Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 415 (2006) 406–425
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

A projection method for model reduction of bilinear dynamical systems

Zhaojun Bai [a,*], Daniel Skoogh [b]

[a] *Department of Computer Science, University of California, One Shield Avenue, Davis, CA 95616, United States*
[b] *Department of Autonomous Systems, Swedish Defence Research Agency, 172 90 Stockholm, Sweden*

Received 27 June 2003; accepted 25 April 2005
Available online 24 August 2005
Submitted by D. Sorensen

Abstract

A Krylov subspace based projection method is presented for model reduction of large scale bilinear systems. A reduced bilinear system is constructed in such a way that it matches a desired number of moments of multivariable transfer functions corresponding to the kernels of Volterra series representation of the original system. Applications to the simulation of dynamical responses of a nonlinear circuit and a micromachined device are presented to illustrate the efficiency of the new method and compare with an approach recently proposed by Phillips.

Keywords: Bilinear systems; Model order reduction; Krylov subspace; Moment-matching

1. Introduction

Bilinear systems are a special class of nonlinear systems that are linear in input and linear in state but not jointly linear in state and input. Specifically, a time invariant single-input and single-output (SISO) bilinear system, symbolically denoted as Σ, has a state form as follows:

* Corresponding author.
E-mail addresses: bai@cs.ucdavis.edu (Z. Bai), skoogh@foi.se (D. Skoogh).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2005.04.032

$$\Sigma: \quad \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{N}\mathbf{x}(t)u(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t), \end{cases} \tag{1}$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$. Here t is the time variable, $\mathbf{x}(t) \in \mathscr{R}^N$ is the state of the system, N is the dimension of the state space. $u(t)$ and $y(t)$ are input and output scalar functions. $\mathbf{A}, \mathbf{N} \in \mathscr{R}^{N\times N}$ and $\mathbf{b}, \mathbf{c} \in \mathscr{R}^N$ are constant matrices and vectors.

Bilinear systems arise as natural models for a variety of physical and biomedical processes. The study of such systems goes back to early 1960s. A comprehensive treatment of the bilinear systems, including system characterization, structural properties, stability and applications, can be found in [14]. Our current interest in studying of model reduction of the bilinear system (1) stems from the need of simulation of large scale nonlinear systems of the form

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))u(t), \\ y(t) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t), \end{cases} \tag{2}$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0$, where as above, $\mathbf{x}(t)$ is the state vector of dimension N, $u(t)$ and $y(t)$ are input and output scalar functions. $\mathbf{c}$ is an output measurement array. The nonlinear vector-valued functions $\mathbf{f}, \mathbf{g} : \mathscr{R}^N \to \mathscr{R}^N$ are sufficient smooth and $\mathbf{f}$ has an equilibrium. Origins of such system include the simulation of time-varying nonlinear circuit elements by independent excitation source [7,3], and MEMS devices, such as micro-pressure sensor [15]. The modeling of dynamical behavior of a voltage-controlled parallel-plate electrostatic actuator also derives a set of state equations of the form (2) [24, 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. With a linearization near an equilibrium point of $\mathbf{f}$ or the Carleman bilinearization, the nonlinear system (2) can be approximated by a bilinear system of the form (1). Carleman bilinearization allows high degree nonlinearity to be explicitly and systematically incorporated in the bilinear system approximation. This will be further discussed in Section 5.

In this paper, we develop a computational technique for reduced-order modeling of the bilinear system (1). For a given bilinear system Σ, another bilinear system $\widehat{\Sigma}$ is constructed such that it has a much smaller state dimension, yet still retains the original behavior under investigation to high accuracy. An accurate and effective reduced system thus replaces the original one and can be applied for a variety types of analysis of physical systems it emulates. Consequently, we can significantly reduce design and simulation time to meet today's high demand for short product development times.

The model reduction of bilinear systems was recently studied by Phillips [17,18]. The technique presented in this paper is inspired by his work. We shall clarify theoretical and algorithmic differences of the two approaches in Section 3 and compare their performance on two case studies in Section 5. Karhunen–Loève expansion based methods and methods of balanced truncation are two most well-known methods for model reduction of nonlinear systems. The former methods are methods of least-squares approximation and are known in literature by several names, including proper

orthogonal decomposition (POD) and principal component analysis (PCA). The latter methods are the extensions of successful balanced truncation methods for linear systems to nonlinear systems. The interested reader is referred to [10,23,12,13] and references therein. Krylov subspace based projection methods have made remarkable progresses and successes for model reduction of very large linear systems over the past decade, see for example [5,8,16,6,1]. It is an active research topic to extend these Krylov subspace based projection methods to large scale nonlinear systems [20,3,9,19]. Bilinear systems are a special class of nonlinear systems, and can approximate nonlinear systems of the form (2) in a systematical way up to a desired degree of accuracy by using the Carleman bilinearization procedure. In this paper, we shall examine this approach for two nonlinear systems arising from MEMS and circuit simulation applications.

The outline of the rest of the paper is as follows. In the next section, we begin with a review of basics of Volterra series representation theory of bilinear systems to introduce essential concepts on transfer functions and moments, and to formally define the goal of model reduction. A projection framework and its theoretical properties are presented in Section 3. In Section 4, we consider a practical implementation of the projection framework. Two applications of this implementation for model reduction of nonlinear systems arising from nonlinear circuit and micromachined devices are presented in Section 5. Some concluding remarks are in Section 6.

With a few exceptions, we will follow the notational conventions used in numerical computing. The boldface letters are used to denote vectors and matrices. The identity matrix is denoted by $\mathbf{I}$ and the zero matrix by $\mathbf{0}$. The actual dimensions of these matrices should be apparent from the context. $\mathbf{v}_i$ denotes the ith column of the matrix $\mathbf{V}$. $\mathbf{V}_{[p]}$ denotes the first p columns of $\mathbf{V}$. The sets of real numbers, column N-vectors and $N \times N$ matrices are denoted by $\mathscr{R}$, $\mathscr{R}^N$ and $\mathscr{R}^{N\times N}$, respectively. Finally, $\mathscr{K}_q(\mathbf{A}, \mathbf{R})$ denotes a (block) Krylov subspace generated by a matrix $\mathbf{A}$ with a block of starting vectors $\mathbf{R}$, i.e., $\mathscr{K}_q(\mathbf{A}, \mathbf{R}) = \text{span}\{\mathbf{R}, \mathbf{AR}, \mathbf{A}^2\mathbf{R}, \ldots, \mathbf{A}^{q-1}\mathbf{R}\}$.

2. Bilinear systems and reduced-order modeling

The Volterra series representation of nonlinear systems is a commonly used system characterization, which generalizes the concept of the linear system impulse response and its Laplace transform, transfer function and moments. For a linear time-invariant SISO system of the form

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t), \end{cases}$$

with the assumption of zero initial condition, the input–output behavior can be described by the following convolution representation:

$$y(t) = \int_0^t h(\sigma)u(t-\sigma)\,\mathrm{d}\sigma,$$

where $u(t)$ is the input signal, $y(t)$ is the output signal, and $h(t) = \mathbf{c}^{\mathrm{T}}\mathrm{e}^{\mathbf{A}t}\mathbf{b}$ is the impulse response which is also called a *kernel*. For a complete theory of linear systems, see for example [11].

