MAX-PLANCK-GESELLSCHAFT

**Max Planck Institute Magdeburg**
**Preprints**

Pawan Goyal    Mian Ilyas Ahmad    Peter Benner

# Krylov Subspace-based Model Reduction for a Class of Bilinear Descriptor Systems

MAX–PLANCK–INSTITUT
FÜR DYNAMIK KOMPLEXER
TECHNISCHER SYSTEME
MAGDEBURG

**MPIMD/15-07**    **May 18, 2015**

**Abstract**

We consider model order reduction of bilinear descriptor systems using an interpolatory projection framework. Such nonlinear descriptor systems can be represented by a series of generalized linear descriptor systems (also called subsystems) by utilizing the Volterra-Wiener approach [22]. Standard projection techniques for bilinear systems utilize the generalized transfer function of these subsystems to construct an interpolating approximation. However, the resulting reduced-order system may not match the polynomial part of the generalized transfer functions. This may result in an unbounded error in terms of $\mathcal{H}_2$ or $\mathcal{H}_\infty$ norms. In this paper, we derive an explicit expression for the polynomial part of each subsystem by assuming a special structure of the bilinear system which reduces to an index-1 linear DAE if the bilinear term is zero. This allows us to propose an interpolatory technique for bilinear DAEs which not only achieves interpolation but also retains the polynomial part of the bilinear system. The approach extends the interpolatory technique for index-1 linear DAEs [18] to bilinear DAEs. Numerical examples are used to illustrate the theoretical results.

**Impressum:**

**Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg**

**Publisher:**
Max Planck Institute for
Dynamics of Complex Technical Systems

**Address:**
Max Planck Institute for
Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

www.mpi-magdeburg.mpg.de/preprints

# 1 Introduction

The importance of model order reduction arises in the analysis of high order mathematical models that describe complex dynamical systems. These high order models are often expensive to observe and therefore, they are replaced by reduced-order systems to simulate the approximate behavior of the actual system. Various approaches have been developed for model order reduction, see, e.g. [1, 4, 10, 23]. In case of linear systems, balanced truncation [20], moment-matching methods [14] and the iterative rational Krylov method [17] are well-used and well-established model reduction methods. However, most practical systems have nonlinearities and model reduction of such systems, particularly models described by differential algebraic equations (DAEs), also called descriptor systems, are less developed and require further research.

In this paper, we investigate Krylov projection methods for bilinear descriptor systems. In general, a bilinear descriptor system has the form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + \sum_{i=1}^{m} N^{(i)}x(t)u_i(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are the state, input and output vectors, respectively. The matrices $E,\ A,\ N^{(i)},\ i = 1, \ldots, m,\ B,\ C$ and $D$ are all real with dimensions determined by those of $x(t)$, $u(t)$ and $y(t)$. Notice that the bilinear terms in the system, involving the product of state and inputs, make it a special class of nonlinear systems. Also the matrix $E$ might be singular, but it is assumed that the matrix pencil $(A, E)$ is regular and stable, that is $\det(sE - A) \not\equiv 0$ and all finite eigenvalues of the matrix pencil $(sE-A)$ have strictly negative real parts, respectively.

Model reduction aims at deriving another system with much smaller state-space dimension $r \ll n$, similar to (1), i.e.,

$$\begin{aligned} E_r\dot{x}_r(t) &= A_r x_r(t) + \sum_{i=1}^{m} N_r^{(i)}x_r(t)u_i(t) + B_r u(t), \\ y_r(t) &= C_r x_r(t) + D_r u(t), \end{aligned} \tag{2}$$

such that the output behavior and some important properties of (1) are retained by (2) for an admissible set of input functions $u(t)$. The reduced-order system (2) can be obtained via projections as follows:

- Construct basis matrices $V \in \mathbb{R}^{n\times r}$ and $W \in \mathbb{R}^{n\times r}$ for the subspaces $\mathcal{V}$ and $\mathcal{W}$ respectively.
- Approximate $x(t)$ by $Vx_r(t)$.
- Ensure the Petrov-Galerkin condition:

$$W^T\left(EVx_r(t) - AVx_r(t) - \sum_{i=1}^{m} N^{(i)}Vx_r(t)u_i(t) - Bu(t)\right) = 0.$$

As a result, the state matrices associated with the reduced-order system (2) are given by

$$\begin{aligned}&E_r = W^T E V, \quad A_r = W^T A V, \quad N_r^{(i)} = W^T N^{(i)} V,\\ &B_r = W^T B, \quad C_r = CV.\end{aligned}$$

Clearly, for a given system, the reduced-order system obtained via projection depends on the choice of $V$ and $W$, or equivalently, on the subspaces $\mathcal{V}$ and $\mathcal{W}$. If the matrix $E$ is the identity matrix or nonsingular, these basis matrices and the resulting reduced-order system can be computed by extending the standard balanced truncation and interpolatory projection methods from linear to bilinear systems [2, 3, 7, 9, 11, 13, 21, 24]. The bilinear version of balanced truncation involves the solution of two generalized Lyapunov equations, which is known to be computationally complex [9]. In [8] effective methods for solving these Lyapunov equations are suggested. Their extension to the descriptor case is an open problem, though. Therefore, we focus on interpolatory projection methods for descriptor systems.

Recently, it was shown in [18] that for a linear descriptor system, it is necessary for interpolatory techniques to compute a reduced-order system which not only interpolates the actual transfer function of the system, but also retains its polynomial part, in order to ensure a bounded error in terms of the $\mathcal{H}_2$-norm. We extend this observation to bilinear descriptor systems and to our knowledge this has not been considered before. The idea is to compute a reduced-order system for a given bilinear DAE system such that the generalized transfer functions associated with the reduced-order and the actual bilinear system not only interpolate at some predefined interpolation points, but also match the corresponding polynomial parts. This involves, first, identifying the generalized transfer functions of the bilinear DAE system which is possible by using the Volterra series representation [22]. Secondly, we construct the basis matrices $V$ and $W$, where the first $k$ generalized transfer functions are used similar to the standard interpolatory subspaces [11]. Then, we identify the polynomial part of each generalized transfer function and finally project the bilinear DAE system to obtain the required reduced-order system.

It is not straightforward to identify explicitly the polynomial part of the generalized transfer functions. In this paper, we assume a special structure of bilinear systems which allows us to compute explicitly a constant polynomial part of each generalized transfer function. The special structure reduces to an index-1 linear DAE system, if the bilinear term is zero. In Section 2, we first discuss interpolatory techniques for such index-1 linear DAE systems. Its extension with the required modifications to the special class of bilinear DAE systems is shown in Section 3, where an expression for the polynomial parts of each generalized transfer function is also derived. In Section 4, we discuss computational issues arising in the interpolatory technique proposed in Section 3. Finally, in Section 5, we present numerical results to illustrate the implementation of our approach.

## 2 Interpolatory Model Reduction for Linear DAEs

In this section, we briefly review interpolatory projection methods for model reduction of linear descriptor systems. Note that the system in (1) reduces to a linear descriptor system for $N^{(i)} = 0$:

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t). \end{aligned} \tag{3}$$

Throughout our discussion in this paper, we assume single-input single-output systems, although the results can be extended to the multi-input multi-output case. This means that $B$, $C^T$ represent column vectors and $D$ is a scalar. Also, we denote the transfer function $\mathcal{G}(s) := C(sE - A)^{-1}B + D$ which can be decomposed into strictly proper ($\mathcal{G}_{sp}(s)$) and polynomial ($\mathcal{P}(s)$) parts, i.e., $\mathcal{G}(s) = \mathcal{G}_{sp}(s) + \mathcal{P}(s)$.

The problem of reducing the above linear descriptor system by interpolatory projection has been considered recently in [18]. It is shown there that the standard interpolatory techniques for model reduction of linear DAEs will generically produce an interpolating ODE system (reduced-order) and will not necessarily match the polynomial part of the DAE system. This may result in an unbounded $\mathcal{H}_2$ error. To overcome this issue, an idea was proposed to identify explicitly the polynomial part of the transfer function and ensure that the reduced-order system retains this polynomial part by using modified interpolatory subspaces for projection [18]. For special descriptor systems of index-1, this can be achieved without modifying the standard interpolatory subspaces [18]. The approach is based on the idea given in [5, 19], where the reduced transfer function $\mathcal{G}_r(s) = C_r(sE_r - A_r)^{-1}B_r + D_r$ interpolates $\mathcal{G}(s) = C(sE - A)^{-1}B + D$ with $D_r \neq D$, unlike for standard interpolation methods where $D_r = D$. In the following, we review this idea of interpolating with $D_r \neq D$ for linear index-1 descriptor systems [18].

