynamical systems is a relatively open problem. There has been much current interest in developing such techniques. We focus on a bi-linearization method, which extends Krylov subspace techniques for linear systems. In this approach, the nonlinear system is first approximated by a bilinear system through Carleman bilinearization. Then a reduced-order bilinear system is constructed in such a way that it matches certain number of multimoments corresponding to the first few kernels of the Volterra-Wiener representation of the bilinear system. It is shown that the two-sided Krylov subspace technique matches significant more number of multimoments than the corresponding one-side technique.

1 Introduction

Several model reduction techniques for nonlinear dynamical systems have been studied by researchers in various fields. Two of the most well-known methods are the Karhunen-Loève decomposition based methods and methods of balanced truncation. Karhunen-Loève decomposition based methods are also known as proper orthogonal decomposition (POD) methods. Methods of balanced truncation extend the success of balanced truncation of linear systems to nonlinear systems. The interested reader is referred to [6] and [14]. The latest work includes [7] and [11]. Means of applying Krylov subspace techniques for adaptively extracting accurate reduced-order models of large-scale nonlinear dynamical systems is a relatively open problem. There has been much current interest in developing such techniques. We will briefly discuss two methods, which extend Krylov subspace techniques for linear dynamical systems.

We consider single-input single-output nonlinear dynamical systems of the form:

$$\left\{ \begin{array}{rcl} \dot{\mathbf{x}} & = & \mathbf{f}(\mathbf{x}) + \mathbf{b}u, \\ y & = & \mathbf{l}^{\mathrm{T}}\mathbf{x} \end{array} \right. \tag{1.1}$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$, where $\mathbf{x} \in \mathcal{R}^N$ is the state variables, N is the dimension of the state space. $u \in \mathcal{R}$ and $y \in \mathcal{R}$ are inputs and outputs, respectively. $\mathbf{b} \in \mathcal{R}^N$ is the input distribution array. $\mathbf{l} \in \mathcal{R}^N$ is the output measurement array. We assume that the nonlinear state evolution function $\mathbf{f}(\mathbf{x}) : \mathcal{R}^N \to \mathcal{R}^N$ is smooth, i.e., C^∞, and has an equilibrium. Without loss of generality we take this equilibrium at $\mathbf{0}$, i.e., $\mathbf{f}(\mathbf{0}) = \mathbf{0}$.

Examples of the origins of nonlinear dynamical systems of the form (1.1) include the simulation of time-varying nonlinear circuit elements by independent excitation source [4, 3], and MEMS, such as

micro-pressure sensor [8]. The modeling of the dynamical behavior of a voltage-controlled parallel-plate electrostatic actuator also derives a set of state equations of the form (1.1) [15, p.138]. Such an electrostatic actuator invokes multi-domain parameters, such as mass, stiffness and damping in the mechanical domain, and an excitation force network in the electrical domain.

2 Linearization method

We will discuss two methods for the reduced-order modeling of the nonlinear system (1.1). The first method is called the *linearization method*. It linearizes the system around the equilibrium point, and then extracts a Krylov subspace for reduced-order modeling. Specifically, suppose that the power series expansion of $\mathbf{f}(\mathbf{x})$ about the equilibrium point $\mathbf{0}$ is written as

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}_1\mathbf{x} + \mathbf{A}_2(\mathbf{x}\otimes\mathbf{x}) + \mathbf{A}_3(\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}) + \cdots \tag{2.2}$$

where $\mathbf{A}_1 \in \mathcal{R}^{N\times N}$ is the Jacobian or the first derivative of $\mathbf{f}$, and $\mathbf{A}_2 \in \mathcal{R}^{N\times N^2}$ is the second derivative matrix of $\mathbf{f}$, and so on. $\otimes$ is the Kronecker product. We linearize the original nonlinear system (1.1) by only using the first term in the expansion (2.2) of $\mathbf{f}$, and obtain a linear system:

$$\left\{\begin{array}{rcl} \dot{\widehat{\mathbf{x}}} &=& \mathbf{A}_1\widehat{\mathbf{x}} + \mathbf{b}u, \\ \widetilde{y} &=& \mathbf{l}^{\mathrm{T}}\widehat{\mathbf{x}}. \end{array}\right. \tag{2.3}$$

