

# Krylov Subspace Methods for Model Order Reduction of Bilinear Control Systems

Tobias Breiten<sup>\*,a,1</sup>, Tobias Damm<sup>a,1</sup>

<sup>a</sup>*Mathematics Department, TU Kaiserslautern, Germany*

---

## Abstract

We discuss the use of Krylov subspace methods with regard to the problem of model order reduction. The focus lies on bilinear control systems, a special class of nonlinear systems, which are closely related to linear systems. While most existent approaches are based on series expansions around zero, we will extend the underlying ideas to a more general context and show that there exist several ways to reduce bilinear systems. Besides, we will briefly address the problem of stability preserving model reduction and further explain the benefit of using two-sided projection methods. By means of some numerical examples, we will illustrate the performance of the presented reduction methods.

*Key words:* Bilinear systems, Multimoment-matching, Krylov subspaces, Model order reduction, Carleman bilinearization

---

## 1. Introduction

Detailed dynamical models of complex systems frequently involve a very large number of coupled first order differential or difference equations, while describing only the transfer behaviour between a few inputs and outputs. Since the numerical simulation of such a system may be prohibitively expensive, one tries to approximate the transfer behaviour by a reduced-order system, i.e. a system with less equations.

For linear systems, the problem is well-studied (see e.g. [1]). Generally speaking, there are two types of methods, one being based on the singular value decomposition, the other using Krylov subspace projections.

Unfortunately, practically all real-life dynamics include nonlinearities, and theory as well as reduction techniques are less developed for nonlinear problems. In this paper we deal with Krylov subspace projection methods for bilinear control systems, which have been initiated rather recently in e.g. [2, 5, 9, 10, 6].

In general, a system of the form

---

\*Corresponding author

*Email addresses:* Tobias.Breiten@gmx.de (Tobias Breiten),  
damm@mathematik.uni-kl.de (Tobias Damm)

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + \sum_{j=1}^m N_j x(t) u_j(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = x_0, \end{cases}$$

with  $A, N_j \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{p \times n}$  is called a continuous time-invariant bilinear control system. Here,  $x(t) \in \mathbb{R}^n$  is the  $n$ -dimensional state of the system,  $u(t) = [u_1(t), \dots, u_m(t)]^T \in \mathbb{R}^m$  is the  $m$ -dimensional input and  $y(t) \in \mathbb{R}^p$  is the  $p$ -dimensional output. Bilinear control systems represent a special class of nonlinear systems that are linear in input and linear in state, but not jointly linear in state and input. They exhibit a close connection to linear systems which can be exploited in order to transfer and generalize, respectively, successful linear reduction techniques. A detailed analysis of bilinear systems can be found in [4] and [8].

Our goal will be the construction of a reduced system

$$\hat{\Sigma} : \begin{cases} \dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \sum_{j=1}^m \hat{N}_j \hat{x}(t) u_j(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t) \end{cases}$$

with  $\hat{A}, \hat{N}_j \in \mathbb{R}^{\hat{n} \times \hat{n}}$ ,  $\hat{B} \in \mathbb{R}^{\hat{n} \times m}$ ,  $\hat{C} \in \mathbb{R}^{p \times \hat{n}}$  such that  $\hat{y}(t) \approx y(t)$  and  $\hat{n} \ll n$ .

We proceed as follows. After briefly reviewing the Volterra series representation and transfer functions of a bilinear system, we introduce the concept of multimoment-matching for different kinds of expansion points. Together with an oblique projection method, this leads to various model reduction approaches, and we discuss their relation to existing methods. In Section 3, we propose a stability-preserving reduction technique and analyze the use of two-sided projection methods. In the last section, we illustrate our results using examples of flow model and a nonlinear electric circuit.

## 2. Multimoment-Matching

For simplicity of notation, we concentrate on single-input/single-output systems with zero initial condition

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Nx(t)u(t) + bu(t), \\ y(t) = c^T x(t), \quad x(0) = 0. \end{cases} \quad (1)$$

However, all our considerations can easily be extended to the general case.

The input-output behaviour of  $\Sigma$  can be described by the infinite sum of

convolution integrals

$$y(t) = \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} \underbrace{c^T e^{At_k} N \dots e^{At_2} N e^{At_1} b}_{h(t_1, \dots, t_k)} u(t - t_1 - \dots - t_k) \dots \cdot u(t - t_k) dt_k \dots dt_1$$

The explicit derivation of this so-called *Volterra series representation* is based on the Picard iteration and can be found in [11]. For  $N = 0$  we obtain the familiar impulse response of a linear system. Each term in the sum may be interpreted as the response of a certain subsystem. Accordingly, the result of a multivariable Laplace transform of the *degree- $k$*  kernel  $h(t_1, \dots, t_k)$  might be called the  $k$ -th transfer function

$$H(s_1, \dots, s_k) = c^T (s_k I - A)^{-1} N \dots (s_2 I - A)^{-1} N (s_1 I - A)^{-1} b$$

Taking arbitrary  $\sigma_i \in \mathbb{R}$ ,  $i = 1, \dots, k$  and using the general left product convention  $\prod_{j=2}^k A_j = A_k \prod_{j=2}^{k-1} A_j$ , we can rewrite this as

$$\begin{aligned} H(s_1, \dots, s_k) &= c^T \left( \prod_{j=2}^k (s_j I - A)^{-1} N \right) (s_1 I - A)^{-1} b \\ &= (-1)^k c^T \left( \prod_{j=2}^k (A - \sigma_j I - (s_j - \sigma_j) I)^{-1} N \right) \\ &\quad \cdot (A - \sigma_1 I - (s_1 - \sigma_1) I)^{-1} b \\ &= (-1)^k c^T \left( \prod_{j=2}^k (I - (s_j - \sigma_j) (A - \sigma_j I)^{-1})^{-1} (A - \sigma_j I)^{-1} N \right) \\ &\quad \cdot (I - (s_1 - \sigma_1) (A - \sigma_1 I)^{-1})^{-1} (A - \sigma_1 I)^{-1} b \end{aligned}$$

Using Neumann expansions for  $s_j$  around  $\sigma_j$  we can substitute

$$\left( I - (s_j - \sigma_j) (A - \sigma_j I)^{-1} \right)^{-1} = \sum_{i=0}^{\infty} (s_j - \sigma_j)^i (A - \sigma_j I)^{-i}$$

and obtain

$$\begin{aligned} H(s_1, \dots, s_k) &= (-1)^k c^T \left( \prod_{j=2}^k \left( \sum_{i=0}^{\infty} (s_j - \sigma_j)^i (A - \sigma_j I)^{-(i+1)} \right) N \right) \\ &\quad \cdot \left( \sum_{i=0}^{\infty} (s_1 - \sigma_1)^i (A - \sigma_1 I)^{-(i+1)} \right) b. \end{aligned}$$

Finally, the use of a multivariable power series notation leads to

$$H(s_1, \dots, s_k) = \sum_{l_k=1}^{\infty} \dots \sum_{l_1=1}^{\infty} m(l_1, \dots, l_k) (s_1 - \sigma_1)^{l_1-1} \dots (s_k - \sigma_k)^{l_k-1},$$

where

$$m(l_1, \dots, l_k) = (-1)^k c^T (A - \sigma_k I)^{-l_k} N \dots (A - \sigma_2 I)^{-l_2} N (A - \sigma_1 I)^{-l_1} b$$

are called *multimoments* corresponding to the  $k$ -th transfer function.