Consider a linear descriptor system of index-1,

$$\begin{aligned} E_{11}\dot{x}_1(t) + E_{12}x_2(t) &= A_{11}x_1(t) + A_{12}x_2(t) + B_1u(t), \\ 0 &= A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t), \\ y(t) &= C_1x_1(t) + C_2x_2(t) + Du(t), \end{aligned} \tag{4}$$

where $x_1(t) \in \mathbb{R}^{n_1}$ and $x_2(t) \in \mathbb{R}^{n_2}$. By the index-1 assumption, the matrices $A_{22}$ and $E_{11} - E_{12}A_{22}^{-1}A_{21}$ are invertible. For an index-1 descriptor system, the polynomial part $\mathcal{P}(s)$ is constant and can be determined by the following result.

**Lemma 2.1.** *[18]. Let $\mathcal{G}(s)$ be the transfer function of the linear descriptor system* (4) *in which $A_{22}$ and $E_{11} - E_{12}A_{22}^{-1}A_{21}$ are both nonsingular. Then, the polynomial part of $\mathcal{G}(s)$ can be written as*

$$\mathcal{P} = CMB + D, \tag{5}$$

*where*

$$M = \lim_{s\to\infty}(sE - A)^{-1} = \begin{bmatrix} 0 & E_A^{-1}E_{12}A_{22}^{-1} \\ 0 & -A_{22}^{-1}\left(I + A_{21}E_A^{-1}E_{12}A_{22}^{-1}\right) \end{bmatrix} \tag{6}$$

*with $E_A = E_{11} - E_{12}A_{22}^{-1}A_{21}$ and $s := 2\pi jf$ is the Laplace variable in which $f$ is the frequency and $j$ is the imaginary unit.*

To ensure bounded error, the reduced transfer function $\mathcal{G}_r(s) = \hat{\mathcal{G}}_{sp}(s)+\hat{\mathcal{P}}(s)$ should not only interpolate $\mathcal{G}(s)$ but also match the constant part, $\hat{\mathcal{P}}(s) = \mathcal{P}(s) = \mathcal{P}$. This means that the problem reduces to identify $\hat{\mathcal{G}}_{sp}(s)$, which interpolates $\mathcal{G}_{sp}(s)$. Note that it is easy to identify an interpolating $\hat{\mathcal{G}}_{sp}(s)$, once we have an explicit expression for $\mathcal{G}_{sp}(s)$. However, the goal is to identify an interpolating $\mathcal{G}_r(s)$ without explicitly constructing $\mathcal{G}_{sp}(s)$ utilizing a special structure of the pencil $(A, E)$. The following theorem provides a possible solution.

**Theorem 2.1.** *[5, 18] Let $\mathcal{G}(s)$ be the transfer function of the linear descriptor system (3). Assume that the interpolation points $\sigma$ and $\mu$ are given such that $sE - A$ and $sE_r - A_r$ are invertible for $s = \sigma, \mu$. Define the projection matrices $V \in \mathbb{C}^{n\times r}$ and $W \in \mathbb{C}^{n\times r}$ such that*

- $\operatorname{range}(V) = \mathcal{K}_q\left((\sigma E - A)^{-1}E, (\sigma E - A)^{-1}B\right)$,
- $\operatorname{range}(W) = \mathcal{K}_q\left((\mu E - A)^{-T}E^T, (\mu E - A)^{-T}C^T\right)$,

*where $\mathcal{K}_q(\mathcal{A}, \mathcal{B}) = \operatorname{span}\left\{\mathcal{B}, \mathcal{A}\mathcal{B}, \ldots, \mathcal{A}^{q-1}\mathcal{B}\right\}$. Also let $F \in \mathbb{C}^{n\times 1}$ and $G \in \mathbb{C}^{n\times 1}$ be solutions to*

$$F^T V = (e_1^r)^T \quad \text{and} \quad W^T G = e_1^r, \tag{7}$$

*where $e_1^r$ is the first column of an $r \times r$ identity matrix. Then, projection of the intermediate system $\tilde{\mathcal{G}}(s) = \tilde{C}(s\tilde{E} - \tilde{A})^{-1}\tilde{B} + \tilde{D}$,*

$$\tilde{E} = E, \quad \tilde{A} = A + G\hat{D}F^T, \quad \tilde{B} = B - G\hat{D}, \quad \tilde{C} = C - \hat{D}F^T, \quad \tilde{D} = \mathcal{P},$$

*where $\hat{D} = \mathcal{P} - D$, results in a reduced-order system $\mathcal{G}_r(s) = C_r(sE_r - A_r)^{-1}B_r + D_r$, in which:*

$$E_r = W^T\tilde{E}V, \quad A_r = W^T\tilde{A}V, \quad B_r = W^T\tilde{B}, \quad C_r = \tilde{C}V_r, \quad D_r = \tilde{D}.$$

*Assuming that $E_r$ is invertible, the polynomial parts of $\mathcal{G}_r(s)$ and $\mathcal{G}(s)$ match, that is $D_r = \mathcal{P}$. Also $\mathcal{G}_r(s)$ satisfies the following interpolation conditions*

$$\mathcal{G}^{(l)}(\sigma) = \mathcal{G}_r^{(l)}(\sigma), \quad \mathcal{G}^{(l)}(\mu) = \mathcal{G}_r^{(l)}(\mu), \quad l = 0, \ldots, q-1,$$

*If $\sigma = \mu$, then*

$$\mathcal{G}^{(l)}(\sigma) = \mathcal{G}_r^{(l)}(\sigma), \quad l = 0, \ldots, 2q-1.$$

**Remark 2.1.** *The reduced transfer function $\mathcal{G}_r(s)$ is not only interpolating $\mathcal{G}(s)$ with $D_r \neq D$ (unlike for standard interpolation methods), but also matches the polynomial part of $\mathcal{G}(s)$, that is $D_r = \mathcal{P}$. This is possible by first computing the constant polynomial part $\mathcal{P}$ of $\mathcal{G}(s)$, then constructing the intermediate system $\tilde{\mathcal{G}}(s)$ with $\tilde{D} = \mathcal{P}$, and then, applying oblique projection to $\tilde{\mathcal{G}}(s)$ with standard interpolatory subspaces $V$ and $W$ associated with $\mathcal{G}(s)$.*

**Remark 2.2.** *Theorem 2.1 does not require the explicit computation of $F$ and $G$ in order to compute the reduced-order system. The expressions for $W^T G$ and $F^T V$ can be substituted directly from* (7)*.*

**Remark 2.3.** *In case of Hermite interpolation with $m$ distinct interpolation points (i.e., using $\sigma_i$ and $\mu_i$, $i = 1, \ldots, m$), the conditions on $F$ and $G$ become*

$$F^T V = [\, \underbrace{1, 1, \ldots, 1}_{m \text{ times}} \,] \quad \textit{and} \quad W^T G = [\, \underbrace{1, 1, \ldots, 1}_{m \text{ times}} \,]^T.$$

# 3 Interpolatory Model Reduction for Bilinear Descriptor Systems

In this section, we extend the interpolatory technique with $D_r \neq D$ as discussed before for index-1 linear DAEs to a special class of bilinear descriptor systems. Consider a bilinear descriptor system, where the matrix pencil $(A, E)$ has a structure analogous to the index-1 linear DAE, given in (4). That is,

$$\begin{aligned} E_{11}\dot{x}_1(t) + E_{12}\dot{x}_2(t) &= A_{11}x_1(t) + A_{12}x_2(t) + N_{11}x_1(t)u(t) + N_{12}x_2(t)u(t) + B_1 u(t), \\ 0 &= A_{21}x_1(t) + A_{22}x_2(t) + N_{21}x_1(t)u(t) + N_{22}x_2(t)u(t) + B_2 u(t), \\ y(t) &= C_1 x(t) + C_2 x_2(t) + Du(t), \end{aligned}$$

where $A_{22}$ and $E_{11} - E_{12}A_{22}^{-1}A_{21}$ are invertible. In frequency domain, the input-output representation of the bilinear system is given by the Volterra series representation of the system. Each term of the Volterra series can be considered as a subsystem of the bilinear descriptor system and involves generalized multivariate transfer functions. The structure of these multivariate transfer functions corresponding to the $kth$ subsystem in the regular form is given as

$$H(s_1, \ldots, s_k) = C(s_k E - A)^{-1} N (s_{k-1}E - A)^{-1} N \cdots N(s_1 E - A)^{-1} B + D\delta(k-1). \quad (8)$$

where $\delta(l) = 1$, if $l = 0$ and $\delta(l) = 0$, otherwise. The next lemma shows that each subsystem of (8) has a constant polynomial part:

**Lemma 3.1.** *Let $H(s_1, \ldots, s_k)$ be defined as in* (8)*, that is the regular multivariate Laplace transform of the degree-k kernel associated to $\Sigma$, where $A_{22}$ and $E_{11} - E_{12}A_{22}^{-1}A_{21}$ are both nonsingular. Then, the constant polynomial part of $H(s_1, \ldots, s_k)$ is given by*

$$D_k = C(MN)^{k-1}MB + D\delta(k-1), \quad (9)$$

*where $M$ is as defined in* (6)*.*

*Proof.* Let $F(\mathcal{S}_k) = F(s_1, \ldots, s_k)$ be the multivariable function

$$F(\mathcal{S}_k) = (s_k E - A)^{-1} N (s_{k-1}E - A)^{-1} N \cdots N(s_1 E - A)^{-1} B, \quad (10)$$

then the polynomial part of $H(\mathcal{S}_k)$ is given by

$$D_k = C \lim_{\mathcal{S}_k \to \infty} F(\mathcal{S}_k) + D\delta(k-1). \tag{11}$$