A generalization of the convolution representation for the bilinear system (1) is given by the Volterra series

$$y(t) = \sum_{k=1}^{\infty} y_k(t),$$

where $y_k(t)$ is the kth subsystem convolution representation given by

$$y_k(t) = \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} h(t_1, t_2, \ldots, t_k) u(t - t_1 - t_2 - \cdots - t_k) \cdots \times u(t - t_k)\, \mathrm{d}t_k \cdots \mathrm{d}t_1,$$

and the associated *degree-k* (*regular*) *kernel* $h(t_1, t_2, \ldots, t_k)$ is given by

$$h(t_1, t_2, \ldots, t_k) = \mathbf{c}^{\mathrm{T}}\mathrm{e}^{\mathbf{A}t_k}\mathbf{N}\mathrm{e}^{\mathbf{A}t_{k-1}}\mathbf{N} \cdots \mathrm{e}^{\mathbf{A}t_2}\mathbf{N}\mathrm{e}^{\mathbf{A}t_1}\mathbf{b}, \tag{3}$$

see for example [21,14,22]. Here for simplicity, we assume that $\mathbf{x}(0) = \mathbf{x}_0 = \mathbf{0}$.

The multivariable Laplace transform of the degree-k kernel (3) defines the *kth transfer function*:

$$H(s_1, s_2, \ldots, s_k) = \mathbf{c}^{\mathrm{T}}(s_k\mathbf{I} - \mathbf{A})^{-1}\mathbf{N}(s_{k-1}\mathbf{I} - \mathbf{A})^{-1}\mathbf{N} \cdots \cdot (s_2\mathbf{I} - \mathbf{A})^{-1}\mathbf{N}(s_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}. \tag{4}$$

By rewriting the above transfer function as

$$H(s_1, s_2, \ldots, s_k) = (-1)^k\mathbf{c}^{\mathrm{T}}(\mathbf{I} - s_k\mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{N} \cdots \cdot (\mathbf{I} - s_2\mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{N}(\mathbf{I} - s_1\mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{b}$$

and by making use of the Neumann expansion

$$(\mathbf{I} - s\mathbf{A}^{-1})^{-1} = \mathbf{I} + s\mathbf{A}^{-1} + s^2\mathbf{A}^{-2} + s^3\mathbf{A}^{-3} + \cdots,$$

the transfer function $H(s_1, s_2, \ldots, s_k)$ of the kth subsystem can be expanded in a multivariable Maclaurin series

$$H(s_1, s_2, \ldots, s_k) = \sum_{\ell_k=1}^{\infty} \cdots \sum_{\ell_1=1}^{\infty} m(\ell_1, \ell_2, \ldots, \ell_k) s_1^{\ell_1-1} s_2^{\ell_2-1} \cdots s_k^{\ell_k-1},$$

where

$$m(\ell_1, \ell_2, \ldots, \ell_k) = (-1)^k\mathbf{c}^{\mathrm{T}}\mathbf{A}^{-\ell_k}\mathbf{N} \cdots \mathbf{A}^{-\ell_2}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b} \tag{5}$$

are called (*low frequency*) *multimoments* of the kth subsystem. For example, the transfer function $H(s_1)$ of the first subsystem can be written in the series

$$H(s_1) = \mathbf{c}^{\mathrm{T}}(s_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} = \sum_{\ell_1=1}^{\infty} m(\ell_1) s_1^{\ell_1-1},$$

where the associated (low frequency) moments are $m(\ell_1) = -\mathbf{c}^{\mathrm{T}}\mathbf{A}^{-\ell_1}\mathbf{b}$. Similarly, the transfer function $H(s_1, s_2)$ of the second subsystem can be written in the series

$$H(s_1, s_2) = \mathbf{c}^{\mathrm{T}}(s_2\mathbf{I} - \mathbf{A})^{-1}\mathbf{N}(s_1\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} = \sum_{\ell_2=1}^{\infty}\sum_{\ell_1=1}^{\infty} m(\ell_1, \ell_2)s_1^{\ell_1-1}s_2^{\ell_2-1},$$

where the corresponding (low frequency) moments are $m(\ell_1, \ell_2) = \mathbf{c}^{\mathrm{T}}\mathbf{A}^{-\ell_2}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b}$.

Alternatively, one can expand the transfer function $H(s_1, s_2, \ldots, s_k)$ at $s_i = \infty$ for $i = 1, 2, \ldots, k$, and obtain the so-called *high frequency moments*. Furthermore, one can also expand the transfer function at zero for some variables s_i and at infinity for other variables s_j, and obtain the so-called *mixed frequency moments*.

The goal of reduced-order modeling is that for the given bilinear system Σ (1), find a reduced bilinear system $\widehat{\Sigma}$ of the same form but with many fewer states, such that the output behavior of the original system Σ is faithfully retained by the reduced system $\widehat{\Sigma}$ for an admissible set of inputs $u(t)$. Specifically, in this paper, we study the problem to be solved is thus: for given matrices $\mathbf{A}, \mathbf{N} \in \mathscr{R}^{N\times N}$ and vectors $\mathbf{b}, \mathbf{c} \in \mathscr{R}^N$ of the bilinear system Σ, find matrices $\widehat{\mathbf{A}}, \widehat{\mathbf{N}} \in \mathscr{R}^{n\times n}$ and vectors $\hat{\mathbf{b}}, \hat{\mathbf{c}} \in \mathscr{R}^n$, such that $n < N$ and a prescribed number of moments of the reduced bilinear system

$$\widehat{\Sigma}: \quad \begin{cases} \dot{\hat{\mathbf{x}}}(t) = \widehat{\mathbf{A}}\hat{\mathbf{x}}(t) + \widehat{\mathbf{N}}\hat{\mathbf{x}}(t)u(t) + \hat{\mathbf{b}}u(t), \\ \hat{y}(t) = \hat{\mathbf{c}}^{\mathrm{T}}\hat{\mathbf{x}}(t), \end{cases} \tag{6}$$

matches the corresponding moments of the original bilinear system Σ. Such a prescribed number is specified by two positive integers r and q, where r corresponds to the transfer functions $H(s_1)$, $H(s_1, s_2)$ to $H(s_1, s_2, \ldots, s_r)$ of the first r subsystems, and q corresponds to the order of approximation to these transfer functions. Namely, for given r and q, we want to construct a reduced bilinear system $\widehat{\Sigma}$ satisfying the following moment-matching condition

$$m(\ell_1, \ell_2, \ldots, \ell_k) = \hat{m}(\ell_1, \ell_2 \ldots, \ell_k) \tag{7}$$

for $k = 1, 2, \ldots, r$ and $\ell_1, \ell_2, \ldots, \ell_k = 1, 2, \ldots, q$. Here

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

are the moments of the transfer function of the kth subsystem of the reduced system $\widehat{\Sigma}$ defined in the following multivariable Maclaurin series:

$$\widehat{H}(s_1, s_2, \ldots, s_k) = \sum_{\ell_k=1}^{\infty}\cdots\sum_{\ell_1=1}^{\infty}\hat{m}(\ell_1, \ell_2, \ldots, \ell_k)s_1^{\ell_1-1}s_2^{\ell_2-1}\cdots s_k^{\ell_k-1}.$$

The total number of the moments matched is $q + q^2 + \cdots + q^r$. The condition (7) implies the following orders of approximations in terms of transfer functions:

$$H(s_1, s_2, \ldots, s_k) = \widehat{H}(s_1, s_2, \ldots, s_k) + \mathscr{O}\big(s_1^{q_1}s_2^{q_2}\cdots s_k^{q_k}\big) \tag{8}$$

for $k = 1, 2, \ldots, r$, where $q_1, q_2, \ldots, q_k \leqslant q$ and at least one of them is equal to q.

We note that the approximation orders in (8) are generally not necessary to be the same for all transfer functions $H(s_1, s_2, \ldots, s_k)$. In Section 4, this condition will be relaxed in a practical algorithm.

3. A projection framework

In this section, we present a projection framework for constructing the reduced bilinear system $\widehat{\Sigma}$ (6) with the desired moment-matching property (7). At the end of this section, we clarify the differences between our approach and a previous one proposed by Phillips [17,18].