In our considerations, we assume that the first  $k-1$  expansion points of the  $k$ -th subsystem coincide with those of the preceding subsystem. This is a limitation induced by the recursive nature of the construction. But, it is far more general than the common choice of *low frequency multimoments*  $\sigma_1 = \dots = \sigma_k = 0$ .

A slightly different procedure with  $s_i = \xi_i^{-1}$  also allows to expand the transfer functions at  $\sigma_i = \infty$ ,

$$\begin{aligned} H(s_1, \dots, s_k) &= c^T (s_k I - A)^{-1} N \dots (s_2 I - A)^{-1} N (s_1 I - A)^{-1} b \\ &= c^T (\xi_k^{-1} I - A)^{-1} N \dots (\xi_2^{-1} I - A)^{-1} N (\xi_1^{-1} I - A)^{-1} b \\ &= c^T \xi_k (I - \xi_k A)^{-1} N \dots \xi_2 (I - \xi_2 A)^{-1} N \xi_1 (I - \xi_1 A)^{-1} b \end{aligned}$$

Again, for  $\xi$  close to 0 ( $s_i$  close to  $\infty$ ) we can make use of the Neumann expansion and obtain

$$\begin{aligned} H(s_1, \dots, s_k) &= \sum_{l_k=1}^{\infty} \dots \sum_{l_1=1}^{\infty} m(l_1, \dots, l_k) \xi_1^{l_1} \dots \xi_k^{l_k} \\ &= \sum_{l_k=1}^{\infty} \dots \sum_{l_1=1}^{\infty} m(l_1, \dots, l_k) s_1^{-l_1} \dots s_k^{-l_k}, \end{aligned}$$

where

$$m(l_1, \dots, l_k) = c^T A^{l_k-1} N \dots A^{l_2-1} N A^{l_1-1} b$$

are called *high frequency multimoments*. These obviously generalize the *Markov parameters* of a linear system. Altogether, we record that the multimoments characterize the system output at least locally. Hence, if we construct a reduced bilinear system that preserves certain predefined multimoments, then we expect it to approximate the original system in some way, although it is still unclear, what a suitable choice of the expansion points would be.

Assuming that we want to match  $q+q^2+\dots+q^r$  multimoments corresponding to the first  $r$  subsystems, we can now state our problem formally.

For a given bilinear system  $\Sigma$  we want to find another system  $\hat{\Sigma}$  with state dimension  $\hat{n} \ll n$ , such that

$$m(l_1, \dots, l_k) = \hat{m}(l_1, \dots, l_k),$$

for  $k = 1, \dots, r$  and  $l_j = q, j = 1, \dots, k$ . This implies the following approximation orders in terms of the transfer functions:

$$\sigma_i \neq \infty: H(s_1, \dots, s_k) = \hat{H}(s_1, \dots, s_k) + \mathcal{O}((s_1 - \sigma_1)^{p_1} \dots (s_k - \sigma_k)^{p_k}),$$

$$\sigma_i = \infty: H(s_1, \dots, s_k) = \hat{H}(s_1, \dots, s_k) + \mathcal{O}\left(s_1^{-(p_1+1)} \dots s_k^{-(p_k+1)}\right),$$

where  $p_j \leq q$  and at least one of them is equal to  $q$ .

### 3. Model Order Reduction

Together with system (1) we consider its *Petrov-Galerkin projection* of the form

$$\hat{\Sigma} : \begin{cases} \hat{x}(t) = \underbrace{W^T AV}_{\hat{A}} \hat{x}(t) + \underbrace{W^T NV}_{\hat{N}} \hat{x}(t)u(t) + \underbrace{W^T b}_{\hat{b}} u(t), \\ \hat{y}(t) = \underbrace{c^T V}_{\hat{c}^T} \hat{x}(t), \quad W, V \in \mathbb{R}^{n \times k} \end{cases}$$

For later reference, we record an obvious property of projection operators.

**Lemma 1.** *Let  $z \in \text{span}\{V\}$ ,  $V \in \mathbb{R}^{n \times k}$  and  $W^T V = I_k$ . Then  $z = VW^T z$ .*

#### One-Sided Projection Methods

Recall that the  $q$ -th (block) Krylov subspace corresponding to a matrix  $A \in \mathbb{R}^{n \times n}$  and a set of starting vectors  $B = [b_1, \dots, b_m] \in \mathbb{R}^{n \times m}$  is defined as follows:

$$\mathcal{K}_q(A, B) = \text{span}\{b_1, \dots, b_m, Ab_1, \dots, Ab_m, \dots, A^{q-1}b_1, \dots, A^{q-1}b_m\}.$$

If we now construct  $V$  as the basis of the union of specific Krylov subspaces, a reduced system will match some of the multimoments of the original system.

**Theorem 1.** *Let a bilinear SISO system  $\Sigma$  be given. Let  $\hat{\Sigma}$  be a reduced bilinear system constructed by an oblique projection  $P = VW^T$ ,  $W^T V = I$  with  $V$  given as follows:*

$$\begin{aligned} \text{span}\{V^{(1)}\} &= \mathcal{K}_q((A - \sigma_1 I)^{-1}, (A - \sigma_1 I)^{-1}b), \\ \text{span}\{V^{(k)}\} &= \mathcal{K}_q((A - \sigma_k I)^{-1}, (A - \sigma_k I)^{-1}NV^{(k-1)}), \quad k = 2, \dots, r \\ \text{span}\{V\} &= \text{span}\left\{\bigcup_{k=1}^r \text{span}\{V^{(k)}\}\right\}. \end{aligned}$$

*Then for  $k = 1, \dots, r$  the first  $q^k$  multimoments of  $\Sigma$  and  $\hat{\Sigma}$ , resulting from series expansions at  $\sigma_i \in \mathbb{R}^n$ ,  $i = 1, \dots, k$ , coincide, i.e.*

$$m(l_1, \dots, l_k) = \hat{m}(l_1, \dots, l_k),$$

*for  $k = 1, \dots, r$  and  $l_1, \dots, l_k = 1, \dots, q$ .*

*Proof.* First, we record that for  $l_1, \dots, l_r = 1, \dots, q$  the following vectors are contained in  $\text{span } V$ :

$$\begin{aligned} &(A - \sigma_1 I)^{-l_1} b, \\ &(A - \sigma_2 I)^{-l_2} N (A - \sigma_1 I)^{-l_1} b, \\ &\quad \dots, \\ &(A - \sigma_r I)^{-l_r} N \dots (A - \sigma_2 I)^{-l_2} N (A - \sigma_1 I)^{-l_1} b, \end{aligned}$$