Note that for $k = 1$, (10) reduces to the linear case and using (6) we have

$$\lim_{\mathcal{S}_1 \to \infty} F(s_1) = \lim_{s_1 \to \infty} (s_1 E - A)^{-1} B = MB. \tag{12}$$

It is easy to see from (11) that (9) holds for $k = 1$ (analog to the linear case). Now for $k = j \geq 1$, assume that

$$\lim_{\mathcal{S}_j \to \infty} F(\mathcal{S}_j) = (MN)^{j-1} MB. \tag{13}$$

We need to show that the above equation holds for $k = j + 1$. Note that

$$F(\mathcal{S}_{j+1}) = (s_{j+1} E - A)^{-1} N F(\mathcal{S}_j).$$

Taking the limit $\mathcal{S}_{j+1} \to \infty$, we have

$$\begin{aligned}\lim_{\mathcal{S}_{j+1} \to \infty} F(\mathcal{S}_{j+1}) &= \lim_{s_{j+1} \to \infty} (s_{j+1} E - A)^{-1} N \lim_{\mathcal{S}_j \to \infty} F(\mathcal{S}_j),\\ &= \lim_{s_{j+1} \to \infty} (s_{j+1} E - A)^{-1} N (MN)^{j-1} MB,\end{aligned}$$

where the last equation follows from (13). Now, we define $\mathcal{B}_{MN} = N(MN)^{j-1} MB$ and use (6) to obtain

$$\lim_{\mathcal{S}_{j+1} \to \infty} F(\mathcal{S}_{j+1}) = \lim_{s_{j+1} \to \infty} (s_{j+1} E - A)^{-1} \mathcal{B}_{MN} = M\mathcal{B}_{MN} = (MN)^j MB$$

Thus, (11) implies that (9) holds. □

Lemma 3.1 suggests that if

$$N = \begin{bmatrix} N_{11} & N_{12} \\ 0 & 0 \end{bmatrix}, \quad \text{where } N_{11} \in \mathbb{R}^{n_1 \times n_1} \text{ and } N_{12} \in \mathbb{R}^{n_1 \times n_2}, \tag{14}$$

then $D_k = 0$ for $k > 1$ and $\mathcal{P} = D_1 = CMB + D$. This means that only the first subsystem has polynomial part and all other subsystems have zero polynomial part. The next subsection addresses the issue of retaining the polynomial part, $D_1$, in the first subsystem of the reduced bilinear system.

## 3.1 Interpolating a bilinear descriptor system and retaining $D_1$

We begin with outlining the standard interpolatory projection, where $D_r = D$.

**Theorem 3.1.** *[11] Consider arbitrary interpolation points $\sigma_i, \mu_i \in \mathbb{C}$ such that $sE-A$ and $sE_r - A_r$ are invertible for $s = \sigma_i, \mu_i,\ i = 1, \ldots, k$. Define the projection matrices $V$ and $W$ as follows*

$$\begin{aligned}
\text{range}\left(V^{(1)}\right) &= \mathcal{K}_q\left((\sigma_1 E - A)^{-1}E, (\sigma_1 E - A)^{-1}B\right),\\
\text{range}\left(V^{(i)}\right) &= \mathcal{K}_q\left((\sigma_i E - A)^{-1}E, (\sigma_i E - A)^{-1}NV^{(i-1)}\right), \quad i = 2, \ldots, k,\\
\text{range}\left(W^{(1)}\right) &= \mathcal{K}_q\left((\mu_1 E - A)^{-T}E^T, (\mu_1 E - A)^{-T}C^T\right),\\
\text{range}\left(W^{(i)}\right) &= \mathcal{K}_q\left((\mu_i E - A)^{-T}E^T, (\mu_i E - A)^{-T}N^T W^{(i-1)}\right), \quad i = 2, \ldots, k,\\
\text{range}\,(V) = \bigcup_{i=1}^{k} &\left\{\,\text{range}\left(V^{(i)}\right)\right\}, \quad \text{range}\,(W) = \bigcup_{i=1}^{k}\left\{\,\text{range}\left(W^{(i)}\right)\right\}.
\end{aligned}$$

*Assume $V$ and $W$ are full column rank matrices. Construct the reduced-order system matrices as*

$$\begin{array}{lll}
E_r = W^T E V, & A_r = W^T A V, & N_r = W^T N V,\\
B_r = W^T B, & C_r = CV, & D_r = D,
\end{array}$$

*then*

$$H(\mathcal{S}_k) = H_r(\mathcal{S}_k) + \mathcal{O}\left((s_1-\mu_1)^q \cdots (s_k-\mu_k)^q (s_1-\sigma_1)^q \cdots (s_k-\sigma_k)^q\right).$$

**Remark 3.1.** *It was shown in [11] that with the choice of the projection matrices $V$ and $W$ in Theorem 3.1 also yields the matching of additional moments which involve $W^T EV$ and $W^T NV$.*

Our aim is to utilize the basis matrices $V$ and $W$ as given in Theorem 3.1 and extend the approach used in Theorem 2.1 to the bilinear descriptor system (8), with $N_{21} = N_{22} = 0$ as in (14). The following theorem provides a possible solution:

**Theorem 3.2.** *Let $V$ and $W$ be as defined in Theorem 3.1 and assume that the structure of the matrix $N$ is as in* (14)*. Also let $F \in \mathbf{C}^{n\times 1}$ and $G \in \mathbf{C}^{n\times 1}$ be solutions to*

$$F^T V = (e_1^r)^T \quad \text{and} \quad W^T G = e_1^r, \tag{15}$$

*where $r$ is the order of the reduced-order system. Then, projection of the intermediate system $\tilde{\Sigma}(\tilde{E}, \tilde{A}, \tilde{N}, \tilde{B}, \tilde{C}, \tilde{D})$,*

$$\begin{array}{lll}
\tilde{E} = E, & \tilde{A} = A + G\hat{D}F^T, & \tilde{N} = N,\\
\tilde{B} = B - G\hat{D}, & \tilde{C} = C - \hat{D}F^T, & \tilde{D} = CMB + D,
\end{array}$$

*where $\hat{D} = CMB$, results in a reduced-order system $\Sigma_r(E_r, A_r, N_r, B_r, C_r, D_r)$ in which*

$$E_r = W^T \tilde{E} V, \qquad A_r = W^T \tilde{A} V, \qquad N_r = W^T \tilde{N} V,$$
$$B_r = W^T \tilde{B}, \qquad C_r = \tilde{C} V_r, \qquad D_r = \tilde{D}.$$

*Assuming nonsingular $E_r$, the polynomial parts of the first subsystem associated with the reduced and original bilinear systems are matched. Also*

$$H(\mathcal{S}_k) = H_r(\mathcal{S}_k) + \mathcal{O}\left((s_1 - \mu_1)^q \cdots (s_k - \mu_k)^q (s_1 - \sigma_1)^q \cdots (s_k - \sigma_k)^q\right).$$

*Proof.* For the first subsystem, the result reduces to Theorem 2.1. However for completeness, we derive its proof in the following:

$$\begin{aligned} H(\sigma_1) - H_r(\sigma_1) = C\left((\sigma_1 E - A)^{-1} B - V(\sigma_1 E_r - A_r)^{-1} B_r\right) \\ - \hat{D} + \hat{D} F^T V(\sigma_1 E_r - A_r)^{-1} B_r. \end{aligned} \tag{16}$$

Hence,

$$\begin{aligned} V(\sigma_1 E_r - A_r)^{-1} B_r &= V(\sigma_1 E_r - A_r)^{-1} W^T (B - G\hat{D}), \\ &= V(\sigma_1 E_r - A_r)^{-1} W^T (\sigma_1 E - A - G\hat{D}F^T)(\sigma_1 E - A)^{-1} B. \end{aligned}$$

Note that $P_\sigma = V(\sigma E_r - A_r)^{-1} W^T (\sigma E - A - G\hat{D}F^T)$ is an oblique projector onto range$(V)$ and let $z \in$ range$(V)$, then $P_\sigma z = z$. This implies,

$$V(\sigma_1 E_r - A_r)^{-1} B_r = P_{\sigma_1}(\sigma_1 E - A)^{-1} B = (\sigma_1 E - A)^{-1} B. \tag{17}$$

Thus (16) becomes

$$H(\sigma_1) - H_r(\sigma_1) = -\hat{D} + \hat{D} F^T (\sigma_1 E - A)^{-1} B.$$