We can then apply a reduced-order modeling method for the linearized system (2.3), and obtain a *linear* reduced-order model. The output $\widetilde{y}$ is an approximation of the output y of the original system (1.1). If we are interested in a small region of the state space near the equilibrium point, or so-called small-signal analysis, then as demonstrated in [4], this approach provides an efficient tool for analyzing the nonlinear system (1.1).

Alternatively, one may also use the linearized model (2.3) to extract a Krylov projection subspace spanned by $\mathbf{V}_n$. Then, by substituting $\mathbf{x} \approx \mathbf{V}_n\mathbf{z}$ into the original nonlinear system (1.1), a *nonlinear* reduced-order model is obtained:

$$\left\{\begin{array}{rcl} \dot{\mathbf{z}} &=& \mathbf{g}(\mathbf{z}) + \mathbf{b}_n u, \\ \hat{y} &=& \mathbf{l}_n^{\mathrm{T}}\mathbf{z} \end{array}\right.$$

where $\mathbf{g}(\mathbf{z}) = \mathbf{V}_n^{\mathrm{T}}\mathbf{f}(\mathbf{V}_n\mathbf{z})$, $\mathbf{b}_n = \mathbf{V}_n^{\mathrm{T}}\mathbf{b}$ and $\mathbf{l}_n = \mathbf{V}_n^{\mathrm{T}}\mathbf{l}$. We assume that $\mathbf{V}_n$ is an orthonormal basis of the projection subspace. One of the issues associated with this approach is that one must have a representation of $\mathbf{g}(\mathbf{z}) = \mathbf{V}_n^{\mathrm{T}}\mathbf{f}(\mathbf{V}_n\mathbf{z})$ that can be efficiently stored and evaluated. The challenge of this issue is highlighted in [8]. If $\mathbf{f}$ has a certain structure, then one may exploit such structure to derive an efficient representation of $\mathbf{g}$. For example, in [3, 2], $\mathbf{f}$ is considered as a quadratic function $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{J}(\mathbf{x}\otimes\mathbf{x})$, and in [5], $\mathbf{f}$ is represented as a gradient of a scalar function $\mathbf{f}(\mathbf{x}) = \nabla_{\mathbf{x}}\phi(\mathbf{x})$.

It is often the case that in order to obtain some pre-knowledge about the dynamical behavior of the full-order nonlinear system, we intentionally linearize a system even if it is not near the equilibrium and accept some degree of error rather than confront the full-order nonlinear system. To understand the limitation of the linearization approach, namely, when a reduced-order model strictly based on the Jacobian of the nonlinear state evaluation function $\mathbf{f}$ is accurate enough for a particular application, we may invoke the tool of variational analysis to analyze the contribution of the higher order nonlinear term [12, p.113]. As a by-product, we may also use the resulting sequence of linearized systems to develop a technique for the reduced-order model of the nonlinear dynamical system, as reported in [9].

3 Bilinearization method

The second approach is intended to explicitly incorporate the higher order nonlinear terms in the power series expansion (2.2) of $\mathbf{f}$ into the construction of a Krylov projection subspace. The approach is based on the Carleman bilinearization of a nonlinear system. The following one-sided Krylov method is similar to the method presented in [10].