From the expression of moments (5) and the practice in Krylov subspace based projection methods for linear systems, we first define the following sequence of Krylov subspaces for the desired number of moment-matching specified by r and q:

$$\text{span}\{\mathbf{V}^{(k)}\} = \mathscr{K}_q\left(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{N}\mathbf{V}^{(k-1)}\right) \tag{9}$$

for $k = 2, 3, \ldots, r$ with $\text{span}\{\mathbf{V}^{(1)}\} = \mathscr{K}_q(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{b})$. Then define $\mathbf{V}$ as the basis of a union of these subspaces:

$$\text{span}\{\mathbf{V}\} = \text{span}\left\{\bigcup_{k=1}^{r} \text{span}\{\mathbf{V}^{(k)}\}\right\}. \tag{10}$$

The subspace spanned by $\mathbf{V}$ will be the projection subspace, denoted as $\mathscr{P}$, for reduced-order modeling of the bilinear system (1). Before we proceed, a few remarks are in order. First, if there is no deflation, namely, the column vectors of $\mathbf{V}^{(k)}$ are linearly independent, then the dimension of the projection subspace $\mathscr{P}$ is $n = \sum_{k=1}^{r} q^k$. This is equal to the desired number of moments to be matched, see (7). Second, as we state at the end of Section 2, the dimensions of the block Krylov subspaces $\mathscr{K}_q$ in (9) do not have to be the same for all k. Similarly, the number of starting vectors for defining $\mathscr{K}_q$ in (9) can just take a portion of columns of $\mathbf{V}^{(k-1)}$. In Section 4, we will present a practical algorithm to provide these options.

We now derive a reduced bilinear system using a projection formulation. We assume that the basis $\mathbf{V}$ of the projection subspace $\mathscr{P}$ is orthonormal, i.e., $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$. First by multiplying $\mathbf{A}^{-1}$ to the first equation in the original model (1) from the left, it yields

$$\begin{cases} \mathbf{A}^{-1}\dot{\mathbf{x}}(t) = \mathbf{x}(t) + \mathbf{A}^{-1}\mathbf{N}\mathbf{x}(t)u(t) + \mathbf{A}^{-1}\mathbf{b}u(t), \\ y(t) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t). \end{cases} \tag{11}$$

Recall that the concept of projecting the states of the original systems onto the subspace $\mathscr{P}$ spanned by $\mathbf{V}$ can be viewed as performing a change of variables

$$\mathbf{x}(t) \approx \mathbf{V}\hat{\mathbf{x}}(t), \tag{12}$$

where $\hat{\mathbf{x}}(t) \in \mathscr{R}^n$. Substituting (12) into (11) and multiplying the first equation of (11) with $\mathbf{V}^{\mathrm{T}}$ from the left yield

$$\begin{cases} \mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V}\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{x}}(t) + \mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{N}\mathbf{V}\hat{\mathbf{x}}(t)u(t) + \mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{b}u(t), \\ \widehat{y}(t) = \mathbf{c}^{\mathrm{T}}\mathbf{V}\,\hat{\mathbf{x}}(t). \end{cases} \tag{13}$$

Then a reduced-order model of the bilinear system (1) is naturally defined as

$$\widehat{\Sigma}: \quad \begin{cases} \dot{\hat{\mathbf{x}}}(t) = \widehat{\mathbf{A}}\hat{\mathbf{x}}(t) + \widehat{\mathbf{N}}\hat{\mathbf{x}}(t)u(t) + \hat{\mathbf{b}}u(t), \\ \hat{y}(t) = \hat{\mathbf{c}}^{\mathrm{T}}\hat{\mathbf{x}}(t), \end{cases} \tag{14}$$

where

$$\widehat{\mathbf{A}} = (\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{-1}, \quad \widehat{\mathbf{N}} = \widehat{\mathbf{A}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{N}\mathbf{V}, \quad \hat{\mathbf{b}} = \widehat{\mathbf{A}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{b}, \quad \hat{\mathbf{c}} = \mathbf{V}^{\mathrm{T}}\mathbf{c}. \tag{15}$$

The forms of the matrices $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{N}}$ in (15) are quite unusual. However, this formulation is essential to achieve the maximum level of moment matching from the projection method. This will be justified by the following theorem and comments. In numerical simulation, the explict inverse of $\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V}$ can be avoided if we work with the reduced system in the form (13).

Theorem 3.1. *For $k = 1, 2, \ldots, r$, the q^k moments $\hat{m}(\ell_1, \ell_2, \ldots, \ell_k)$ of the kth subsystems of the reduced system $\widehat{\Sigma}$ matches the q^k moments $m(\ell_1, \ell_2, \ldots, \ell_k)$ of the original system Σ, where $\ell_1, \ell_2, \ldots, \ell_k = 1, 2, \ldots, q$.*

Proof. In the following, we shall prove the moment-matching property for the first and second subsystems in detail. The proof for higher degree kernels can be shown straightforwardly by induction.

First we note that by the construction of the basis $\mathbf{V}$ shown in (10), the following vectors are in the projection subspace $\mathscr{P}$:

$$\begin{aligned}
&\mathbf{A}^{-\ell_1}\mathbf{b},\\
&\mathbf{A}^{-\ell_2}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b},\\
&\cdots,\\
&\mathbf{A}^{-\ell_k}\mathbf{N}\mathbf{A}^{-\ell_{k-1}}\mathbf{N}\cdots\mathbf{A}^{-\ell_1}\mathbf{b},\\
&\cdots,\\
&\mathbf{A}^{-\ell_r}\mathbf{N}\mathbf{A}^{-\ell_{r-1}}\mathbf{N}\cdots\cdots\cdots\mathbf{A}^{-\ell_1}\mathbf{b},
\end{aligned}$$

where $\ell_1, \ell_2, \ldots, \ell_r = 1, 2, \ldots, q$. Under the assumption of nondeflation, the total number of vectors in the basis is $n = \sum_{k=1}^{r} q^k$.

Next, we note the fact that if the basis $\mathbf{V}$ of $\mathscr{P}$ is orthonormal, then for any vector $\mathbf{z} \in \mathscr{P}$,

$$\mathbf{z} = \mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{z}. \tag{16}$$

We now prove that the moments $\hat{m}(\ell_1) = -\hat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}}$ of the first subsystem of the reduced system $\widehat{\Sigma}$ matches the corresponding moments $m(\ell_1)$ of the original system Σ for $\ell_1 = 1, 2, \ldots, q$. Since the vectors $\mathbf{A}^{-\ell_1}\mathbf{b}$ are in $\mathscr{P}$, by (16), we have

$$\mathbf{A}^{-\ell_1}\mathbf{b} = \mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-\ell_1}\mathbf{b}, \tag{17}$$

Repeatedly using (17), it yields

$$\begin{aligned}
\mathbf{V}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}} &= \mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{\ell_1}\widehat{\mathbf{A}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{b}\\
&= \mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{\ell_1-1}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{b}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{\ell_1-2}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V}\mathbf{V}^{T}\mathbf{A}^{-1}\mathbf{b}\\
&= \mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{\ell_1-2}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-2}\mathbf{b}\\
&\vdots\\
&= \mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-\ell_1}\mathbf{b} = \mathbf{A}^{-\ell_1}\mathbf{b}. \qquad (18)
\end{aligned}
$$

Multiplying the first and last terms in (18) with $\mathbf{c}^{\mathrm{T}}$ from the left, it yields the moment-matching property for the first subsystem

$$\hat{m}(\ell_1) = -\hat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}} = -\mathbf{c}^{\mathrm{T}}\mathbf{A}^{-\ell_1}\mathbf{b} = m(\ell_1),$$

for $\ell_1 = 1, 2, \ldots, q$.

Next we prove that the moments $\hat{m}(\ell_1, \ell_2) = \hat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}}$ of the second subsystem of the reduced system $\widehat{\Sigma}$ match the corresponding moments $m(\ell_1, \ell_2)$ of the original system Σ for $\ell_1, \ell_2 = 1, 2, \ldots, q$.