Now from (15), we have $F^T(\sigma_1 E - A)^{-1} B = 1$ and this proves the matching at $\sigma_1$ for the first subsystem. Similarly, $H(\mu_1) = H_r(\mu_1)$ holds. Next, we consider the second subsystem

$$\begin{aligned} H(\sigma_1, \sigma_2) - H_r(\sigma_1, \sigma_2) = C\big(&(\sigma_2 E - A)^{-1} N (\sigma_1 E - A)^{-1} B \\ &- V(\sigma_2 E_r - A_r)^{-1} N_r (\sigma_1 E_r - A_r)^{-1} B_r\big) \\ &+ \hat{D} F^T V(\sigma_2 E_r - A_r)^{-1} N_r (\sigma_1 E_r - A_r)^{-1} B_r. \end{aligned} \tag{18}$$

Since

$$V \underbrace{(\sigma_2 E_r - A_r)^{-1} N_r (\sigma_1 E_r - A_r)^{-1} B_r}_{z_r} = V(\sigma_2 E_r - A_r)^{-1} W^T N V (\sigma_1 E_r - A_r)^{-1} B_r,$$

it holds

$$V z_r = V(\sigma_2 E_r - A_r)^{-1} W^T N (\sigma_1 E - A)^{-1} B,$$

by using (17). Now using (15), the above equation becomes

$$
\begin{aligned}
Vz_r &= V(\sigma_2 E_r - A_r)^{-1} W^T \left((\sigma_2 E - A) - G\hat{D}F^T\right)(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B \\
&= P_{\sigma_2}(\sigma_2 E - A)^{-1} N(\sigma_2 E - A)^{-1} B = (\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B.
\end{aligned}
$$

Using this in (18), we have

$$H(\sigma_1, \sigma_2) - H_r(\sigma_1, \sigma_2) = \hat{D}F^T(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B.$$

From (15), we have $F^T(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B = 0$. Hence $H(\sigma_1, \sigma_2) = H_r(\sigma_1, \sigma_2)$. Similarly $H(\mu_1, \mu_2) = H_r(\mu_1, \mu_2)$. Using the same steps, we can deal with subsystems of higher order and higher derivatives. □

## 3.2 Interpolating a bilinear descriptor system and retaining the polynomial part of the first $k$ subsystems

So far, we have discussed how an interpolatory technique can retain the polynomial part of the first subsystem in the reduced bilinear system by assuming that the higher order subsystems have zero polynomial part. In this section, we consider a general case where the higher order subsystems also have non-zero polynomial parts. The goal is to construct a reduced bilinear system that retains the polynomial parts (non-zero) of the first $k$ subsystems associated with the original bilinear system, in addition to interpolating these subsystems. As discussed in the preceding section, the structure of the $kth$ subsystem of the reduced bilinear system can be written as

$$
\begin{aligned}
H_r(s_1, \ldots, s_k) = &C_r(s_k E_r - A_r)^{-1} N_r (s_{k-1} E_r - A_r)^{-1} N_r \cdots \\
&\cdots N_r(s_1 E_r - A_r)^{-1} B_r + D_k,
\end{aligned}
\tag{19}
$$

in which $E_r$ is nonsingular and $D_k$ is the polynomial part of the $kth$ subsystem of the original bilinear system. This means that the reduced bilinear system ensures matching of the polynomial parts of the first $k$ subsystems corresponding to the original bilinear system. However, we also need to ensure interpolation for these subsystems. The following theorem provides our main result for achieving this.

**Theorem 3.3.** *Let $V$ and $W$ be as defined in Theorem 3.1 and define intermediate matrices:*

$$
\begin{aligned}
&\tilde{E} = E, \qquad &&\tilde{A} = A + L_A, \qquad &&\tilde{N} = N - L_N, \\
&\tilde{B} = B - L_B, \qquad &&\tilde{C} = C - L_C,
\end{aligned}
$$

*where $L_A, L_N, L_B, L_C$ are solutions to the following equations:*

$$W^T L_B = \left[\hat{D}_1(e_1^q)^T, D_2(e_1^{q^2})^T, \ldots, D_k(e_1^{q^k})^T\right]^T, \tag{20}$$

$$L_C V = \left[\hat{D}_1(e_1^q)^T, D_2(e_1^{q^2})^T, \ldots, D_k(e_1^{q^k})^T\right], \tag{21}$$

$$L_A V = \left[L_B(e_1^q)^T, L_N\left[V^1(I_q \otimes (e_1^q)^T), \ldots, V^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T))\right]\right], \tag{22}$$

$$W^T L_A = \left[L_C^T(e_1^q)^T, L_N^T[W^1(I_q \otimes (e_1^q)^T), \ldots, W^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T)]\right]^T, \tag{23}$$

*in which* $\hat{D}_1 = D_1 - D$ *and* $e_1^{q^j}$ *is the first column of an identity matrix of size* $q^j \times q^j$*. Then, projection of the intermediate system results in a reduced-order system:*

$$\begin{aligned} E_r &= W^T\tilde{E}V, & A_r &= W^T\tilde{A}V, & N_r &= W^T\tilde{N}V, \\ B_r &= W^T\tilde{B}, & C_r &= \tilde{C}V, \end{aligned}$$

*that satisfies*

$$H(\mathcal{S}_k) = H_r(\mathcal{S}_k) + \mathcal{O}\left((s_1-\mu_1)^q \cdots (s_k-\mu_k)^q (s_1-\sigma_1)^q \cdots (s_k-\sigma_k)^q\right).$$

*Proof.* Consider the first subsystem at $s_1 = \sigma_1$:

$$\begin{aligned} H(\sigma_1) - H_r(\sigma_1) = C\big((s_1E-A)^{-1}B - V(s_1E_r-A_r)^{-1}B_r\big) \\ + L_CV(s_1E_r-A_r)^{-1}B_r - (\underbrace{D_1-D}_{\hat{D}_1}). \end{aligned} \tag{24}$$

Since, $V(\sigma_1E_r-A_r)^{-1}B_r = V(\sigma_1E_r-A_r)^{-1}W^T(B-L_B)$ and from (22), $L_A(\sigma_1E-A)^{-1}B = L_B$, therefore

$$V(\sigma_1E_r-A_r)^{-1}B_r = V(\sigma_1E_r-A_r)^{-1}W^T((\sigma_1E-A)-L_A)(\sigma_1E-A)^{-1}B.$$

Now, introducing the oblique projector $P_\sigma = V(\sigma E_r-A_r)^{-1}W^T\left((\sigma E-A)-L_A\right)$ and utilizing $P_\sigma z = z$ for $z \in \text{range}(V)$, we get

$$V(\sigma_1E_r-A_r)^{-1}B_r = P_{\sigma_1}(\sigma_1E-A)^{-1}B = (\sigma_1E-A)^{-1}B. \tag{25}$$

Using this in (24), we get

$$H(\sigma_1) - H_r(\sigma_1) = L_C(s_1E-A)^{-1}B - \hat{D}_1.$$

From (21), $L_C(\sigma_1E-A)^{-1}B = \hat{D}_1$. Hence $H(\sigma_1) = H_r(\sigma_1)$. Similarly, $H(\mu_1) = H_r(\mu_1)$ holds. Now, consider the second subsystem

$$\begin{aligned} H(\sigma_1,\sigma_2) - H_r(\sigma_1,\sigma_2) = C\big(&(\sigma_2E-A)^{-1}N(\sigma_1E-A)^{-1}B \\ &- V(\sigma_2E_r-A_r)^{-1}N_r(\sigma_1E_r-A_r)^{-1}B_r\big) \\ &+ L_CV(\sigma_2E_r-A_r)^{-1}N_r(\sigma_1E_r-A_r)^{-1}B_r - D_2. \end{aligned} \tag{26}$$

Hence,

$$\begin{aligned} V\underbrace{(\sigma_2E_r-A_r)^{-1}N_r(\sigma_1E_r-A_r)^{-1}B_r}_{\hat{z}_r} &= V(\sigma_2E_r-A_r)^{-1}W^T(N-L_N)V(\sigma_1E_r-A_r)^{-1}B_r, \\ &= V(\sigma_2E_r-A_r)^{-1}W^T(N-L_N)(\sigma_1E-A)^{-1}B, \end{aligned}$$

where the last equation follows from (25). Now (22) implies

$$L_A(\sigma_2E-A)^{-1}N(\sigma_1E-A)^{-1}B = L_N(\sigma_1E-A)^{-1}B.$$

Thus,

$$\begin{aligned} V\hat{z}_r &= V(\sigma_2 E_r - A_r)^{-1} W^T((\sigma_2 E - A) - L_A)(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B, \\ &= P_{\sigma_2}(\sigma_2 E - A) N(\sigma_2 E - A)^{-1} B = (\sigma_2 E - A) N(\sigma_1 E - A)^{-1} B. \end{aligned}$$

Using this in (18), we have

$$H(\sigma_1, \sigma_2) - H_r(\sigma_1, \sigma_2) = L_C(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B - D_2.$$