Thus, Lemma 1 applies to these vectors and for the first subsystem we have

$$\begin{aligned}
V(\hat{A} - \sigma_1 I)^{-l_1} \hat{b} &= V(W^T A V - \sigma_1 I)^{-l_1} W^T b && \text{def. of } \hat{A}, \hat{b} \\
&= V(W^T A V - \sigma_1 W^T V)^{-l_1} W^T b && W^T V = I \\
&= V(W^T (A - \sigma_1 I) V)^{-l_1} W^T b \\
&= V(W^T (A - \sigma_1 I) V)^{-l_1} W^T (A - \sigma_1 I) \\
&\quad \cdot (A - \sigma_1 I)^{-1} b \\
&= V(W^T (A - \sigma_1 I) V)^{-l_1} W^T (A - \sigma_1 I) \\
&\quad \cdot V W^T (A - \sigma_1 I)^{-1} b && \text{Lemma 1} \\
&= V(W^T (A - \sigma_1 I) V)^{-l_1} (W^T (A - \sigma_1 I) V) \\
&\quad \cdot W^T (A - \sigma_1 I)^{-1} b \\
&= V(W^T (A - \sigma_1 I) V)^{-(l_1-1)} W^T (A - \sigma_1 I)^{-1} b \\
&\quad \vdots \\
&= V W^T (A - \sigma_1 I)^{-l_1} b \\
&= (A - \sigma_1 I)^{-l_1} b && \text{Lemma 1}
\end{aligned}$$

which, in fact, is well-known from the linear case. Using this result, we obtain for the second subsystem

$$\begin{aligned}
&V(\hat{A} - \sigma_2 I)^{-l_2} \hat{N}(\hat{A} - \sigma_1 I)^{-l_1} \hat{b} \\
&= V(\hat{A} - \sigma_2 I)^{-l_2} W^T N V(\hat{A} - \sigma_1 I)^{-l_1} \hat{b} && \text{def. of } \hat{N} \\
&= V(\hat{A} - \sigma_2 I)^{-l_2} W^T N (A - \sigma_1 I)^{-l_1} b && \text{1st subsystem} \\
&= V(W^T A V - \sigma_2 I)^{-l_2} W^T N (A - \sigma_1 I)^{-l_1} b && \text{def. of } \hat{A} \\
&= V(W^T A V - \sigma_2 W^T V)^{-l_2} W^T N (A - \sigma_1 I)^{-l_1} b && W^T V = I \\
&= V(W^T (A - \sigma_2 I) V)^{-l_2} W^T (A - \sigma_2 I) \\
&\quad \cdot (A - \sigma_2 I)^{-1} N (A - \sigma_1 I)^{-l_1} b \\
&= V(W^T (A - \sigma_2 I) V)^{-l_2} W^T (A - \sigma_2 I) \\
&\quad \cdot V W^T (A - \sigma_2 I)^{-1} N (A - \sigma_1 I)^{-l_1} b && \text{Lemma 1} \\
&= V(W^T (A - \sigma_2 I) V)^{-(l_2-1)} W^T (A - \sigma_2 I)^{-1} N (A - \sigma_1 I)^{-l_1} b \\
&\quad \vdots \\
&= V W^T (A - \sigma_2 I)^{-l_2} N (A - \sigma_1 I)^{-l_1} b \\
&= (A - \sigma_2 I)^{-l_2} N (A - \sigma_1 I)^{-l_1} b && \text{Lemma 1}
\end{aligned}$$

By an analogue induction step we can deal with subsystems of higher order. Multiplication with  $(-1)^k c^T$  then proves the assertion.  $\square$

**Remark 1.** Assuming that all vectors of the Krylov spaces are linearly independent, the reduced system is of dimension  $\hat{n} = q + q^2 + \dots + q^r$ .

**Theorem 2.** *Let a bilinear SISO system  $\Sigma$  be given. If a reduced bilinear system  $\hat{\Sigma}$  is constructed by an oblique projection  $P = VW^T$ ,  $W^T V = I$  with  $V$  given by*

$$\begin{aligned} \text{span}\{V^{(1)}\} &= \mathcal{K}_q(A, b), \\ \text{span}\{V^{(k)}\} &= \mathcal{K}_q(A, NV^{(k-1)}), \quad k = 2, \dots, r \\ \text{span}\{V\} &= \text{span} \left\{ \bigcup_{k=1}^r \text{span}\{V^{(k)}\} \right\}, \end{aligned}$$

then for  $k = 1, \dots, r$  the first  $q^k$  high frequency multimoments of  $\Sigma$  and  $\hat{\Sigma}$  coincide, i.e.

$$c^T A^{l_k-1} N \dots A^{l_2-1} N A^{l_1-1} b = \hat{c}^T \hat{A}^{l_k-1} \hat{N} \dots \hat{A}^{l_2-1} \hat{N} \hat{A}^{l_1-1} \hat{b},$$

for  $k = 1, \dots, r$  and  $l_1, \dots, l_k = 1, \dots, q$ .

*Proof.* Not surprisingly, the proof is almost identical to the proof of the previous theorem. We make use of the fact that the following vectors belong to the subspace spanned by the columns of  $V$ :

$$\begin{aligned} &A^{l_1-1} b, \\ &A^{l_2-1} N A^{l_1-1} b, \\ &\dots, \\ &A^{l_r-1} N \dots A^{l_2-1} N A^{l_1-1} b. \end{aligned}$$

We show the computations exemplarily for the second order multimoments:

$$\begin{aligned} V \hat{A}^{l_2-1} \hat{N} \hat{A}^{l_1-1} \hat{b} &= V \hat{A}^{l_2-1} W^T N V \hat{A}^{l_1-1} \hat{b} && \text{def. of } \hat{N} \\ &= V \hat{A}^{l_2-1} W^T N A^{l_1-1} b && \text{shown before} \\ &= V (W^T A V)^{l_2-1} W^T N A^{l_1-1} b && \text{def. of } \hat{A} \\ &= V (W^T A V)^{l_2-2} W^T A V W^T N A^{l_1-1} b \\ &= V (W^T A V)^{l_2-2} W^T A N A^{l_1-1} b && \text{Lemma 1} \\ &\vdots \\ &= V (W^T A V) W^T A^{l_2-2} N A^{l_1-1} b \\ &= V W^T A^{l_2-1} N A^{l_1-1} b && \text{Lemma 1} \\ &= A^{l_2-1} N A^{l_1-1} b && \text{Lemma 1} \end{aligned}$$

The induction step again is analogue. □

**Remark 2.** *Note that the dimensions of the Krylov subspaces as well as the number of starting vectors do not have to be the same for all  $k$ . However, for reasons of clarity we assumed all of them to be equal to  $q$ .*

Our theorems generalize other results from the literature. To see this consider the original bilinear system,