By the definitions of $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{N}}$ (15) and Eqs. (16) and (18), we have

$$
\begin{aligned}
\mathbf{V}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}} &= \mathbf{V}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-\ell_1}\mathbf{b} \quad \text{by (18)}\\
&= \mathbf{V}\widehat{\mathbf{A}}^{-\ell_2}\cdot\widehat{\mathbf{A}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{N}\mathbf{V}\cdot\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-l_1}\mathbf{b} \quad \text{by definition of } \widehat{\mathbf{N}}\\
&= \mathbf{V}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{A}}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b} \quad \text{by (16)}\\
&= \mathbf{V}(\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{V})^{\ell_2-1}\mathbf{V}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b} \quad \text{by definition of } \widehat{\mathbf{A}}\\
&= \mathbf{A}^{-\ell_2}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b}, \qquad (19)
\end{aligned}
$$

where the last equality is obtained by using the same recursive derivation as shown in (18). After multiplying the first and last terms in (19) with $\mathbf{c}^{\mathrm{T}}$ from the left, we immediately have

$$\hat{m}(\ell_1, \ell_2) = \hat{\mathbf{c}}^{\mathrm{T}}\widehat{\mathbf{A}}^{-\ell_2}\widehat{\mathbf{N}}\widehat{\mathbf{A}}^{-\ell_1}\hat{\mathbf{b}} = \mathbf{c}^{\mathrm{T}}\mathbf{A}^{-\ell_2}\mathbf{N}\mathbf{A}^{-\ell_1}\mathbf{b} = m(\ell_1, \ell_2)$$

for $\ell_1, \ell_2 = 1, 2, \ldots, q$. Thus the q^2 moments for the second subsystem matches those of the original system. □

The work presented in this section is inspired by the work of Phillips [17,18]. In [17,18], the projection subspace $\mathscr{P}$ is defined as a union of the following Krylov subspaces:

$$\mathrm{span}\{\mathbf{V}^{(k)}\} = \mathscr{K}_q\left(\mathbf{A}^{-1}, \mathbf{N}\mathbf{V}^{(k-1)}\right)$$

for $k = 2, 3, \ldots$, with $\mathrm{span}\{\mathbf{V}^{(1)}\} = \mathscr{K}_q(\mathbf{A}^{-1}, \mathbf{b})$. Let $\mathbf{V}$ be an orthonormal basis of $\mathscr{P}$, then the matrices $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{N}}$ and the vectors $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ for the reduced system are defined by

$$\widehat{\mathbf{A}} = \mathbf{V}^{\mathrm{T}}\mathbf{A}\mathbf{V}, \quad \widehat{\mathbf{N}} = \mathbf{V}^{\mathrm{T}}\mathbf{N}\mathbf{V}, \quad \hat{\mathbf{b}} = \mathbf{V}^{\mathrm{T}}\mathbf{b}, \quad \hat{\mathbf{c}} = \mathbf{V}^{\mathrm{T}}\mathbf{c}.$$

These are simpler ways to define the projection subspace and the reduced system than our definitions of the projection subspace (9) and the reduced system (14) and (15). However, by a careful analysis, it can be shown that the reduced system defined in

this way does not satisfy the desired moment-matching and approximation properties (7) and (8), and the similar ones in [17,18]. In fact, the reduced system constructed in this way matches some low frequency moments, some high frequency moments, and even some mixed frequency moments. Numerical examples in Section 4 show that the resulting reduced system in this way is less accurate and stable.

4. A practical algorithm

In this section, we describe a practical implementation of the projection framework to match moments $m(\ell_1)$ and $m(\ell_1, \ell_2)$ corresponding to the first and second degree kernels of the bilinear system Σ. This is our anticipation of the situations where this algorithm is of the most frequent use.

To match some of moments of the first and second subsystems, by Section 3, the projection subspace $\mathscr{P}$ should be defined as follows:

$$\mathscr{P} = \text{span}\{\mathbf{V}\} = \text{span}\big\{\text{span}\{\mathbf{V}^{(1)}\} \cup \text{span}\{\mathbf{V}^{(2)}\}\big\}, \tag{20}$$

where

$$\text{span}\{\mathbf{V}^{(1)}\} = \mathscr{K}_{q_1}(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{b}) \quad \text{and} \quad \text{span}\{\mathbf{V}^{(2)}\} = \mathscr{K}_{q_2}\big(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{N}\mathbf{V}^{(1)}_{[p_2]}\big).$$

The nonnegative integers q_1, q_2 and p_2 are prescribed parameters with $p_2 \leqslant q_1$. $\mathbf{V}^{(1)}_{[p_2]}$ denotes the first p_2 columns of the matrix $\mathbf{V}^{(1)}$. For $r = 2$, this is a more general case than the projection subspace defined in (9), which corresponds to the special choice $q_1 = p_2 = q_2 = q$. If there is no deflation, namely, the column vectors of $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$ are linearly independent, then the dimension of the projection subspace $\mathscr{P}$ is $n = q_1 + p_2 q_2$. Following the discussion in Section 3, it is immediately seen that the following moments are preserved:

$$\begin{aligned} &m(\ell_1) = \hat{m}(\ell_1) \quad \text{for } \ell_1 = 1, 2, \ldots, q_1, \\ &m(\ell_1, \ell_2) = \hat{m}(\ell_1, \ell_2) \quad \text{for } \ell_1 = 1, 2, \ldots, p_2 \text{ and } \ell_2 = 1, 2, \ldots, q_2. \end{aligned}$$

This implies that we have the following approximations of the transfer functions of the first and second subsystems:

$$\begin{aligned} &H(s_1) = \widehat{H}(s_1) + \mathscr{O}(s_1^{q_1}), \\ &H(s_1, s_2) = \widehat{H}(s_1, s_2) + \mathscr{O}(s_1^{p_2} \odot s_2^{q_2}). \end{aligned}$$

Here $s_1^{p_2} \odot s_2^{q_2}$ stands for the product of the powers of the variables s_1 and s_2 with either $s_1^{p_2}$ or $s_2^{q_2}$ true.

An algorithm template for generating an orthonormal basis $\mathbf{V}$ of the projection subspace $\mathscr{P}$ goes as follows.

Algorithm for generating an orthonormal basis V of $\mathscr{P}$
Input: $\mathbf{A}$, $\mathbf{N}$, $\mathbf{b}$, q_1, q_2 and p_2 with $p_2 \leqslant q_1$

Output: $\mathbf{V}$ as defined in (20) with $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$.

1. $\mathbf{r} = \mathbf{A}^{-1}\mathbf{b}$
2. $\mathbf{v}_1^{(1)} = \mathbf{r}/\|\mathbf{r}\|_2$
3. **for** $i = 1 : q_1 - 1$
4. $\quad \mathbf{r} = \mathbf{A}^{-1}\mathbf{v}_i^{(1)}$
5. $\quad \mathbf{h} = (\mathbf{V}_{[i]}^{(1)})^{\mathrm{T}}\mathbf{r}$
6. $\quad \mathbf{r} = \mathbf{r} - \mathbf{V}_{[i]}^{(1)}\mathbf{h}$
7. $\quad$ if $\|\mathbf{r}\|_2 = 0$, deflation
8. $\quad \mathbf{v}_{i+1}^{(1)} = \mathbf{r}/\|\mathbf{r}\|_2$
9. **end**
10. $\mathbf{G} = \mathbf{A}^{-1}\mathbf{N}\mathbf{V}_{[p_2]}^{(1)}$
11. $\mathbf{V}^{(2)} = \mathbf{orth}(\mathbf{G})$
12. **for** $i = 1 : p_2(q_2 - 1)$
13. $\quad \mathbf{r} = \mathbf{A}^{-1}\mathbf{v}_i^{(2)}$
14. $\quad \mathbf{h} = \left(\mathbf{V}_{[p_2+i-1]}^{(2)}\right)^{\mathrm{T}}\mathbf{r}$
15. $\quad \mathbf{r} = \mathbf{r} - \mathbf{V}_{[p_2+i-1]}^{(2)}\mathbf{h}$
16. $\quad$ if $\|\mathbf{r}\|_2 = 0$, deflation
17. $\quad \mathbf{v}_{p_2+i}^{(2)} = \mathbf{r}/\|\mathbf{r}\|_2$
18. **end**
19. $\mathbf{V} = \mathbf{orth}([\mathbf{V}^{(1)}\mathbf{V}^{(2)}])$

The following comments are in order.