By Carleman bilinearization (see, for example, [12, 13]), the nonlinear system (1.1) can be approximated by a bilinear system given in the following form

$$\begin{cases} \dot{\widehat{\mathbf{x}}} &= \widehat{\mathbf{A}}\widehat{\mathbf{x}} + \widehat{\mathbf{N}}\,\widehat{\mathbf{x}}\,u + \widehat{\mathbf{b}}u, \\ \widehat{y} &= \widehat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{x}}. \end{cases} \tag{3.4}$$

The Volterra-Wiener representation of the bilinear systems (3.4) with the kernel in regular form is given by $y(t) = \sum_{k=1}^{\infty} y_k(t)$, where the degree-$k$ subsystem $y_k(t)$ is given by

$$y_k(t) = \int_0^t h_{reg}(t_1, \ldots, t_k) u(t - t_1 - \cdots - t_k) u(t - t_2 - \cdots - t_k) \cdots u(t - t_k) dt_1 \cdots dt_k$$

with the associated k-th degree regular kernel

$$h_{reg}(t_1, \ldots, t_k) = \widehat{\mathbf{c}}^T e^{\widehat{\mathbf{A}} t_k} \widehat{\mathbf{N}} \cdots \widehat{\mathbf{N}} e^{\widehat{\mathbf{A}} t_2} \widehat{\mathbf{N}} e^{\widehat{\mathbf{A}} t_1} \widehat{\mathbf{b}}.$$

The multi-dimensional Laplace transform of $h_{reg}(t_1, \ldots, t_k)$ yields the transfer function

$$H_k(s_1, \ldots, s_k) = \widehat{\mathbf{c}}^{\mathrm{T}} (s_k \mathbf{I} - \widehat{\mathbf{A}})^{-1} \widehat{\mathbf{N}} \cdots \widehat{\mathbf{N}} (s_2 \mathbf{I} - \widehat{\mathbf{A}})^{-1} \widehat{\mathbf{N}} (s_1 \mathbf{I} - \widehat{\mathbf{A}})^{-1} \widehat{\mathbf{b}}. \tag{3.5}$$

From the power series expansion of $(s_j \mathbf{I} - \widehat{\mathbf{A}})^{-1}$, it is natural to define the corresponding *multi-moments* as

$$m(\ell_1, \ell_2, \ldots, \ell_k) = (-1)^k \widehat{\mathbf{c}}^{\mathrm{T}} \widehat{\mathbf{A}}^{-\ell_k} \widehat{\mathbf{N}} \ldots \widehat{\mathbf{N}} \widehat{\mathbf{A}}^{-\ell_2} \widehat{\mathbf{N}} \widehat{\mathbf{A}}^{-\ell_1} \widehat{\mathbf{b}}, \tag{3.6}$$

where ℓ_i are nonnegative integers. The expressions of the transfer function (3.5) and the associated multi-moments (3.6) suggest that in order to match the moments for the degree-k kernel, we can first generate the subspace $\mathbf{V}^{(k)}$ of nested Krylov subspaces with depth k defined by

$$\mathrm{span}\{\mathbf{V}^{(k)}\} = \mathcal{K}_m \left(\cdots\cdots \mathcal{K}_m \left(\widehat{\mathbf{A}}^{-1}, \widehat{\mathbf{A}}^{-1} \mathbf{N} \cdot \mathcal{K}_m (\widehat{\mathbf{A}}^{-1}, \widehat{\mathbf{A}}^{-1} \mathbf{b}) \right) \cdots \right), \tag{3.7}$$

for $k = 1, 2, \ldots, p$, and then take a union of the subspaces

$$\mathrm{span}\{\mathbf{V}_n\} = \bigcup_{k=1}^{p} \mathrm{span}\{\mathbf{V}^{(k)}\}. \tag{3.8}$$

Once the basis $\mathbf{V}_n$ of the projection subspace is extracted, we can approximate the state vector $\widehat{\mathbf{x}}(t)$ by another state vector $\mathbf{z}(t)$ constrained to the subspace $\mathrm{span}\{\mathbf{V}_n\}$, i.e., let $\widehat{\mathbf{x}}(t) \approx \mathbf{V}_n \mathbf{z}(t)$. This yields a reduced-order model of the bilinear system (3.4):

$$\begin{cases} \dot{\mathbf{z}} = \widehat{\mathbf{A}}_n \mathbf{z} + \widehat{\mathbf{N}}_n \,\mathbf{z}\, u + \widehat{\mathbf{b}}_n u, \\ \widetilde{y} = \widehat{\mathbf{c}}_n^{\mathrm{T}} \mathbf{z}. \end{cases} \tag{3.9}$$

This approach can explicitly incorporate higher order nonlinear terms of the state evolution function $\mathbf{f}$.