Now from (21), $L_C(\sigma_2 E - A)^{-1} N(\sigma_1 E - A)^{-1} B = D_2$ and therefore $H(\sigma_1, \sigma_2) = H_r(\sigma_1, \sigma_2)$. Similarly, $H(\mu_1, \mu_2) = H_r(\mu_1, \mu_2)$. Using similar steps, we can also deal with higher subsystems and higher derivatives. □

# 4 Computational Issues and Time-domain Representation of the Reduced-order System

In this section, we discuss the computational issues associated with the intermediate system $\tilde{\Sigma}$. As discussed before, the intermediate system requires the solution of (20)–(23) for matrices $L_A, L_N, L_B$ and $L_C$. Since $L_B$ and $L_C$ are independent of other unknowns, they can be easily computed. However, the main issue is the computation of $L_A$ and $L_N$. These matrices require the simultaneous solution of (22) and (23) for given $L_B$ and $L_C$. To ensure the existence of the simultaneous solution, we derive a necessary condition, called the compatibility condition. This follows by equating the right hand sides of (22) and (23) after pre-multiplying by $W^T$ and post-multiplying by $V$, respectively:

$$\begin{aligned} W^T \left[L_B (e_1^q)^T, L_N \left[V^1 (I_q \otimes (e_1^q)^T), \ldots, V^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T))\right]\right] = \\ [L_C^T (e_1^q)^T, L_N^T [W^1 (I_q \otimes (e_1^q)^T), \ldots, W^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T)]]^T V. \end{aligned} \tag{27}$$

The following theorem guarantees that the above compatibility condition is satisfied.

**Theorem 4.1.** *Assume $L_B$ and $L_B$ satisfy* (20) *and* (21)*, respectively, and let $L_N \in \mathbb{R}^{n\times n}$ satisfy,*

$$W^T L_N V = \mathcal{T} \begin{bmatrix} D_2 & \cdots & D_{k+1} \\ \vdots & \ddots & \vdots \\ D_{k+1} & \cdots & D_{2k} \end{bmatrix} \mathcal{T}^T, \tag{28}$$

*where $\mathcal{T} = \sum_{i=0}^{k-1} (e_{1+q^i}^r)^T \otimes e_{i+1}^k$, $r = q + \cdots + q^k$ is the order of the reduced-order system and $D_j$ is the polynomial part of the $j$th subsystem. Then, the compatibility condition* (27) *is satisfied.*

*Proof.* Consider the first row of the block matrix given in (27):

$$\left(W^1\right)^T \left[L_B (e_1^q)^T, L_N \left[V^1 (I_q \otimes (e_1^q)^T), \ldots, V^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T))\right]\right] = e_1^q L_C V.$$

To show that the above equation holds, we use (20) and (28)

$$\begin{aligned}
&\left(W^1\right)^T \left[L_B(e_1^q)^T, L_N\left[V^1(I_q \otimes (e_1^q)^T), \dots, V^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T)\right)\right]\right] \\
&\quad = [e_1^q \hat{D}_1(e_1^q)^T, e_1^q D_2(e_1^q)^T(I_q \otimes (e_1^q)^T), \dots, e_1^q D_k(e_1^{q^{k-1}})^T(I_{q^{k-1}} \otimes (e_1^q)^T)], \\
&\quad = e_1^q[\hat{D}_1(e_1^q)^T, D_2(e_1^q)^T \otimes (e_1^q)^T, \dots, D_k(e_1^{q^{k-1}})^T \otimes (e_1^q)^T], \\
&\quad = e_1^q[\hat{D}_1(e_1^q)^T, D_2(e_1^{q^2})^T, \dots, D_k(e_1^{q^k})^T], \\
&\quad = e_1^q L_C V,
\end{aligned}$$

where the last equality follows from (21). Now consider the *ith* row of the block matrix in (27):

$$\begin{aligned}
&\left(W^i\right)^T \left[L_B(e_1^q)^T, L_N\left[V^1(I_q \otimes (e_1^q)^T), \dots, V^{k-1}(I_{q^{k-1}} \otimes (e_1^q)^T)\right)\right]\right] \\
&\qquad = \left[L_N^T W^{i-1}(I_{q^{i-1}} \otimes (e_1^q)^T)\right]^T V, \\
&\qquad = \left[e_1^{q^i} D_i(e_1^q)^T, e_1^{q^i} D_{i+1}(e_1^q)^T(I_q \otimes (e_1^q)^T), \dots, e_1^{q^i} D_{i+k-1}(e_1^{q^{k-1}})^T(I_{q^{k-1}} \otimes (e_1^q)^T)\right], \\
&\qquad = (I_{q^{i-1}} \otimes (e_1^q)^T)^T e_1^{q^{i-1}} \left[D_i(e_1^q)^T, D_{i+1}(e_1^{q^2})^T, \dots, D_{i+k-1}(e_1^{q^k})^T\right], \\
&\qquad =: \mathcal{R}.
\end{aligned}$$

Using the condition on $L_N$ given in (28), we obtain

$$\begin{aligned}
\mathcal{R} &= (I_{q^{i-1}} \otimes (e_1^q)^T)^T \left[(W^{i-1})^T L_N V^1, (W^{i-1})^T L_N V^2, \dots, (W^{i-1})^T L_N V^k\right], \\
&= (I_{q^{i-1}} \otimes (e_1^q)^T)^T \left[(W^{i-1})^T L_N V^1, (W^{i-1})^T L_N V^2, \dots, (W^{i-1})^T L_N V^k\right], \\
&= \left[L_N^T W^{i-1}(I_{q^{i-1}} \otimes (e_1^q)^T)^T\right]^T V.
\end{aligned}$$

This means that each row of the block matrix corresponding to the left and right side of the compatibility condition given in (27) are equal. Therefore, if $L_N$ is chosen to satisfy the assumption (28), then it is ensured that (27) holds. □

**Remark 4.1.** *It is interesting to see that in order to compute the reduced model, we do not need to compute explicitly the matrices $L_A$, $L_N$, $L_B$ and $L_C$. We only require the expressions for $W^T L_B$, $L_C V$, $W^T L_A V$ and $W^T L_N V$. One can substitute $W^T L_B$ and $L_C V$ directly from* (20) *and* (21)*. The expression of $W^T L_N V$ can be easily identified using* (28)*. Similarly, one can obtain the expression of $W^T L_A V$ without explicitly computing $L_A$ by pre-multiplying (22) by $W^T$ and use* (28) *and* (20)*.*

Now, we summarize the complete methodology of computing the reduced-order system for the system (8) in the following algorithm.

**Algorithm 1** Interpolatory Model-Order Reduction for Bilinear Systems

1: **Input:** $E, A, N, B, C, D, [\sigma_1, \cdots, \sigma_k], [\mu_1, \cdots, \mu_k], q$.

2: **Output:** $E_r, A_r, N_r, B_r, C_r$.

3: Construct $V$ and $W$ from Theorem 3.1.

4: Compute the polynomial part of the $kth$ subsystem
$$D_k = C(MN)^{k-1}MB + D\delta(k-1).$$

5: Identify the expression of $W^T L_B$, $L_C V$ , $W^T L_A V$ and $W^T L_N V$ as
$$W^T L_B = \left[\hat{D}_1(e_1^q)^T, D_2(e_1^{q^2})^T, \ldots, D_k(e_1^{q^k})^T\right]^T =: R_B, \quad \hat{D}_1 = D_1 - D.$$
$$W^T L_B = \left[\hat{D}_1(e_1^q)^T, D_2(e_1^{q^2})^T, \ldots, D_k(e_1^{q^k})^T\right] =: R_C.$$
$$W^T L_N V = \mathcal{T} \begin{bmatrix} D_2 & \cdots & D_{k+1} \\ \vdots & \ddots & \vdots \\ D_{k+1} & \cdots & D_{2k} \end{bmatrix} \mathcal{T}^T =: R_N,$$
where $\mathcal{T} = \sum\limits_{i=0}^{k-1} (e_{1+q^i}^r)^T \otimes e_{i+1}^k$ and $r = \sum_{i=1}^k q^i + \cdots + q^k$.
$$\begin{aligned} W^T L_A V = \big[&R_B(e_1^q)^T, R_N(:, 1:q)(I_q \otimes (e_1^q)^T), \ldots, \\ &R_N(:, q + \cdots + q^{k-1} + (1:q^k))(I_{q^{k-1}} \otimes (e_1^q)^T)\big] =: R_A \end{aligned}$$

6: Compute the reduced model as
$$E_r = W^T E V, \qquad A_r = W^T A V + R_A, \quad N_r = W^T N V - R_N,$$
$$B_r = W^T B - R_B, \quad C_r = CV - R_C.$$