- Steps 1 to 9 consist of the well-known classical Gram–Schmidt process to generate an orthonormal basis $\mathbf{V}^{(1)}$ of the Krylov subspace $\mathscr{K}_{q_1}(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{b})$. They are also known as the Arnoldi process. In this case, only the basis vectors are saved. Similarly, steps 11 to 18 consist of the block classical Gram–Schmidt process, executed sequentially, to generate an orthonormal basis $\mathbf{V}^{(2)}$ of the Krylov subspace $\mathscr{K}_{q_2}\left(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{N}\mathbf{V}_{[p_2]}^{(1)}\right)$.
- In steps 11 and 19, we use function $\mathbf{orth}(\mathbf{X})$ to stand for the Gram–Schmidt process or QR decomposition for generating an orthonormal basis for the range of $\mathbf{X}$.
- The matrix-vector products in steps 1, 4, 10 and 13 should be implemented by solving the linear systems of equations with the coefficient matrix $\mathbf{A}$.
- When the deflation happens at step 7 or 16, step 8 or 17 is skipped. In a working program, the deflation should be properly handled, see for example, the `gsreorthog` procedure in [25, p. 287].

Once the orthonormal basis $\mathbf{V}$ is generated, the reduced system (14) is derived by computing $\widehat{\mathbf{A}}$, $\widehat{\mathbf{N}}$, $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ as defined in (15). Substantial computational savings could be obtained by exploiting the structures of $\mathbf{A}$ and $\mathbf{N}$ in a working program. The

explicit inverse in defining $\widehat{\mathbf{A}}$ can be avoided if we work with the reduced bilinear system as in the form (13).

5. Numerical examples

In this section, we present two case studies for the model reduction of nonlinear systems based on bilinearization approach presented in the previous sections. All numerical simulations were run in Matlab on a SUN 440 MHz Ultra 10 workstation. We used Matlab's `ode15s` for solving ordinary differential equations under investigation.

Example 1. We use an electrostatic gap-closing actuator, a frequently used micromachined device shown in Fig. 1, as an example to compare the accuracy of our approach presented in Section 4 and Phillips' method [17,18] as described at the end of Section 3.

The governing equation for simulating the dynamical response of the actuator in SUGAR 2.0 [4] is given by

$$\begin{cases} \mathbf{M}\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{b}u(\mathbf{q}, t) \\ y = \mathbf{l}^{\mathrm{T}}\mathbf{q}, \end{cases} \tag{21}$$

where $\mathbf{q}$ is a state vector of length N_o. $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}$ are the $N_o \times N_o$ multi-energy domain system matrices, which are analogous to the mass, damping and stiffness of a purely mechanical system. $u(\mathbf{q}, t)$ is the input excitation source including nonlinear electrostatic force. $y(t)$ is the output of the system. Vector $\mathbf{b}$ is an input influence array to indicate the position input excitation. $\mathbf{l}$ is chosen to extract the components of the state vector of interest.

In this example, a 2 μm by 100 μm flexure is attached to a 5 μm by 100 μm moveable plate that extends from node b to node c. The parallel plate approximation is used to calculate the total force to the plate. $\mathbf{b}$ only has two ones in the components corresponding to the direction of force at nodes b and c. $\mathbf{l}$ only has one in the component corresponding to the node c.

$$u(\mathbf{q}, t) = -\frac{1}{4}\epsilon_0 A \frac{v(t)^2}{\mathrm{gap}(\mathbf{q})^2},$$

where ϵ_0 is the permitivity of free space, $\mathrm{gap}(\mathbf{q})$ is the distance between the 2 plate electrodes, A is the area facing the gap, and $v(t)$ is the voltage between the electrodes.

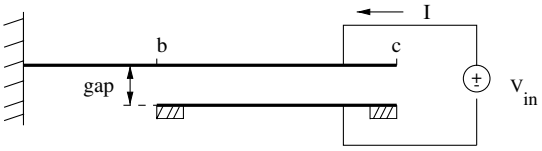

Fig. 1. An electrostatic gap-closing actuator.

M and **K** are derived using linear beam theory, while **D** is based on simple Couette damping and is proportional to **M**. All structures are fabricated in a 2 μm polysilicon layer. The order of the model (21) is $N_o = 30$.

We now show an approximation of the model (21) by a bilinear model (1). First, we note that the model (21) can be equivalently cast in the following form:

$$\begin{cases} \mathbf{C}\dot{\mathbf{x}} + \mathbf{G}\mathbf{x} = \mathbf{b}\,u(\mathbf{x}, t), \\ y = \mathbf{l}^{\mathrm{T}}\mathbf{x}, \end{cases} \tag{22}$$

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \dot{\mathbf{q}} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{D} & \mathbf{M} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix}, \quad \mathbf{b} := \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{l} := \begin{bmatrix} \mathbf{l} \\ \mathbf{0} \end{bmatrix}.$$

Furthermore, the system (22) can be rewritten in the form of (2). By a linearization near an equilibrium point $\mathbf{x}_e$ of the state evolution function, we obtain a bilinear system of the form (1).

In numerical simulation, a piecewise linear voltage function $v(t)$ is applied across the gap. The voltage $v(t)$ ramps from 5 V at $t = 10$ μs to 12 V at $t = 500$ μs, and then drops to 0 V. The displacement component of node c in the direction of force is observed. The initial voltage step causes the device to oscillate. As the voltage increases at a linear rate, the gap decreases at a nonlinear rate due to the electrostatic force increasing proportionally to $1/\mathrm{gap}^2$. This force also causes the period of oscillation to increase. Once the voltage is removed, the actuator exponentially decays back to equilibrium due to viscous air damping.

This phenomenon is captured in the numerical simulation as shown in Figs. 2 and 3. Fig. 2 shows the displacements over the time interval $[0, 3 \times 10^{-4}]$ by three different approaches, namely, solving the original system (22), solving the reduced bilinear system $\widehat{\Sigma}$ generated by the algorithm described in Section 4 and Phillips' method [17,18] as described at the end of Section 3. The reduced system $\widehat{\Sigma}$ (6) by the new algorithm faithfully reproduces the displacement behavior of the original system Σ. In contrast, Phillips' method does poorly. Fig. 3 shows the decay back to equilibrium simulated by the reduced system (6). Comparing with the results reported in [4,2], Fig. 3 shows that the reduced bilinear system captures the essential behavior of the original system. Just for the record, it took 16 seconds to produce the results shown Fig. 3. On the other hand, after more than 10 h, we still could not solve the original system over the same time interval.