To match a desired number of multimoments, the dimension of the one-sided Krylov subspace $\text{span}\{\mathbf{V}_n\}$ can become quite large even for lower degree kernels, which results in a high dimension of the reduced-order bilinear model. For example, to match the multimoments of the first, second and third order kernels up to order m:

$$\widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-\ell_1}\widehat{\mathbf{b}}, \quad \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\widehat{\mathbf{b}}, \quad \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-\ell_3}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\widehat{\mathbf{b}},$$

where $\ell_1, \ell_2, \ell_3 = 1, \ldots, m$, it requires a subspace of dimension $m + m^2 + m^3$. To avoid such rapid growth of the dimension, we propose to build a pair of biorthogonal bases for both left and right Krylov subspaces to match the multimoments in a more efficient way. The result is more dramatic than the difference between one-sided and two-sided Krylov methods for linear systems.

In the two-sided method, the right subspace is the same as (3.8). For the left subspace, we first construct the p nested Krylov subspaces with depth k:

$$\text{span}\{\mathbf{W}^{(k)}\} = \mathcal{K}_m\left(\cdots\cdots\mathcal{K}_m\left(\widehat{\mathbf{A}}^{-T}, \mathbf{N}^T\widehat{\mathbf{A}}^{-T}\cdot\mathcal{K}_m(\widehat{\mathbf{A}}^{-T}, \mathbf{c})\right)\cdots\right),$$

where $k = 1, \ldots, p$. The left projection subspace $\mathbf{W}_n$ is then taken as the union of these subspaces

$$\text{span}\{\mathbf{W}_n\} = \bigcup_{k=1}^{p} \text{span}\{\mathbf{W}^{(k)}\}.$$

Furthermore, the bases $\mathbf{V}_n$ and $\mathbf{W}_n$ are constructed to be biorthogonal. The system matrices in the reduced-order bilinear model (3.9) are then defined as

$$\widehat{\mathbf{A}}_n^{-1} = \mathbf{W}_n^T\widehat{\mathbf{A}}^{-1}\mathbf{V}_n, \;\; \widehat{\mathbf{N}}_n = \widehat{\mathbf{A}}_n\mathbf{W}_n^T\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{N}}\mathbf{V}_n, \;\; \widehat{\mathbf{b}}_n = \widehat{\mathbf{A}}_n\mathbf{W}_n^T\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{b}}, \;\; \widehat{\mathbf{c}}_n = \mathbf{V}_n^T\widehat{\mathbf{c}}.$$

It can be shown that the reduced-order model matches all multimoments that can be represented through the scalar product

$$\mathbf{s}^T\mathbf{r} = (-1)^k\widehat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{A}}^{-\ell_k}\widehat{\mathbf{N}}\ldots\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\widehat{\mathbf{b}},$$

where $\mathbf{r} \in \text{span}\{\mathbf{V}_n\}$ and $\mathbf{s} \in \text{span}\{\mathbf{W}_n\}$. It matches more number of multimoments than the total number of basis vectors, whereas by using only one-sided basis, it generally only matches the same number of multimoments as the number of basis vectors.