**Remark 4.2.** *As shown in [11], two-sided projections might lead to much better approximation since more multi-moments are matched for higher order subsystems. The same holds for the proposed modified Krylov subspace technique. To see this we consider an example similar to the one used in [11]. Let us assume the projection subspaces $V$ and $W$ are as follows*

$$\begin{aligned} \mathrm{span}(V) &= \mathrm{span}\left\{A^{-1}B, \ldots, (A^{-1}E)^5 A^{-1}B, A^{-1}NA^{-1}B, A^{-1}N(A^{-1}E)A^{-1}B\right\}, \\ \mathrm{span}(W^T) &= \mathrm{span}\left\{CA^{-1}, \ldots, C(A^{-1}E)^5 A^{-1}, CA^{-1}NA^{-1}, C(A^{-1}E)A^{-1}NA^{-1}\right\}. \end{aligned}$$

*According to Theorem 3.3, the reduced model preserves* 12 *multi-moments of the first subsystem*

$$C(A^{-1}E)^{l_1}A^{-1}B + D\delta(l_1) = C_r^T(A_r^{-1}E_r)^{l_1}A_r^{-1}B_r + D_1\delta(l_1),$$

*where $l_1 = 0, \ldots, 11$. For the second subsystem, 29 multi-moments match*

$$C(A^{-1}E)^{l_2}A^{-1}N(A^{-1}E)^{l_1}A^{-1}B = C_r^T(A_r^{-1}E_r)^{l_2}A_r^{-1}N_r(A_r^{-1}E_r)^{l_1}A_r^{-1}B_r + D_2\delta(l_1)\delta(l_2),$$

*where, $l_1, l_2 = 0, 1, \ldots, 5$ or $l_1 = 6,\ l_2 = 0, 1$ and $l_1 = 0, 1,\ l_2 = 6$. For the third subsystem, 37 multi-moments match*

$$\begin{aligned} C(A^{-1}E)^{l_3}A^{-1}N \cdots N(A^{-1}E)^{l_1}A^{-1}B = {}& C_r^T(A_r^{-1}E_r)^{l_3}A_r^{-1}N_r \cdots N_r(A_r^{-1}E_r)^{l_1}A_r^{-1}B_r \\ &+ D_3\delta(l_1)\delta(l_2)\delta(l_3), \end{aligned}$$

*where* $l_1 = 0, 1, \dots, 5, l_2 = 0, l_3 = 0, 1$ *or* $l_1 = 0, 1, l_2 = 0, l_3 = 2, 3, 4, 5$ *or* $l_1 = 0, 1, l_2 = 1, l_3 = 0, 1$. *For the fourth subsystem, 4 multi-moments match*

$$\begin{aligned} C((A^{-1}E)^{l_4}A^{-1}N\cdots N(A^{-1}E)^{l_1}A^{-1}B = {} & C_r((A_r^{-1}E_r)^{l_4}A_r^{-1}N_r\cdots N_r(A_r^{-1}E_r)^{l_1}A_r^{-1}B_r \\ & + D_4\delta(l_1)\delta(l_2)\delta(l_3)\delta(l_4), \end{aligned}$$

*where* $l_1 = 0, 1, l_2 = 0, l_3 = 0, l_4 = 0, 1$.

## 4.1 Time-Domain Representation

Till now we have shown how to achieve interpolation for the leading $k$ subsystems along with retaining their polynomial parts. In this subsection we show the time domain representation of the reduced bilinear system whose $kth$ order subsystem is of the form given in (19). The following theorem summarizes our results.

**Theorem 4.2.** *Given a bilinear system, whose kth order transfer function has the form given in* (19)*. Then, the time domain representation of this bilinear system can be written as*

$$\begin{aligned} E_r\dot{x}_r(t) &= A_rx_r(t) + N_rx_r(t)u(t) + B_ru(t), \\ y_r(t) &= C_rx_r(t) + \sum_{k=1}^{\infty}D_ku^k(t). \end{aligned} \tag{29}$$

*Proof.* We begin with the $kth$ order transfer function

$$\begin{aligned} H_r(s_1,\dots,s_k) &= C_r(s_kE_r - A_r)^{-1}N_r(s_{k-1}E_r - A_r)^{-1}N_r\cdots \\ &\qquad \cdots N_r(s_1E_r - A_r)^{-1}B_r + D_k, \\ &= \hat{C}(s_kI_r - \hat{A})^{-1}\hat{N}(s_{k-1}I_r - \hat{A})^{-1}\hat{N}\cdots \\ &\qquad \cdots \hat{N}(s_1I_r - \hat{A})^{-1}\hat{B} + D_k, \end{aligned} \tag{30}$$

where

$$\hat{A} = E_r^{-1}A_r, \quad \hat{N} = E_r^{-1}N_r, \quad \hat{B} = E_r^{-1}B_r, \quad \text{and} \quad \hat{C} = C_r. \tag{31}$$

Now by utilizing the multivariate inverse Laplace transform on (30), we obtain the regular Volterra kernel as

$$h_k(t_1,t_2,\dots,t_k) = \hat{C}e^{\hat{A}t_k}\hat{N}e^{\hat{A}t_{k-1}}\hat{N}\cdots\hat{N}e^{\hat{A}t_1}\hat{B} + D_k\delta(t_k)\delta(t_{k-1})\cdots\delta(t_1). \tag{32}$$

As discussed in [22], the output $y_r(t)$ of a nonlinear system can be described in terms of the Volterra kernel $h_k(t_1,t_2,\dots,t_k)$ and input $u(t)$ as follows:

$$y_r(t) = \sum_{k=1}^{\infty}\int_0^{t_1}\int_0^{t_2}\cdots\int_0^{t_k}h_k(t_1,t_2,\dots,t_k)u(t-\sum_{i=1}^{k}t_i)\cdots u(t-t_k)dt_k\cdots dt_1.$$

Substituting (32) in the above equation, we can write

$$y_r(t) = y_r^{(1)}(t) + y_r^{(2)}(t),$$

where

$$y_r^{(1)}(t) = \sum_{k=1}^{\infty} \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_k} \hat{C} e^{\hat{A}t_k} \hat{N} \cdots \hat{N} e^{\hat{A}t_2} \hat{N} e^{\hat{A}t_1} \hat{B} u(t - \sum_{i=1}^{k} t_i) \cdots u(t - t_k) dt_k \cdots dt_1,$$

$$y_r^{(2)}(t) = \sum_{k=1}^{\infty} \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_k} D_k \delta(t_k)\delta(t_{k-1}) \cdots \delta(t_1) u(t - \sum_{i=1}^{k} t_i) \cdots u(t - t_k) dt_k \cdots dt_1.$$

The response $y_r^{(1)}(t)$ is simply the Volterra series representation of a bilinear ODE system with zero initial condition [22]. This means that corresponding to $y_r^{(1)}(t)$, we have

$$\begin{aligned} \dot{x}_r(t) &= \hat{A}x_r(t) + \hat{N}x_r(t)u(t) + \hat{B}u(t), \\ y_r^{(1)}(t) &= \hat{C}x_r(t), \qquad x_r(0) = 0. \end{aligned} \tag{33}$$

For $y_r^{(2)}(t)$, we use the properties of the Dirac delta function [12] which leads to

$$y_r^{(2)}(t) = \sum_{k=1}^{\infty} D_k u(t) \cdots u(t) = \sum_{k=1}^{\infty} D_k u^k(t).$$

By combining the responses $y_r^{(1)}(t)$ and $y_r^{(2)}(t)$ and substituting the expression of $\hat{A}, \hat{N}, \hat{B}$ and $\hat{C}$ from (31), we obtain a bilinear system as in (29) and this proves the theorem. $\square$

Since the output equation in (29) contains the sum of an input dependent infinite series, we need to compute the summation at each time step. This increases the computational cost and may suppress the importance of the model reduction procedure. In the following, we discuss some cases where this infinite summation can be computed cheaply.

**Case 1:** For the particular structure of $N$ as in (14), $D_k = 0$ for all $k > 1$. Thus $\sum_{k=1}^{\infty} D_k u^k(t)$ reduces to $D_1 u(t)$, which is computationally cheap.

**Case 2:** There are some applications where the input $u(t)$ can be considered constant or unity ($u(t) = \alpha$ or $u(t) = 1$). These scenarios may appear, for example in the parameter varying systems [6]. In such a case

$$\sum_{k=1}^{\infty} D_k u^k(t) = D_1\alpha + D_2\alpha^2 + D_3\alpha^3 \cdots .$$

Substituting the expression of $D_k$ from Lemma 3.1 in the above equation, we get

$$\sum_{k=1}^{\infty} D_k u^k(t) = (CMB + D)\alpha + C(MN)MB\alpha^2 + C(MN)^2 MB\alpha^3 + \cdots,$$
$$= \alpha C(I + \alpha MN + \alpha^2 (MN)^2 + \cdots)MB + \alpha D.$$

Now, if we assume $\|\alpha MN\|_2 < 1$, we have

$$\sum_{k=1}^{\infty} D_k u^k(t) = (C(I - \alpha MN)^{-1} MB + D)\alpha.$$

Thus we can identify an expression of the convergent series for constant inputs.