Example 2. We consider a nonlinear system of the form

$$\begin{cases} \dot{\mathbf{v}}(t) = \mathbf{f}(\mathbf{v}(t)) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}^{\mathrm{T}}\mathbf{v}(t), \end{cases} \tag{23}$$

with initial condition $\mathbf{v}(0) = \mathbf{v}_0$, where $\mathbf{v} \in \mathscr{R}^{N_o}$ is the state variables. $u(t)$ and $y(t)$ are input and output functions, respectively. $\mathbf{b} \in \mathscr{R}^{N_o}$ is the input distribution array. $\mathbf{c} \in \mathscr{R}^{N_o}$ is the output measurement array. We assume that the nonlinear state evolution

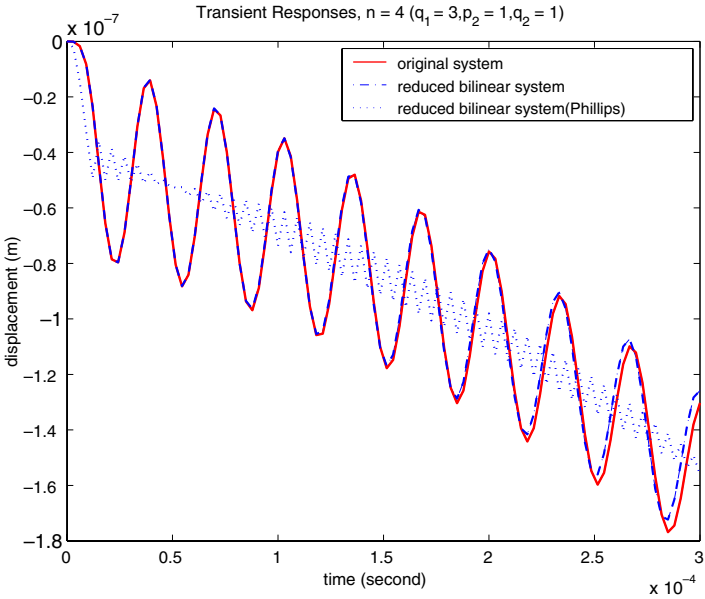

Fig. 2. Displacements of node c computed based the full system, and reduced ones by our method and Phillips' method.

function $\mathbf{f}(\mathbf{x}) : \mathscr{R}^{N_o} \to \mathscr{R}^{N_o}$ possess a sufficient degree of smoothness, and has an equilibrium. Without loss of generality we take this equilibrium at $\mathbf{0}$, i.e., $\mathbf{f}(\mathbf{0}) = \mathbf{0}$.

Suppose that the power series expansion of $\mathbf{f}(\mathbf{x})$ about the equilibrium point $\mathbf{0}$ is written as

$$\mathbf{f}(\mathbf{v}) = \mathbf{A}_1\mathbf{v} + \mathbf{A}_2(\mathbf{v} \otimes \mathbf{v}) + \mathbf{A}_3(\mathbf{v} \otimes \mathbf{v} \otimes \mathbf{v}) + \cdots, \tag{24}$$

where $\mathbf{A}_1 \in \mathscr{R}^{N_o \times N_o}$ is the Jacobian or the first derivative of $\mathbf{f}$, and $\mathbf{A}_2 \in \mathscr{R}^{N_o \times N_o^2}$ is the second derivative matrix of $\mathbf{f}$, and so on. $\otimes$ is the Kronecker product. The linearization of the nonlinear system (23) simply takes the first term in (24) and yields a linear system of the form

$$\begin{cases} \dot{\mathbf{v}}(t) = \mathbf{A}_1\mathbf{v}(t) + \mathbf{b}u(t), \\ y_\ell(t) = \mathbf{c}^{\mathrm{T}}\mathbf{v}(t). \end{cases} \tag{25}$$

If we consider the first two terms of the expansion (24), then by Carleman bilinearization up to the second order, the nonlinear system (23) can be approximated by the following bilinear system:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{N}\mathbf{x}(t)u(t) + \mathbf{b}u(t), \\ y_b(t) = \mathbf{c}^{\mathrm{T}}\mathbf{x}(t), \end{cases} \tag{26}$$

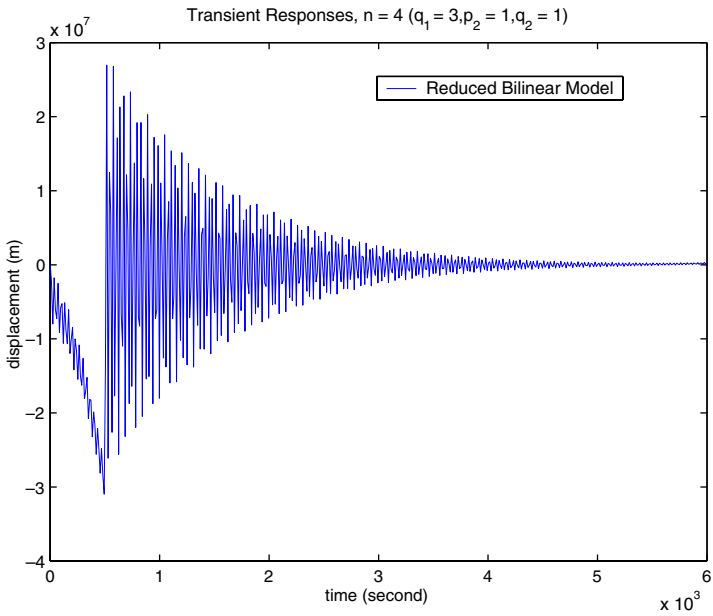

Fig. 3. Displacement of node c, back to equilibrium simulated by the reduced system.

where

$$\mathbf{x} = \begin{bmatrix} \mathbf{v} \\ \mathbf{v} \otimes \mathbf{v} \end{bmatrix}, \quad \mathbf{b} := \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{c} := \begin{bmatrix} \mathbf{c} \\ \mathbf{0} \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{A}_1 \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}_1 \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{b} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{b} & \mathbf{0} \end{bmatrix}.$$

For the details of Carleman bilinearization process, see [21,22]. We note that the dimension of the state space of the resulting bilinear system (26) is $N = N_o + N_o^2$, which is significantly higher than the state dimension of the original system (23). This is the major obstacle to using the Carleman bilinearization for a very large nonlinear system with high degree of nonlinearity necessary. However, in the following, we show that a model reduction technique, we are able to obtain a satisfactory reduced bilinear system of the state dimension n even much smaller than the original dimension N_o.

Let us consider an RC circuit with nonlinear resistors and an independent current source proposed by Chen and White [3], see Fig. 4. Let $u(t)$ be the input-signal to the independent current source and $v_1, v_2, \ldots, v_{N_o}$ be the node voltages. By applying Kirchhoff's current law and assuming $C = 1$ for each capacitor, we obtain the nonlinear system (23), where

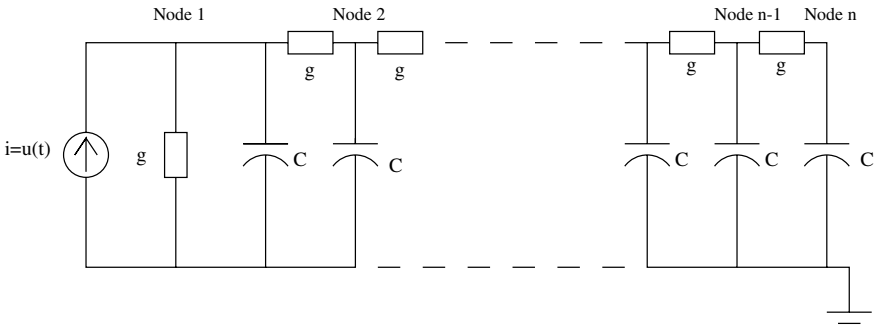

Fig. 4. A nonlinear RC circuit [3].