Example 1. Let the right and left subspaces be

$$\text{span}\{\mathbf{V}\} = \{\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{b}}, \widehat{\mathbf{A}}^{-2}\widehat{\mathbf{b}}, \ldots, \widehat{\mathbf{A}}^{-7}\widehat{\mathbf{b}}, \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{b}}, \ldots, \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-4}\widehat{\mathbf{b}}\}$$

and

$$\text{span}\{\mathbf{W}\}^T = \{\widehat{\mathbf{c}}^T, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-1}, \ldots, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-6}, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{N}}, \ldots, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-4}\widehat{\mathbf{N}}, \}$$

By using the combined information in the basis $\mathbf{V}$ and $\mathbf{W}$, it can be shown that the reduced-order bilinear model matches 13 moments of the degree-1 kernel

$$\widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{b}}, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-2}\widehat{\mathbf{b}}, \ldots, \widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-13}\widehat{\mathbf{b}},$$

40 multimoments of the degree-2 kernel

$$\widehat{\mathbf{c}}^T\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\widehat{\mathbf{b}},$$

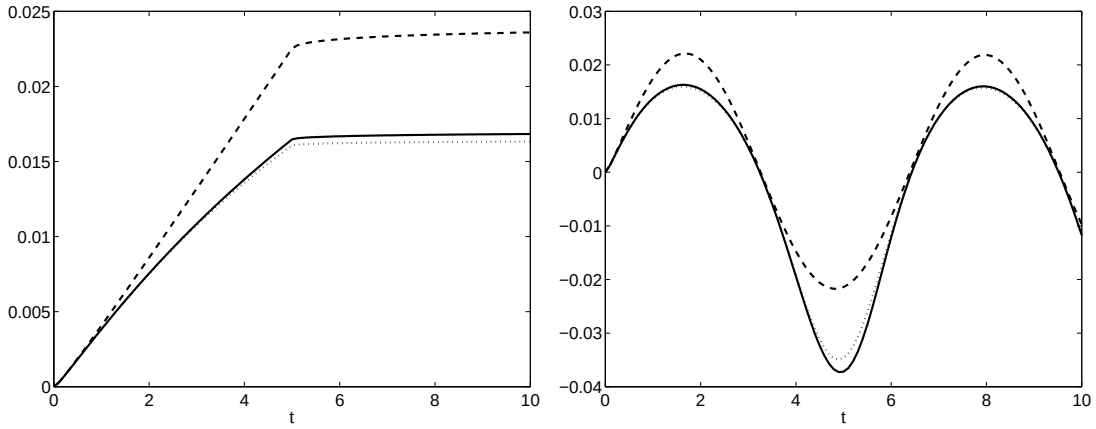

Figure 1: Transient responses of a nonlinear circuit system with full-order system (solid line), linearized system (dash line), and the reduced-order bilinear system (dot-dot line). The plots of full-order bilinear system overlap the plots of the reduced-order one and is omitted. Left: The input signal $u(t)$ is the ramp function ($u(t) = t/5, 0 \le t \le 5,\ u(t) = 1, t > 5$). Right: the input signal is the sin function $u(t) = \sin(t)$.

where $\ell_1, \ell_2 = 1, 2, 3, 4$; $\ell_1 = 5, 6, 7, \ell_2 = 1, 2, 3, 4$; or $\ell_1 = 1, 2, 3, 4, \ell_2 = 5, 6, 7$, and 16 multimoments of the degree-3 kernel

$$\widehat{\mathbf{c}}^T \widehat{\mathbf{A}}^{-\ell_3} \widehat{\mathbf{N}} \widehat{\mathbf{A}}^{-\ell_2} \widehat{\mathbf{N}} \widehat{\mathbf{A}}^{-\ell_1} \widehat{\mathbf{b}},$$

where $\ell_1 = 1, 2, 3, 4, \ell_2 = 1, \ell_3 = 1, 2, 3, 4$. It matches a total of 69 multimoments with only 22 basis vectors. On the other hand, to be able to match the same number of moments using only the left or right subspace, we will need 69 basis vectors. As a result, the dimension of the reduced-order model using two-sided subspaces will be of order 11, not the order of 69 by using one-sided subspace.