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Nx(t)u(t) + bu(t), \\ y(t) &= c^T x(t)\end{aligned}$$

If we multiply the state equation by a non-singular matrix  $E \in \mathbb{R}^{n \times n}$  from the left, we obtain:

$$\begin{aligned}E\dot{x}(t) &= EAx(t) + ENx(t)u(t) + Ebu(t), \\ y(t) &= c^T x(t)\end{aligned}$$

We project the state vector  $x$  onto  $V\hat{x}$  and then multiply with  $V^T$  from the left:

$$\begin{aligned}V^T EV\dot{\hat{x}}(t) &= V^T EAV\hat{x}(t) + V^T ENV\hat{x}(t)u(t) + V^T Ebu(t), \\ \hat{y}(t) &= c^T V\hat{x}(t)\end{aligned}$$

If  $V$  has full column rank then  $V^T EV$  is non-singular and the system can be transferred into

$$\begin{aligned}\dot{\hat{x}}(t) &= W^T AV\hat{x}(t) + W^T NV\hat{x}(t)u(t) + W^T bu(t), \\ \hat{y}(t) &= c^T V\hat{x}(t),\end{aligned}\tag{2}$$

with  $W^T = (V^T EV)^{-1}V^T E$ . Hence, the described procedure leads to a reduced system resulting from an oblique projection of the original system. This implies that we do not have to change the construction of the projection matrix  $V$  in order to guarantee multimoment-matching. Moreover, we can premultiply the system by any arbitrary non-singular matrix  $E$ . Let us point out some special cases. If  $E = A^{-1}$ , we end up with the reduction technique proposed by Z. Bai and D. Skoogh. Contrary to the statement in [2], we record that the reduced model is actually constructed by an oblique projection instead of an orthonormal one. The other special case arises for  $E = I$  and orthonormal  $V$ :

$$W^T = (V^T IV)^{-1}V^T I = V^T.$$

Here, the projection indeed is orthogonal and we obtain the approach that was initially introduced by J.R. Phillips in [9]. Finally, the discussion in [5] corresponds to the situation where all multimoments result from the same expansion point, i.e.  $\sigma_1 = \dots = \sigma_k = \sigma$ .

#### *Stability-Preserving Model Order Reduction.*

There are different notions of stability for bilinear systems. In most cases a minimal requirement would be that the uncontrolled system is stable, i.e.  $\sigma(A) \subset \mathbb{C}_-$ . This also implies that  $L^2$ -input signals are mapped to  $L^2$ -output signals.

Unfortunately, projection methods in general do not preserve this property. We have seen, however, that there is no restriction on the projection matrix  $W$  as long as we have  $W^T V = I$ . This freedom may also be exploited to guarantee stability. The following result stems from a similar consideration for linear systems and can be found in [13].

**Theorem 3.** *Given a bilinear system  $\Sigma$  with  $\text{Re}(\lambda_i(A)) < -\sigma \leq 0$  for all  $i$ . Let  $V \in \mathbb{R}^{n \times k}$  be an arbitrary matrix with full column rank. Assume that  $P = P^T > 0$  satisfies  $A^T P + PA + 2\sigma P < 0$ . Then the reduced model  $\hat{\Sigma}$ , resulting from an oblique projection  $P = VW^T$ , with  $W^T = (V^T P V)^{-1} V^T P$ , satisfies  $\text{Re}(\lambda_i(\hat{A})) < -\sigma$  for all  $i$ .*

*Two-Sided Projection Methods.*

Instead of using the projection matrix  $W$  for stability purposes, the number of preserved multimoments can be increased significantly if we use another sequence of nested Krylov subspaces for the construction of  $W$ . Though the benefit of two-sided projection methods has already been proposed in [3], we will show that the matching conditions stated therein are incomplete. Besides, we want to point out an interesting invariance property which is adapted from linear systems theory. For reasons of comparison, we will focus on low frequency multimoments. The results corresponding to the other cases are equivalent.

**Theorem 4.** *Let a bilinear SISO system  $\Sigma$  be given. If a reduced bilinear system  $\hat{\Sigma}$  is constructed by an oblique projection  $P = VW^T$ ,  $W^T V = I$  with  $V$  and  $W$  given as follows:*

$$\begin{aligned} \text{span}\{V^{(1)}\} &= \mathcal{K}_q(A^{-1}, A^{-1}b), \\ \text{span}\{V^{(k)}\} &= \mathcal{K}_q(A^{-1}, A^{-1}NV^{(k-1)}), \quad k = 2, \dots, r \\ \text{span}\{V\} &= \text{span}\left\{\bigcup_{k=1}^r \text{span}\{V^{(k)}\}\right\} \\ \text{span}\{W^{(1)}\} &= \mathcal{K}_q(A^{-T}, A^{-T}c), \\ \text{span}\{W^{(k)}\} &= \mathcal{K}_q(A^{-T}, A^{-T}N^T W^{(k-1)}), \quad k = 2, \dots, r \\ \text{span}\{W\} &= \text{span}\left\{\bigcup_{k=1}^r \text{span}\{W^{(k)}\}\right\}. \end{aligned}$$

*Then for  $k = 1, \dots, r$  the first  $q^k$  low frequency multimoments of  $\Sigma$  and  $\hat{\Sigma}$  coincide. Additionally, all possible combinations of low frequency multimoments that can be represented as  $w^T v$  or  $w^T N v$ , where  $v \in \text{span}\{V\}$ ,  $w \in \text{span}\{W\}$ , are preserved.*

*Proof.* The first part immediately follows from Theorem 1. Let us consider an arbitrary multimoment of the form  $w^T v$ :

$$\underbrace{c^T A^{-l_k} N A^{-l_{k-1}} \dots N A^{-l_j}}_{w^T} \underbrace{A^{-l_i} N \dots A^{-l_2} N A^{-l_1} b}_{v}, \quad l_i, l_j \leq q$$

Analogue to the discussions for one-sided projections, one can show:

$$\begin{aligned} A^{-l_i} N \dots A^{-l_2} N A^{-l_1} b &= V \hat{A}^{-l_i} \hat{N} \dots \hat{A}^{-l_2} \hat{N} \hat{A}^{-l_1} \hat{b} \\ c^T A^{-l_k} N A^{-l_{k-1}} \dots N A^{-l_j} &= \hat{c}^T \hat{A}^{-l_k} \hat{N} \hat{A}^{-l_{k-1}} \dots \hat{N} \hat{A}^{-l_j} W^T \end{aligned}$$