$$\mathbf{f}(\mathbf{v}) = [f_k(\mathbf{v})] = \begin{bmatrix} -g(v_1) - g(v_1 - v_2) \\ g(v_1 - v_2) - g(v_2 - v_3) \\ \vdots \\ g(v_{k-1} - v_k) - g(v_k - v_{k+1}) \\ \vdots \\ g(v_{N_o-1} - v_{N_o}) \end{bmatrix}, \quad \mathbf{b} = \mathbf{c} := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The output signal $y(t)$ is the voltage between node 1 and ground. The current through each resistor will have the following voltage dependence:

$$g(v) = \exp(40v) + v - 1. \tag{27}$$

With the second-order approximation of $g(v)$, the first component of $\mathbf{f}$ $(\mathbf{v})$ can be written as

$$f_1(\mathbf{v}) = -82v_1 + 41v_2 - 1600v_1^2 + 800v_1v_2 + 800v_2v_1 - 800v_2^2 + \cdots. \tag{28}$$

The second component of $\mathbf{f}$ $(\mathbf{v})$ is

$$\begin{aligned} f_2(\mathbf{v}) = {} & 41v_1 - 82v_2 + 41v_3 + 800v_1^2 - 800v_1v_2 \\ & - 800v_2v_1 + 800v_2v_3 + 800v_3v_2 - 800v_3^2 + \cdots. \end{aligned} \tag{29}$$

In general, the kth component of $\mathbf{f}(\mathbf{v})$ can be written as

$$\begin{aligned} f_k(\mathbf{v}) = {} & 41v_{k-1} - 82v_k + 41v_{k+1} + 800v_{k-1}^2 - 800v_{k-1}v_k \\ & - 800v_kv_{k-1} + 800v_kv_{k+1} + 800v_{k+1}v_k - 800v_{k+1}^2 + \cdots. \end{aligned} \tag{30}$$

Finally, the last Nth component of $\mathbf{f}(\mathbf{v})$ is

$$f_{N_o}(\mathbf{v}) = 41v_{N_o-1} - 41v_{N_o} + 800v_{N_o-1}^2 - 800v_{N_o-1}v_{N_o} - 800v_{N_o}v_{N_o-1} + 800v_{N_o}^2 + \cdots. \tag{31}$$

These expressions can be rewritten in matrix notation in the expansion form (24). The coefficient matrix $\mathbf{A}_1$ of the first term is given by

$$\mathbf{A}_1 = \begin{bmatrix} -82 & 41 & & & \\ 41 & -82 & 41 & & \\ & \ddots & \ddots & \ddots & \\ & & 41 & -82 & 41 \\ & & & 41 & -41 \end{bmatrix}.$$

Now let us consider the coefficient matrix $\mathbf{A}_2$ of the second term in the expansion (24). Note that $\mathbf{A}_2$ is of dimension $N_o \times N_o^2$. If we order the second order terms $v_i v_j$ in the vector $\mathbf{v} \otimes \mathbf{v}$ as the $((i-1)N_o + j)$th component, then from (28)–(31) we see that the nonzero entries of $\mathbf{A}_2$ are as follows:

- in the first row,

$$\mathbf{A}_2(1,1) = -1600, \quad \mathbf{A}_2(1,2) = 800,$$
$$\mathbf{A}_2(1,N_o+1) = 800, \quad \mathbf{A}_2(1,N_o+2) = -800;$$

- in the second row,

$$\mathbf{A}_2(2,1) = 800, \quad \mathbf{A}_2(2,2) = -800,$$
$$\mathbf{A}_2(2,N_o+1) = -800, \quad \mathbf{A}_2(2,N_o+3) = 800,$$
$$\mathbf{A}_2(2,2N_o+2) = 800, \quad \mathbf{A}_2(2,2N_o+3) = -800;$$

- in general, in the kth row, where $2 < k < N_o - 1$,

$$\mathbf{A}_2(k,(k-2)N_o+k-1) = 800, \quad \mathbf{A}_2(k,(k-2)N_o+k) = -800,$$
$$\mathbf{A}_2(k,(k-1)N_o+k-1) = -800, \quad \mathbf{A}_2(k,(k-1)N_o+k+1) = 800,$$
$$\mathbf{A}_2(k,kN_o+k) = 800, \quad \mathbf{A}_2(k,kN_o+k+1) = -800;$$

- and in the last row

$$\mathbf{A}_2(N_o,(N_o-2)N_o+N_o-1) = 800, \quad \mathbf{A}_2(N_o,(N_o-2)N_o+N_o) = -800,$$
$$\mathbf{A}_2(N_o,(N_o-1)N_o+N_o-1) = -800, \quad \mathbf{A}_2(N_o,(N_o-1)N_o+N_o) = 800.$$

We observe that the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are extremely sparse. For example, the number of nonzero entries of the $N_o \times N_o^2$ matrix $\mathbf{A}_2$ is approximately equal to $6N_o$. Furthermore, the matrices $\mathbf{A}$ and $\mathbf{N}$ in the bilinear system (26) are highly structured. These properties can be exploited in depth for computational efficiency in both storage and CPU time. A detailed discussion on these issues is beyond the scope of this paper.

In our numerical simulation, we used $N_o = 200$. As a result, the bilinear system (26) is of the dimension $N = N_o + N_o^2 = 40,200$. Figs. 5 and 6 show the outputs $y(t)$ for input functions $u(t) = \mathrm{e}^{-t}$ and $u(t) = (\cos(2\pi t/10) + 1)/2$, respectively. Numerical results of four approaches are reported in these figures. The first approach is to solve the original system (23) directly. The second approach is to use the linearized system (25). The third and fourth approaches are model reduction algorithms, namely, the new algorithm presented in Section 4, and Phillips' method [17,18]. We highlight the following two observations: (a) The reduced bilinear system approach is significantly better than the linearization and the new algorithm is more accurate than Phillips' method. (b) As a result of Carleman bilinearization, the state dimension of the bilinear system (26) is increased dramatically. For example, when the state dimension of the original system is $N_o = 200$, the state dimension of the bilinear system at the second-order approximation is increased to $N = 40{,}200$. Fortunately, the state dimension of the reduced systems can be in fact very small, even much smaller than the original system N_o. In this example, the reduced system of order 21 essentially reproduces the output behaviors of the order original system of order 200. Finally, we note that in our experiments, we observed that the stability of the reduced system defined by Phillips' method is highly sensitive to the choice of the parameters q_1, q_2 and p_2, as illustrated in Fig. 7.

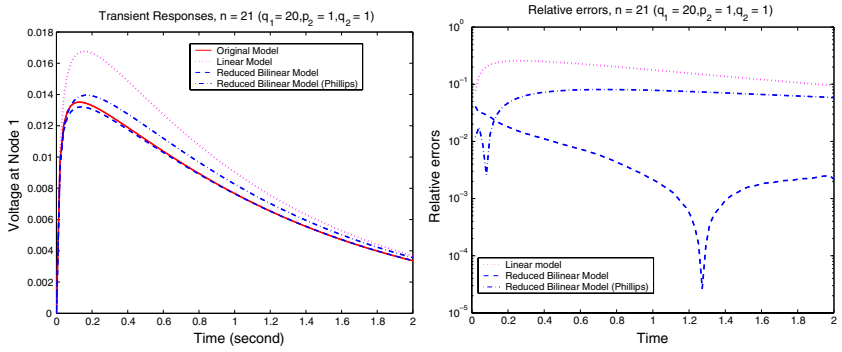

Fig. 5. Left: Output $y(t)$ of the nonlinear RC circuit for $u(t) = \mathrm{e}^{-t}$. Right: Relative errors with respect to the original system.