**Case 3:** In this case, we assume convergence for $\|D_k\|$, i.e. $\sum_{k=j+1}^{\infty} \|D_k\| < \tau \ll 1$. Then for bounded inputs, we can truncate the infinite summation after the $jth$ term. That is

$$\sum_{k=1}^{\infty} D_k u^k(t) \approx \sum_{k=1}^{j} D_k u^k(t).$$

Thus, we can save the computations associated with $\sum_{k=j+1}^{\infty} D_k u^k(t)$.

# 5 Numerical Experiments

In this section we present numerical results for model reduction of a structured bilinear DAE system using different approaches. The reduced-order system can be computed either by direct implementation of Theorem 3.1, without matching the polynomial part in the reduced-order system (classical interpolatory technique) or by our proposed methodology which achieves matching of the polynomial part in addition to interpolation. All the numerical results were simulated in MATLAB® version 7.11.0.584(R2010b) 64-bit (glnza64) on Intel(R) Core(TM)2 Quad CPU Q9550 @ 2.83GHz, 6 MB cache, 4GB RAM, openSUSE 12.1 (x86-64).

## 5.1 Artificial Example

The bilinear DAE system, that is to be reduced, is generated randomly with order $n = 100$ and with partitioning $n_1 = 90,\ n_2 = 10$. It is ensured that the structure of the pencil is similar to the index-1 pencil of a linear DAE. The polynomial parts of the first 4 subsystems of the bilinear system are $D_1 = 0.1472, D_2 = 5 \times 10^{-3}, D_3 = 1.92 \times 10^{-4}, D_4 = 7.35 \times 10^{-6}$, where $D_i$ is the polynomial part of the $ith$ subsystem. The interpolation points are selected as $\sigma = \mu = [0, 0.5]$ with multiplicity $q = 1$ resulting in a reduced-order system of order $r = 4$. We truncate the infinite summation in Lemma 4.1 after first 4 terms since $\|D_i\|$ decreases exponentially.

We compute the reduced model using the classical interpolation technique and the proposed methodology using the same interpolation points and multiplicity. The time response of the actual and the reduced bilinear systems, obtained by using the implicit Euler method, is shown in Figure 1 for an exponential input. The relative errors associated with the two approaches are shown in Figure 2.

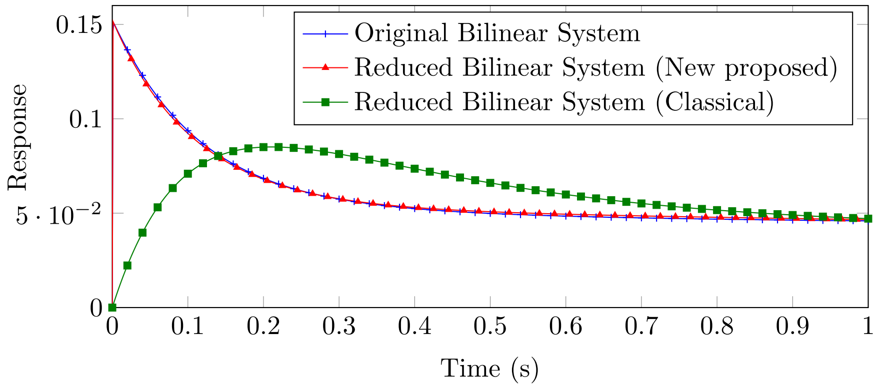


Figure 1: Transient response, $u(t) = e^{-10t}$.

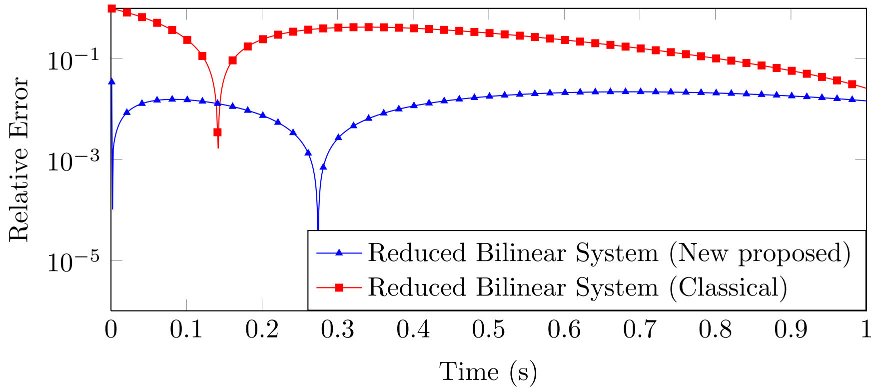


Figure 2: Relative Error, $u(t) = e^{-10t}$.

Certainly, the reduced-order system obtained from the direct implementation shows completely different dynamics whereas the proposed methodology captures the dynamics of the original system well.

## 5.2 Nonlinear RC Circuit

As a second example, we consider a nonlinear RC circuit that represents a modified form of the transmission line circuit proposed in [16]. The circuit includes resistors, capacitors and diodes as shown in Figure 3.

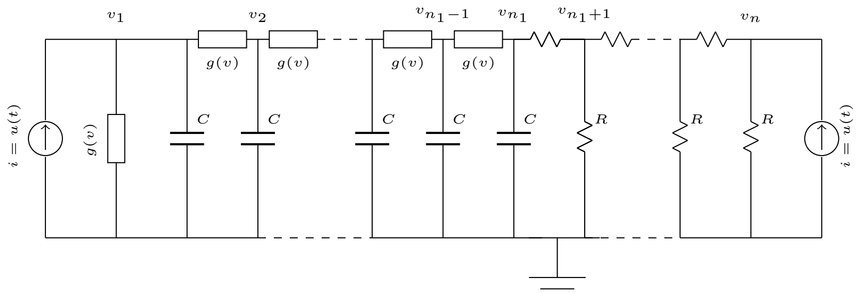


Figure 3: Nonlinear transmission line circuit.

All the resistances and capacities are set to 1 and all the diodes ensure $i_D = e^{40v_D} + v_D - 1$, where $i_D$ represents current and $v_D$ voltage across the diodes. The input $u(t)$ is the current source $i$ and the output $y(t)$ represents the average voltage over all nodes ranging from 1 to $n$.

Using Kirchhoff's current law, at each node we have

$$\begin{aligned}
\dot{v}_1 &= -2v_1 + v_2 + 2 - e^{40v_1} - e^{40(v_1 - v_2)} + u(t), \\
\dot{v}_k &= -2v_k + v_{k-1} + v_{k+1} + e^{40(v_{k-1} - v_k)} - e^{40(v_k - v_{k+1})}, \quad (2 \leq k \leq n_1 - 1) \\
\dot{v}_{n_1} &= -2v_{n_1} + v_{n_1-1} + v_{n_1+1} - 1 + e^{40(v_{n_1-1} - v_{n_1})}, \\
0 &= 3v_k - v_{k-1} - v_{k+1}, \qquad (n_1 + 1 \leq k \leq n - 1) \\
0 &= -2v_n + v_{n-1} + u(t).
\end{aligned} \tag{34}$$

In order to represent the above nonlinear system as a quadratic-bilinear system, we set $v_1$ to $v_{k,k+1}$ $(v_{k,k+1} = v_k - v_{k+1})$, $k = 1, \ldots, n_1 - 1$ and $v_{n_1+1}$ to $v_n$ as the state variables and perform some changes of variables by defining $y_1 = e^{40v_1} - 1$ and $y_k = e^{40(v_{k-1,k})} - 1, \; 2 \leq k \leq n_1$. Together with the differential equations of all $y_k$'s, the above changes lead to the following quadratic system

$$
\begin{aligned}
\dot{v}_1 &= -v_1 - v_{1,2} - y_1 - y_2 + u(t),\\
\dot{v}_{1,2} &= -v_1 - 2v_{1,2} + v_{2,3} - y_1 - 2y_2 + y_3 + u(t),\\
\dot{v}_{k,k+1} &= -2v_{k,k+1} + v_{k-1,k} + v_{k+1,k+2} + y_k - 2y_{k+1} + y_{k+2}, \quad (2 \leq k \leq n_1 - 2)\\
\dot{v}_{n_1-1,n_1} &= -2v_{n_1-1,n_1} + v_{n_1-2,n_1-1} + v_{n_1} - v_{n_1+1} + y_{n_1-1} - 2y_{n_1},\\
0 &= 3v_k - v_{k-1} - v_{k+1}, \qquad (n_1 + 1 \leq k \leq n - 1)\\
0 &= -2v_n + v_{n-1} + u(t),\\
\dot{y_1} &= 40(y_1 + 1)(-v_1 - v_{1,2} - y_1 - y_2 + u(t)),\\
\dot{y_2} &= 40(y_2 + 1)(-v_1 - 2v_{1,2} + v_{2,3} - y_1 - 2y_2 + y_3 + u(t)),\\
\dot{y_k} &= 40(y_k + 1)(-2v_{k-1,k} + v_{k-2,k-1} + v_{k,k+1} + y_{k-1} - 2y_k + y_{k+1}),\\
\dot{y_{n_1}} &= 40(y_{n_1} + 1)(-2v_{n_1-1,l_1} + v_{n_1-2,n_1-1} + v_{n_1} - v_{n_1-1} + y_{n_1-1} - 2y_{n_1}).
\end{aligned}
$$

In the above set of equations, we fixed $v_{n_1}$ to $v_1 - \sum_{i=2}^{n_1} v_{k-1,k}$. This means that the circuit can be modelled by a quadratic-bilinear descriptor system of order $\tilde{n} = n_1 + n$. It is easy to see that the matrix pencil associated with this quadratic system has nilpotency index-1. Next, we utilize the Carleman bilinearization, thus ensuring that the resulting bilinearized system also has an index-1 matrix pencil [15]. The order of the bilinearized DAE system is $N = (n_1 + n)(2n_1 + 1)$. The polynomial part of the first subsystem of the bilinearized system is $D_1 = 0.0333$ and higher order subsystems have zero polynomial part.