This leads to:

$$\begin{aligned} &c^T A^{-l_k} N A^{-l_{k-1}} \dots N A^{-l_j} A^{-l_i} N \dots A^{-l_2} N A^{-l_1} b \\ &= \hat{c}^T \hat{A}^{-l_k} \hat{N} \hat{A}^{-l_{k-1}} \dots \hat{N} \hat{A}^{-l_j} W^T V \hat{A}^{-l_i} \hat{N} \dots \hat{A}^{-l_2} \hat{N} \hat{A}^{-l_1} \hat{b} \\ &= \hat{c}^T \hat{A}^{-l_k} \hat{N} \hat{A}^{-l_{k-1}} \dots \hat{N} \hat{A}^{-l_j} \hat{A}^{-l_i} \hat{N} \dots \hat{A}^{-l_2} \hat{N} \hat{A}^{-l_1} \hat{b}, \quad (W^T V = I) \end{aligned}$$

The procedure for multimoments of the form  $w^T N v$  is quite similar.

$$\begin{aligned} &c^T A^{-l_k} N A^{-l_{k-1}} \dots N A^{-l_j} N A^{-l_i} N \dots A^{-l_2} N A^{-l_1} b \\ &= \hat{c}^T \hat{A}^{-l_k} \hat{N} \hat{A}^{-l_{k-1}} \dots \hat{N} \hat{A}^{-l_j} W^T N V \hat{A}^{-l_i} \hat{N} \dots \hat{A}^{-l_2} N \hat{A}^{-l_1} \hat{b} \\ &= \hat{c}^T \hat{A}^{-l_k} \hat{N} \hat{A}^{-l_{k-1}} \dots \hat{N} \hat{A}^{-l_j} \hat{N} \hat{A}^{-l_i} \hat{N} \dots \hat{A}^{-l_2} N \hat{A}^{-l_1} \hat{b}, \quad \text{def. of } \hat{N} \end{aligned}$$

□

**Remark 3.** Note the crucial fact that the above construction automatically doubles the number of matched subsystems.

Before we can establish the connection to the approach given in [3], we will show that the output of the reduced system is independent from the particular choice of a basis for  $W$ . For this, let us consider the following reduced systems, resulting from two different oblique projections:

$$\begin{aligned} \hat{\Sigma} : &\begin{cases} \dot{\hat{x}} = W^T A V \hat{x} + W^T N V \hat{x} u + W^T b u, \\ \hat{y} = c^T V \hat{x}, \end{cases} \\ \tilde{\Sigma} : &\begin{cases} \dot{\tilde{x}} = (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T A \tilde{V} \tilde{x} + (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T N \tilde{V} \tilde{x} u + (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T b u, \\ \tilde{y} = c^T \tilde{V} \tilde{x}, \end{cases} \end{aligned} \quad (3)$$

with  $V, W, \tilde{V}, \tilde{W} \in \mathbb{R}^{n \times k}$  and  $W^T V = I$ . Indeed,  $\tilde{\Sigma}$  results from an oblique projection, since we have  $(\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T \tilde{V} = I$ .

**Theorem 5.** Let  $\text{span}\{V\} = \text{span}\{\tilde{V}\}$  and  $\text{span}\{W\} = \text{span}\{\tilde{W}\}$ . Assume that  $V$  and  $W$  have full column rank. Then the input-output behaviours of  $\hat{\Sigma}$  and  $\tilde{\Sigma}$  coincide.

*Proof.* Since  $\text{span}\{V\} = \text{span}\{\tilde{V}\}$ , the columns of  $\tilde{V}$  can be written as linear combinations of the columns of  $V$  or, equivalently,  $\tilde{V} = VE$ , where  $E \in \mathbb{R}^{k \times k}$ . By making use of the estimate

$$\text{rank}(\tilde{V}) = \text{rank}(VE) \leq \min(\text{rank}(V), \text{rank}(E))$$

we can conclude that  $E$  is invertible. The same argumentation yields  $\tilde{W} = WF$ , with invertible  $F \in \mathbb{R}^{k \times k}$ . Recall that the input-output behaviour of a bilinear system is given by its Volterra series representation. Hence, if we can prove that the transfer functions of all subsystems of  $\tilde{\Sigma}$  and  $\hat{\Sigma}$  coincide, we have shown the assertion. First of all, we notice that:

$$\tilde{W}^T \tilde{V} = F^T W^T V E = F^T E$$

This is due to the biorthogonality of  $W$  and  $V$ . The statement can be shown by induction. For the transfer function of the linear subsystem, we have:

$$\begin{aligned} (sI - \tilde{A})^{-1} \tilde{b} &= (sI - (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T A \tilde{V})^{-1} (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T b \\ &= (sI - (F^T E)^{-1} \tilde{W}^T A \tilde{V})^{-1} (F^T E)^{-1} \tilde{W}^T b \\ &= (sI - (F^T E)^{-1} F^T W^T A V E)^{-1} (F^T E)^{-1} F^T W^T b \\ &= (sI - E^{-1} F^{-T} F^T W^T A V E)^{-1} E^{-1} F^{-T} F^T W^T b \\ &= (sI - E^{-1} W^T A V E)^{-1} E^{-1} W^T b \\ &= E^{-1} (sE^{-1} - E^{-1} W^T A V)^{-1} E^{-1} W^T b \\ &= E^{-1} (sI - W^T A V)^{-1} W^T b \\ &= E^{-1} (sI - \hat{A})^{-1} \hat{b} \end{aligned}$$

Thus, it follows:

$$\tilde{H}(s) = \tilde{c}^T (sI - \tilde{A})^{-1} \tilde{b} = c^T \tilde{V} (sI - \tilde{A})^{-1} \tilde{b} = c^T V E E^{-1} (sI - \hat{A})^{-1} \hat{b} = \hat{H}(s)$$

In order to prove equality for the second subsystem, we obtain:

$$\begin{aligned} (sI - \tilde{A})^{-1} \tilde{N} &= (sI - (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T A \tilde{V})^{-1} (\tilde{W}^T \tilde{V})^{-1} \tilde{W}^T N \tilde{V} \\ &= (sI - (F^T E)^{-1} \tilde{W}^T A \tilde{V})^{-1} (F^T E)^{-1} \tilde{W}^T N \tilde{V} \\ &= (sI - (F^T E)^{-1} F^T W^T A V E)^{-1} (F^T E)^{-1} F^T W^T N V E \\ &= (sI - E^{-1} F^{-T} F^T W^T A V E)^{-1} E^{-1} F^{-T} F^T W^T N V E \\ &= (sI - E^{-1} W^T A V E)^{-1} E^{-1} W^T N V E \\ &= E^{-1} (sE^{-1} - E^{-1} W^T A V)^{-1} E^{-1} W^T N V E \\ &= E^{-1} (sI - W^T A V)^{-1} W^T N V E \\ &= E^{-1} (sI - \hat{A})^{-1} \hat{N} E \end{aligned}$$

Hence, we have:

$$\begin{aligned} \tilde{H}(s_1, s_2) &= \tilde{c}^T (s_2 I - \tilde{A})^{-1} \tilde{N} (s_1 I - \tilde{A})^{-1} \tilde{b} \\ &= c^T \tilde{V} (s_2 I - \tilde{A})^{-1} \tilde{N} (s_1 I - \tilde{A})^{-1} \tilde{b} \\ &= c^T V E E^{-1} (s_2 I - \hat{A})^{-1} \hat{N} E E^{-1} (s_1 I - \hat{A})^{-1} \hat{b} \\ &= \hat{H}(s_1, s_2) \end{aligned}$$