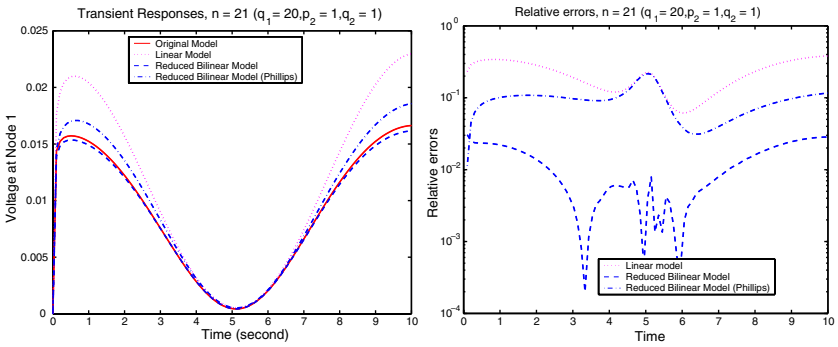

Fig. 6. Left: Output $y(t)$ of the nonlinear RC circuit for $u(t) = (\cos(2\pi t/10) + 1)/2$. Right: Relative errors with respect to the original system.

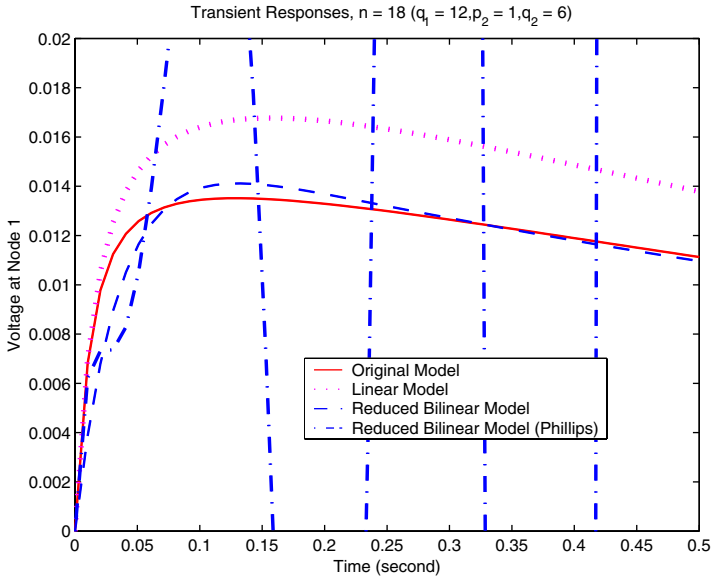

Fig. 7. Sensitivity of Phillips' method to the parameters $p_1, q_2, p_2, u(t) = \mathrm{e}^{-t}$.

6. Concluding remarks

A Krylov subspace based projection method is presented for reduced-order modeling of the bilinear system Σ. Although the system matrices as defined in (14) for the reduced system $\widetilde{\Sigma}$ is unconventional compared to the one proposed in [17,18], we have shown that they are the right ones to use for satisfying the desired moment

preservation and transfer function approximation (7) and (8). We have demonstrated the advantages of this new definition in two applications. It is an open question to connect the accuracy of the output approximation with the optimal degrees of transfer function approximations and the corresponding number of matching-moments.

Carleman bilinearization is a systematical way to approximate a nonlinear system, such as in the form of (2), to a desired degree of nonlinearity. However, the order of the resulting bilinear system increases dramatically. It is essential to exploit the underlying structure. Memory requirement may eventually become a bottleneck. Recently, a piecewise-linear system approximation of a nonlinear system is presented in [19]. It is an interesting project to compare these two approaches.

Acknowledgment

We thank the referees for valuable comments and suggestions to improve the presentation of the paper. Support for this work has been provided in part by NSF ITR grant ACI-0220104. Skoogh was also supported in part by the Foundation BLANCEFLOR Boncompagni-Ludovisi, neé Bildt and the Royal Swedish Academy of Sciences while visiting University of California, Davis.

References

[1] Z. Bai, Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems, Appl. Numer. Math. 43 (2002) 9–44.
[2] Z. Bai, D. Bindel, J. Clark, J. Demmel, K.S.J. Pister, N. Zhou, New numerical techniques and tools in SUGAR for 3D MEMS simulation, in: Technical Proceedings of the Fourth International Conference on Modeling and Simulation of Microsystems, 2000, pp. 31–34.
[3] Y. Chen, J. White, A quadratic method for nonlinear model order reduction, in: International Conference on Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators, San Diego, 2000, pp. 477–480.
[4] J.V. Clark, N. Zhou, D. Bindel, L. Schenato, W. Wu, J. Demmel, K.S.J. Pister, 3D MEMS simulation using modified nodal analysis, in: Proceedings of Microscale Systems: Mechanics and Measurements Symposium, 2000, pp. 68–75.
[5] P. Feldman, R.W. Freund, Efficient linear circuit analysis by Padé approximation via the Lanczos process, IEEE Trans. Computer-Aided Design CAD-14 (1995) 639–649.
[6] R.W. Freund, Reduced-order modeling techniques based on Krylov subspaces and their use in circuit simulation, in: B.N. Datta (Ed.), Applied and Computational Control, Signals, and Circuits, vol. 1, Birkhäuser, Boston, 1999, pp. 435–498.
[7] R.W. Freund, P. Feldman, Small-signal circuit analysis and sensitivity computations with the PVL algorithm, IEEE Trans. Circuits Systems II 43 (1996) 577–585.
[8] E. Grimme, D.C. Sorensen, P. Van Dooren, Model reduction of state space system via an implicitly restarted Lanczos method, Numer. Algorithms 12 (1996) 1–31.
[9] P.K. Gunupudi, M. Nakhla, Nonlinear circuit-reduction of high-speed interconnect network using congruent transformation techniques, IEEE Trans. Advanced Packaging 24 (2001) 317–325.
[10] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge Univ. Press, Cambridge, 1996.

[11] T. Kailath, Linear Systems, Prentice-Hall, New York, 1980.
[12] S. Lall, J.E. Marsden, S. Glavaski, A subspace approach to balanced truncation for model reduction of nonlinear control systems, Internat. J. Rubust Nonlinear Control 12 (2002) 519–535.
[13] M. Rathinam, L.R. Petzold, A new look at proper orthogonal decomposition, SIAM J. Numer. Anal. 41 (2003) 1893–1925.
[14] R.R. Mohler, Nonlinear Systems, Dynamics and Control, vol. I; Applications to Bilinear Control, vol. II, Prentice Hall, Englewood Cliffs, NJ, 1991.
[15] T. Mukherjee, G. Fedder, D. Ramaswamy, J. White, Emerging simulation approaches for micromachined devices, IEEE Trans. CAD 19 (2000) 1572–1588.
[16] A. Odabasioglu, M. Celik, L.T. Pileggi, PRIMA: passive reduced-order interconnect macromodeling algorithm, in: Technical Digest 1997, IEEE/ACM International Conference on Computer-Aided Design, IEEE, 1997, pp. 58–65.
[17] J. Phillips, Projection frameworks for model reduction of weakly nonlinear systems, in: Proceedings of DAC 2000, 2000, pp. 184–189.
[18] J. Phillips, Projection-based approaches for model reduction of weakly nonlinear, time-varying systems, IEEE Trans. Computer-Aided Design Integrated Circuits Systems 22 (2003) 171–187.
[19] M. Rewieński, J. White, A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuit and micromachined devices, IEEE Trans. Computer-Aided Design Integrated Circuits Systems 22 (2003) 155–170.
[20] J. Roychowdhury, Reduced-order modeling time-varying systems, IEEE Trans. Circuits Systems II 46 (1999) 1273–1288.
[21] W.J. Rugh, Nonlinear System Theory, The John Hopkins University Press, Boltimore, 1981.
[22] S. Sastry, Nonlinear Systems: Analysis, Stability and Control, Springer, New York, 1999.
[23] J.M.A. Scherpen, Balancing for nonlinear systems, Systems Control Lett. 21 (1993) 143–153.
[24] S.D. Senturia, Microsystem Design, Kluwer Academic Publishers, Boston, 2001.
[25] G.W. Stewart, Matrix Algorithms, Basic Decompositions, vol. I, SIAM, Philadelphia, 1998.