For our experiment, we choose $n_1 = 10$ and $n = 30$. The bilinearized system is therefore of order $N = 840$. Using Theorem 3.1, we compute the projection matrices such that the reduced-order system guarantees interpolation of the first two subsystems at $\sigma = \mu = [10, 50, 300]$. The multiplicities of all the interpolation points are set to 1. The reduced models of the bilinearized system are computed using the classical and the proposed methodology using the same interpolation points and multiplicities. Since we do not have specific criteria yet to choose this interpolation points and their multiplicities which can ensure a stable reduced-order model for both the modified and the classical method. For our result, it is possible to get stable reduced-order models using this methodology for the same interpolation points and same multiplicities in case of one-sided projection, i.e. $W = V$.

The time response of the resulting reduced-order bilinear systems is shown in Figure 4 by utilizing the implicit Euler method. Also the absolute error ($|y - \hat{y}|$) is shown in Figure 5.

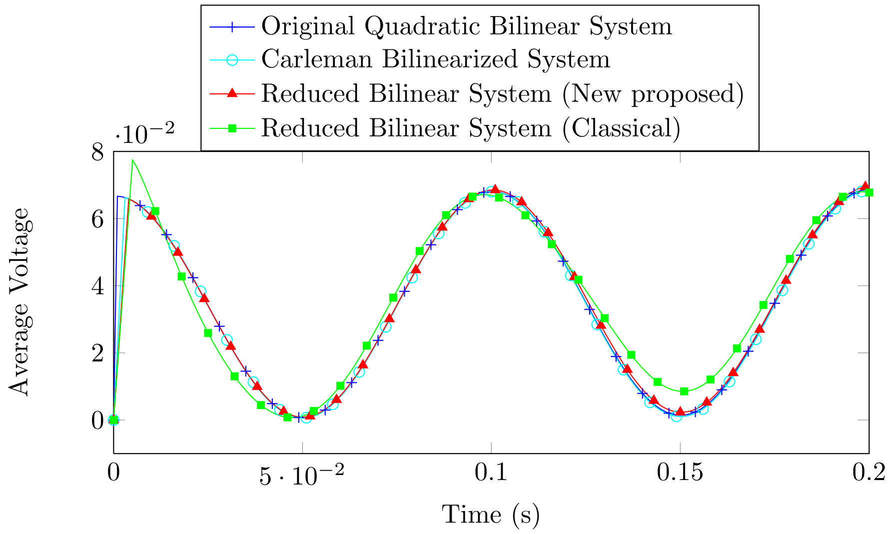


Figure 4: The time response of the original and the reduced models with input $u(t) = cos(20\pi t) + 1$.

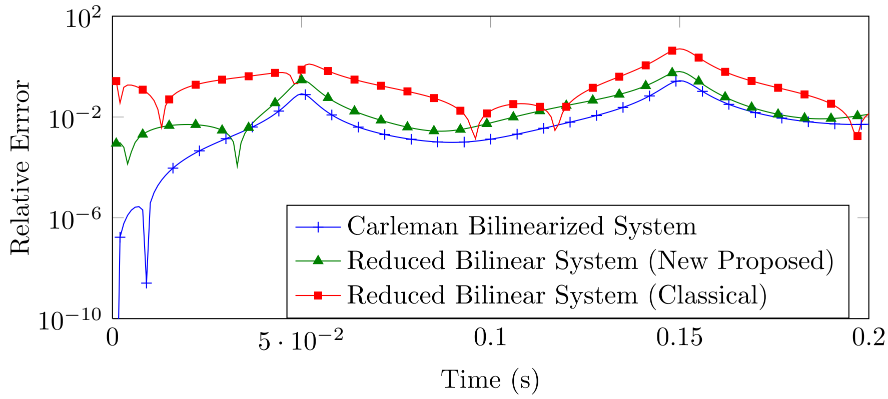


Figure 5: The absolute error for the original bilinearized system and the reduced bilinear systems with input $u(t) = cos(20\pi t) + 1$.

Clearly the proposed interpolatory technique shows substantial improvement in the transient response of the system.

## 6 Conclusions

We proposed interpolatory techniques for a special class of bilinear descriptor system with particular attention to its polynomial part. An expression that explicitly identifies the polynomial part of each subsystem associated with the bilinear system has been derived. This extends the expression for the polynomial part of linear index-1 DAE systems discussed in [18] to bilinear systems. Also, we have derived conditions on interpolatory subspaces that not only guarantees interpolation of the first $k$ subsystems but also retains the polynomial part of the bilinear system. The approach can also be applied to other bilinear DAE structures as long as the polynomial part of each subsystem is constant.

## References


[1] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems.* SIAM Publications, Philadelphia, PA, 2005.

[2] Z. Bai. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Appl. Numer. Math*, 43(1–2):9–44, 2002.

[3] Z. Bai and D. Skoogh. A projection method for model reduction of bilinear dynamical systems. *Linear Algebra Appl.*, 415(2–3):406–425, 2006.

[4] U. Baur, P. Benner, and L. Feng. Model order reduction for linear and nonlinear systems: a system-theoretic perspective. *Archives of Computational Methods in Engineering*, 21(4):331–358, 2014.

[5] C. Beattie and S. Gugercin. Interpolatory projection methods for structure-preserving model reduction. *Systems & Control Letters*, 58(3):225–232, 2009.

[6] P. Benner and T. Breiten. On $\mathcal{H}_2$-model reduction of linear parameter-varying systems. *PAMM*, 11(1):805–806, 2011.

[7] P. Benner and T. Breiten. Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems. *SIAM J. Matrix Anal. Appl.*, 33(3):859–885, 2012.

[8] P. Benner and T. Breiten. Low rank methods for a class of generalized lyapunov equations and related issues. *Numerische Mathematik*, 124(3):441–470, 2013.

[9] P. Benner and T. Damm. Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems. *SIAM J. Cont. Optim.*, 49(2):686–711, 2011.

[10] P. Benner, V. Mehrmann, and D.C. Sorensen. *Dimension Reduction of Large-Scale Systems*, volume 45 of *Lecture Notes in Computational Science and Engineering.* Springer-Verlag, Berlin/Heidelberg, Germany, 2005.

[11] T. Breiten and T. Damm. Krylov subspace methods for model order reduction of bilinear control systems. *Sys. Control Lett.*, 59(10):443–450, 2010.

[12] S. Chakraborty. Some applications of Dirac's delta function in statistics for more than one random variable. *Applications and Applied Mathematics*, 3(1):42–54, 2008.

[13] L. Feng and P. Benner. A note on projection techniques for model order reduction of bilinear systems. *Numerical Analysis and Applied Mathematics, AIP Conference Proceedings*, 936:208–211, 2007.

[14] R. W. Freund. Model reduction methods based on Krylov subspaces. *Acta Numer.*, 12:267–319, 2003.

[15] P. Goyal, M.I. Ahmad, and P. Benner. Model reduction of quadratic-bilinear descriptor systems via Carleman bilinearization. *Proceedings of European Control Conference, to appear*, 2015.

[16] C. Gu. QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 30(9):1307–1320, 2011.

[17] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30(2):609–638, 2008.

[18] S. Gugercin, T. Stykel, and S. Wyatt. Model reduction of descriptor systems by interpolatory projection methods. *SIAM J. Sci. Comput.*, 35(5):B1010–B1033, 2013.

[19] A. J. Mayo and A. C. Antoulas. A framework for the solution of the generalized realization problem. *Linear Algebra Appl.*, 425(2-3):634–662, 2007.

[20] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, AC-26(1):17–32, 1981.

[21] J. R. Phillips. Projection-based approaces for model reduction of weakly nonlinear, time-varying systems. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, 22(2):171–187, 2003.

[22] W. J. Rugh. *Nonlinear System Theory.* Johns Hopkins University Press, Baltimore, 1981.

[23] W. H. A. Schilders, H. A. Van der Vorst, and J. Rommes. *Model Order Reduction: Theory, Research Aspects and Applications.* Springer-Verlag, Berlin, Heidelberg, 2008.

[24] L. Zhang and J. Lam. On $H_2$ model reduction of bilinear systems. *Automatica J. IFAC*, 38(2):205–216, 2002.