Again, assertions for higher subsystems can be proven equivalently.  $\square$

Similar statements for linear systems, can be found in [12]. Let us now take a closer look at the construction proposed in [3]. The reduced model  $\hat{\Sigma}$  is given as follows:

$$\hat{A} = (W^T A^{-1} V)^{-1}, \quad \hat{N} = \hat{A} W^T A^{-1} N V, \quad \hat{b} = \hat{A} W^T A^{-1} b, \quad \hat{c} = V^T c.$$

While the construction for  $V$  is the same as in Theorem 1, the construction of  $W$  requires a slight change in the sequence of Krylov spaces:

$$\begin{aligned} \text{span}\{W^{(1)}\} &= \mathcal{K}_q(A^{-T}, c), \\ \text{span}\{W^{(k)}\} &= \mathcal{K}_q(A^{-T}, N^T A^{-T} W^{(k-1)}), \quad k = 2, \dots, r. \end{aligned}$$

Since the procedure additionally requires  $W^T V = I$ , it holds:

$$\hat{A} = (W^T A^{-1} V)^{-1} = (W^T A^{-1} V)^{-1} W^T V = (W^T A^{-1} V)^{-1} W^T A^{-1} A V.$$

Hence, the reduced system is of the form (3) with  $\tilde{W} = A^{-T} W$ . This now implies that  $\text{span}\{\tilde{W}\}$  fulfills the assumptions of Theorem 1. Based on this conclusion, let us consider the example from [3]. Here, we will use our approach to reduce  $\Sigma$  which leads to slightly different Krylov spaces. However, we have explained that this does not influence the matching conditions.

**Example 1.**

$$\text{span}\{V\} = \{A^{-1}b, \dots, A^{-7}b, A^{-1}NA^{-1}b, \dots, A^{-1}NA^{-4}b\}$$

and

$$\text{span}\{W^T\} = \{c^T A^{-1}, \dots, c^T A^{-7}, c^T A^{-1}NA^{-1}, \dots, c^T A^{-4}NA^{-1}\}$$

According to Theorem 1, the reduced model preserves 14 multimoments of the first subsystem

$$c^T A^{-1}b, \dots, c^T A^{-14}b,$$

57 multimoments of the second subsystem

$$c^T A^{-l_2} N A^{-l_1} b,$$

where  $l_1, l_2 = 1, \dots, 7$  or  $l_1 = 8, l_2 = 1, \dots, 4$  or  $l_1 = 1, \dots, 4, l_2 = 8$ . For the third subsystem 72 multimoments match

$$c^T A^{-l_3} N A^{-l_2} N A^{-l_1} b,$$

where  $l_1 = 1, \dots, 7, l_2 = 1, l_3 = 1, \dots, 4$  or  $l_1 = 1, \dots, 4, l_2 = 1, l_3 = 1, \dots, 7$  or  $l_1 = 1, \dots, 4, l_2 = 2, l_3 = 1, \dots, 4$  and 16 multimoments of the fourth subsystem

$$c^T A^{-l_4} N A^{-l_3} N A^{-l_2} N A^{-l_1} b,$$

where  $l_1 = 1, \dots, 4, l_2, l_3 = 1, l_4 = 1, \dots, 4$ . Altogether, we record that 159 multimoments (not only 69, cf. [3]) are preserved by a reduced model of dimension 11.

In the following section we will see that two-sided projections may lead to much better approximation results than one-sided projections. On the other hand, they are more likely to produce ill-conditioned system matrices as well as unstable systems, which sometimes annihilates their benefit.

#### 4. Numerical Examples

We apply our results to a discretized nonlinear partial differential equation and to a model of an RC-circuit with nonlinear resistor.

##### *A flow model*

We use *Carleman bilinearization* to transform a semi-discretized nonlinear partial differential equation into a large-scale bilinear control system. Consider the one-dimensional viscid Burgers equation

$$\begin{aligned} \frac{\partial w}{\partial t} + w \frac{\partial w}{\partial x} &= \frac{\partial}{\partial x} \left( \nu \frac{\partial w}{\partial x} \right) && \text{for } (x, t) \in (0, L) \times (0, T) && (4) \\ w(x, 0) &= p(x) && \text{for } x \in (0, L) \\ w(0, t) &= u(t) && \text{for } t \in (0, T) \\ w(L, t) &= q(t) && \text{for } t \in (0, T) \end{aligned}$$

Such systems are used to model gas dynamics and traffic flow. The solution  $w(x, t)$  is interpreted as a particular velocity at a point  $x$  and a time  $t$ . In general, the viscosity coefficient  $\nu(x, t)$  depends on space and time as well. In [7], an optimal control problem for this situation is discussed. There, the goal is to find optimal parameters  $(p_{opt}, u_{opt}, q_{opt}, \nu_{opt})$  such that the corresponding solution  $w_{opt}$  models a given quantity  $\hat{w}$  as effectively as possible (data assimilation).

To focus on model order reduction, we make some simplifications. First of all, the viscosity coefficient  $\nu(x, t) = \nu$  is assumed to be a constant parameter. Furthermore, we impose zero initial condition on the system, i.e.  $f(x) = 0$ . Finally, let us concentrate on the case where only the left boundary is subject to a control while the right one is considered to be zero.

Let us perform a spatial discretization of equation (4), using an equidistant step size  $h = \frac{L}{N+1}$ . Here,  $N$  denotes the number of interior points of the interval  $(0, L)$ . Making use of the approximations

$$\frac{\partial w}{\partial x} \approx \frac{w(x+h) - w(x-h)}{2h}, \quad \frac{\partial^2 w}{\partial x^2} \approx \frac{w(x+h) - 2w(x) + w(x-h)}{h^2},$$

we obtain the nonlinear control system

$$\frac{d}{dt} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_i \\ \vdots \\ w_N \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{w_1 w_2}{2h} + \frac{\nu}{h^2}(w_2 - 2w_1) \\ -\frac{w_2}{2h}(w_3 - w_1) + \frac{\nu}{h^2}(w_3 - 2w_2 + w_1) \\ \vdots \\ -\frac{w_i}{2h}(w_{i+1} - w_{i-1}) + \frac{\nu}{h^2}(w_{i+1} - 2w_i + w_{i-1}) \\ \vdots \\ \frac{w_N w_{N-1}}{2h} + \frac{\nu}{h^2}(-2w_N + w_{N-1}) \end{bmatrix}}_{f(w)} + \underbrace{\begin{bmatrix} \frac{w_1}{2h} + \frac{\nu}{h^2} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{g(w)} u.$$