One critical issue associated with the bilinearization method is the growth of the dimension of the bilinear system (3.4) as a result of Carleman linearization. For example, even if we only use the first two terms in the power series expansion (2.2) of $\mathbf{f}$, the order of the resulting bilinear system is about $\mathcal{O}(N^2)$. However, the matrices $\mathbf{A}_i$ in the power series expansion (2.2) of $\mathbf{f}$ are generally extremely sparse, and the matrices $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{N}}$ in the bilinear system (3.4) are highly structured, so one can exploit these facts in a Krylov process, namely through the matrix-vector multiplications during the Lanczos or Arnoldi process, to produce an efficient reduced-order model.

Example 2. In Fig. 1, we show the transient responses of a RC circuit with nonlinear resistors as described in [3]. The dimension of the original full-order nonlinear system is $N = 100$. Although the order of the bilinear system by using second-order approximation is of dimension 10100, the order of the reduced one is only 11. The pair of bases $\mathbf{V}_n$ and $\mathbf{W}_n$ is constructed as described in Example 1.

Further details of the bilinearization-based Krylov subspace techniques for reduced-order modeling of large-scale nonlinear dynamical systems will be reported in [1].

Acknowledgements. Support for this work has been provided in part by NSF under grant ACI-9813362 and DOE under grant DE-FG03-94ER25219, and in part by a MICRO project (#00-005) from University of California and MSC.software Corporation. DS was also supported by the Foundation BLANCEFLOR Boncompagni-Ludovisi, neé Bildt and the Royal Swedish Academy of Sciences while visiting University of California at Davis during the academic year of 2000-2001.

References

[1] Z. Bai and D. Skoogh. Reduced-order modeling of structured nonlinear systems. work in progress. 2002.

[2] J. Chen and S. M. Kang. An algorithm for automatic model-order reduction of nonlinear MEMS devices. In *IEEE Inter. Symp. on Circuits and Systems*, Geneva, Switzerland, 2000.

[3] Y. Chen and J. White. A quadratic method for nonlinear model order reduction. In *Inter. Conf. on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators*, San Diego, 2000.

[4] R. W. Freund and P. Feldman. Small-signal circuit analysis and sensitivity computations with the PVL algorithm. *IEEE Trans. Circuits and Systems – II*, 43:577–585, 1996.

[5] L. D. Gabbay, J. E. Mehner, and S. D. Senturia. Computer-aided generation of nonlinear reduced-order dynamic macromodels – I: Non-stress-stiffened case. *J. of Microelectromechanical Systems*, 9:262–269, 2000.

[6] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge Univ. Press, 1996.

[7] S. Lall, J. E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Inter. J. on Rubust and Nonlinear Control*, 2000. submitted.

[8] T. Mukherjee, G. Fedder, D. Ramaswamy, and J. White. Emerging simulation approaches for micromachined devices. *IEEE Trans. CAD*, 19:1572–1588, 2000.

[9] J. Phillips. Automated extraction of nonlinear circuit macromodels. In *Proceedings of IEEE 2000 Custom Integrated Circuits Conference*, pages 451–452, 2000.

[10] J. Phillips. Projection frameworks for model reduction of weakly nonlinear systems. In *Proceedings of DAC 2000*, pages 184–189, 2000.

[11] M. Rathinam and L. Petzold. An iterative method for simulation of large scale modular systems using reduced order models. In *Proc. IEEE Control and Decision Conference*, Australia, 2000.

[12] W. J. Rugh. *Nonlinear system theory*. The John Hopkins University Press, Boltimore, 1981.

[13] S. Sastry. *Nonlinear systems: Analysis, stability and control*. Springer, New York, 1999.

[14] J. M. A. Scherpen. Balancing for nonlinear systems. *Systems and Control Letters*, 21:143–153, 1993.

[15] S. D. Senturia. *Microsystem design*. Kluwer Academic Publishers, Boston, 2001.