The idea behind the Carleman bilinearization procedure is to approximate the nonlinear functions  $f$  and  $g$  by a Taylor expansion. Since the nonlinearity of  $f$

is given by simple products of the state variables, a second order approximation already yields exact results. If we use a Kronecker product notation, this can be expressed as

$$\begin{aligned} f(w) &= A_1 w + \frac{1}{2} A_2 (w \otimes w), \\ g(w) &= B_0 + B_1 w, \end{aligned}$$

with  $B_0 \in \mathbb{R}^n$ ,  $A_1, B_1 \in \mathbb{R}^{n \times n}$ ,  $A_2 \in \mathbb{R}^{n \times n^2}$ . Here,  $A_1, B_1$  denote the Jacobians of  $f$  and  $g$ , respectively and  $A_2$  stands for the matrix of second derivatives of  $f$ . Thus, we have for the first and second order terms:

$$\begin{aligned} \frac{d}{dt} w &= A_1 w + \frac{1}{2} A_2 (w \otimes w) + (B_0 + B_1 w) u \\ \frac{d}{dt} (w \otimes w) &= \dot{w} \otimes w + w \otimes \dot{w} \\ &= \left( A_1 w + \frac{1}{2} A_2 (w \otimes w) + (B_0 + B_1 w) u \right) \otimes w \\ &\quad + w \otimes \left( A_1 w + \frac{1}{2} A_2 (w \otimes w) + (B_0 + B_1 w) u \right) \\ &= (A_1 \otimes I + I \otimes A_1) (w \otimes w) + (B_0 \otimes I + I \otimes B_0) w u + \dots \end{aligned}$$

By introducing the enlarged state vector

$$x = \begin{bmatrix} w \\ w \otimes w \end{bmatrix},$$

we construct the  $N + N^2$ -dimensional bilinear system

$$\dot{x} = \begin{bmatrix} A_1 & & \frac{1}{2} A_2 \\ 0 & A_1 \otimes I + I \otimes A_1 & \end{bmatrix} x + \begin{bmatrix} B_1 & 0 \\ B_0 \otimes I + I \otimes B_0 & 0 \end{bmatrix} x u + \begin{bmatrix} B_0 \\ 0 \end{bmatrix} u$$

The average value of  $w$  over the interval  $(0, L)$  is assumed to be the measurable output of the system:

$$y = \frac{1}{N} [1 \quad \dots \quad 1 \quad 0 \quad \dots \quad 0] x$$

We report the results for some simulations generated with the values  $L = 1$ ,  $\nu = 0.1$  and  $N = 300$ . The original bilinearized system  $\Sigma$  was of dimension 90300. Relative errors of original and reduced systems for some typical input signals were compared. Similar to recently investigated examples, we stuck to the matching of multimoments of the first and second subsystem. As was indicated in Remark 2, we made use of the fact that the number of preserved first and second order multimoments do not have to be the same. For example, in case of matching low frequency multimoments, we constructed the projection

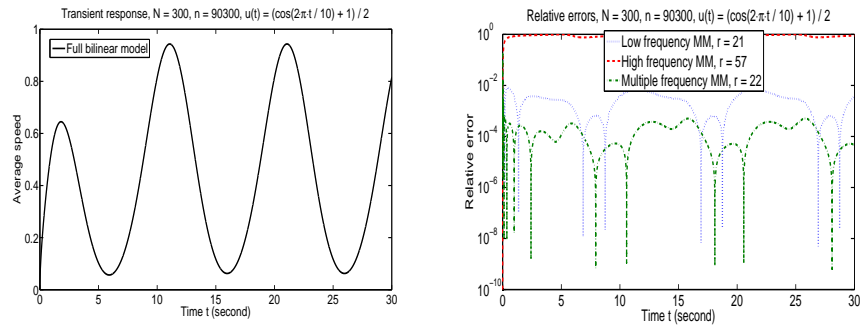


Figure 1: Relative errors depending on expansion points (orthogonal projection).

subspace  $V$  as

$$\begin{aligned} \text{span}\{V^{(1)}\} &= \mathcal{K}_{12}(A^{-1}, A^{-1}b), \\ \text{span}\{V^{(2)}\} &= \mathcal{K}_3(A^{-1}, A^{-1}NV_{[3]}^{(1)}) \\ \text{span}\{V\} &= \text{span}\{V^{(1)}\} \cup \text{span}\{V^{(2)}\}. \end{aligned}$$

Here, the square bracket denotes the first 3 columns of the matrix  $V^{(1)}$ . According to Remark 1, if there is no deflation in the Krylov spaces, we expect the reduced model to be of dimension  $\hat{n} = 12 + 3 \cdot 3 = 21$ . Note that the actual dimensions of the reduced models might be smaller due to necessary orthogonalization steps in the construction of the subspace  $V$ . An equivalent construction was used for high frequency multimoments. However, as shown in figures 1 and 2, even the increase to matching 50 first order multimoments together with some second order multimoments did not yield satisfying approximations that would allow recommending this approach. On the other hand, one should note that the reduction process required no matrix inversions but rather could be achieved by simple matrix products. Based on the ideas pro-

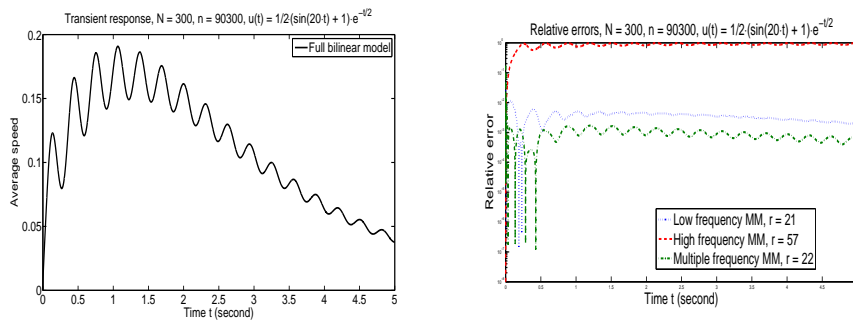


Figure 2: Relative errors depending on expansion points (orthogonal projection).

posed in Section 3, we further investigated a method that may be interpreted

as a bilinear generalization of the rational interpolation problem. Similar to linear systems, matching multimoments using a single interpolation point will approximate the transfer functions only at a specific frequency. Hence, it seems reasonable to perform multimoment-matching at multiple interpolation points in order to cover a broader frequency range. In our case, we decided to use

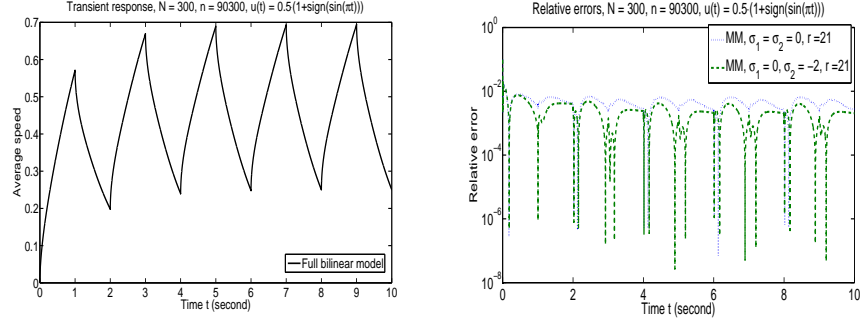


Figure 3: Relative errors for two individual chosen expansion points.

multimoments corresponding to the following expansion points:

$$\sigma_1 = \sigma_2 = 0, 1, 10 \quad (2 + 2 \cdot 2)$$

$$\sigma_1 = \sigma_2 = 100, \infty \quad (1 + 1 \cdot 1)$$

Consequently, the reduced model was of dimension  $3 \cdot 6 + 2 \cdot 2 = 22$ . Referring again to figures 1 and 2, we note that this technique indeed yields significant improvements. Finally, figures 3 and 4 illustrate that using different expansion

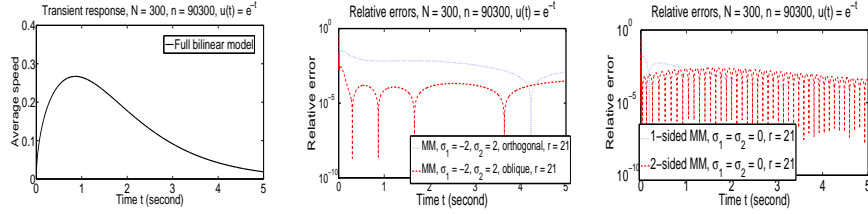


Figure 4: Comparison of orthogonal, oblique ( $E = A^{-1}$ ) and two-sided projection.

points for the first and second subsystem are legitimate. Moreover, in figure 3 this variation slightly increased accuracy of the reduced model. Figure 4 reveals the influence of different projection methods that may be used for reduction. Although the number of preserved multimoments of orthogonal and oblique projection coincided, the results for the latter one are surprisingly better. In contrast to that, we would probably have expected more variance between one-sided and two-sided projection methods.

### A nonlinear RC circuit

Our second example is a large-scale system resulting from the Carleman bilinearization of a nonlinear RC ladder network. This application can be seen as a standard test case since it has been studied in most papers focusing on bilinear model order reduction, see e.g. [2, 5, 9]. For our purposes, we used a circuit consisting of 500 nonlinear resistors, leading to a bilinear system of dimension 250500. In case of low frequency and high frequency multimoments, respectively, we preserved  $12 + 3 \cdot 3$  multimoments. The expansion points used for the combined multimoment-matching were chosen as

$$\sigma_1 = \sigma_2 = 0, \infty \quad (4 + 2 \cdot 2),$$

$$\sigma_1 = \sigma_2 = 1, 10, 100 \quad (1 + 1 \cdot 1).$$

In contrast to the flow model, we conclude from figure 5 that matching high frequency multimoments as well as using multiple interpolation points seems superior to the common ansatz ( $\sigma = 0$ ). Especially for the high-frequency input signal, the new approaches should be preferred.

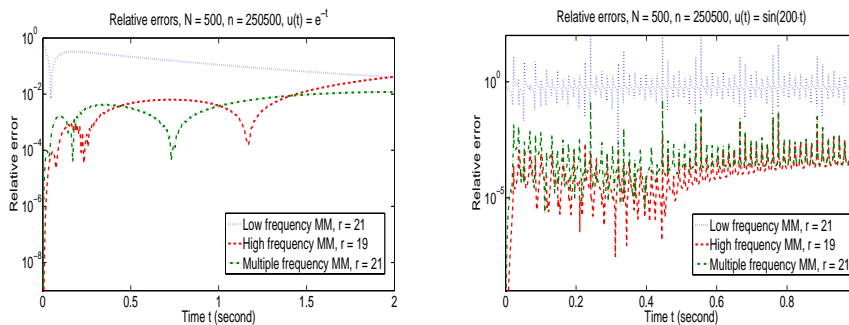


Figure 5: Relative errors for RC circuit.

## 5. Conclusions

We have generalized recent results on Krylov-subspace based model order reduction of bilinear control systems. It has been shown that the transfer functions corresponding to the Volterra series of a bilinear system can be expanded at arbitrary points. The existence of an infinite sequence of transfer functions allows the use of different expansion points for different subsystems. The use of multiple interpolation points provides approximations over a broad frequency range. Furthermore, we have seen how to extend the concept of Markov parameters to bilinear systems. A construction for a reduced model that matches certain high frequency multimoments of a given bilinear system was presented. We have established the connection to an existent approach by help of using a specific oblique projection. The freedom in choosing an arbitrary left inverse

without influencing the multimoment-matching property yielded a possibility to preserve the stability of the system matrix  $A$ . Finally, the theoretical benefit of using two-sided projection methods has been pointed out. We have shown that these constructions automatically double the number of matched subsystems. By means of two numerical examples we demonstrated the superiority of some of the proposed methods in certain situations.

## References

- [1] A.C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Cambridge University Press, 2005.
- [2] Z. Bai, D. Skoogh, A projection method for model reduction of bilinear dynamical systems, *Linear Algebra Appl.* 415 (2006) 406-425.
- [3] Z. Bai, D. Skoogh, Krylov subspace techniques for reduced-order modeling of nonlinear dynamical systems, *Appl. Numer. Math* 43 (2002) 9-44.
- [4] C. Bruni, G. DiPillo, G. Koch, On the mathematical models of bilinear systems, *Ricerche di Automatica* 2 (1971) 11-26.
- [5] M. Condon, R. Ivanov, Krylov subspaces from bilinear representations of nonlinear systems, *COMPEL* 26 (2007) 399-406.
- [6] L. Feng, P. Benner, A note on projection techniques for model order reduction of bilinear systems, in: *International Conference of Numerical Analysis and Applied Mathematics*. AIP Conference Proceedings 936, Corfu, 2007, pp. 208-211.
- [7] Y. Leredde et al., On initial, boundary conditions and viscosity coefficient control for Burgers' equation, *Int. J. Numer. Methods Fluids* 28 (1998) 113-128.
- [8] R.R. Mohler, *Bilinear Control Processes*, New York Academic Press, 1973.
- [9] J.R. Phillips, Projection-based approaches for model reduction of weakly nonlinear time-varying systems, *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.* 22 (2003), 171-183.
- [10] J.R. Phillips, Projection frameworks for model reduction of weakly nonlinear systems, *Proceedings of DAC 2000* (2000) 184-189.
- [11] W.J. Rugh, *Nonlinear System Theory - The Volterra/Wiener Approach*, The Johns Hopkins University Press, 1981.
- [12] B. Salimbahrami, B. Lohmann, Krylov subspace methods in linear model order reduction: Introduction and invariance properties, *Sci. Rep. Inst. of Automation U Bremen* (2002).
- [13] A. Yousefi, *Preserving Stability in Model and Controller Reduction with Application to Embedded Systems*, TU Munich, 2006.