MAX-PLANCK-GESELLSCHAFT

Max Planck Institute Magdeburg
Preprints

Peter Benner Pawan Goyal

Multipoint Interpolation of Volterra Series and $\mathcal{H}_2$-Model Reduction for a Family of Bilinear Descriptor Systems

MAX–PLANCK–INSTITUT
FÜR DYNAMIK KOMPLEXER
TECHNISCHER SYSTEME
MAGDEBURG

MPIMD/15-16 September 30, 2015

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

Abstract

In this paper, we investigate interpolatory model order reduction for large-scale bilinear descriptor systems. Recently, it was shown in [14] for linear descriptor systems that directly extending the standard rational interpolation conditions used in $\mathcal{H}_2$ optimal model reduction to descriptor systems in general yields an unbounded error in the $\mathcal{H}_2$-norm. This is due to the possible mismatch of the polynomial part of the original and reduced-order systems. This conclusion also holds for nonlinear systems as well. In this paper, we deal with bilinear descriptor systems and aim to pay attention to the polynomial part of the bilinear descriptor system along with interpolation. To this end, we have shown in [12] how to determine the polynomial part of each subsystem of the bilinear descriptor system explicitly, by assuming special structures of the system matrices. Considering the same structured bilinear descriptor systems, in this paper we first show how to achieve multipoint interpolation of the underlying Volterra series of bilinear descriptor systems while retaining the polynomial part of each subsystem of the bilinear system. Then, we extend the interpolation based first-order necessary conditions for $\mathcal{H}_2$ optimality to bilinear descriptor systems and propose an iterative scheme to obtain an $\mathcal{H}_2$ optimal reduced-order system. By mean of two numerical examples, we demonstrate the efficiency of the proposed model-order reduction technique and compare it with reduced bilinear systems obtained by using linear *IRKA*.

Key words: Interpolatory model reduction, bilinear descriptor systems, Volterra series, $\mathcal{H}_2$ optimality.

Author's addresses:

Peter Benner
Computational Methods in Systems and Control Theory
Max Planck Institute for Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg, Germany
email address: `benner@mpi-magdeburg.mpg.de`

Pawan Goyal
Computational Methods in Systems and Control Theory
Max Planck Institute for Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg, Germany
email address: `goyalp@mpi-magdeburg.mpg.de`

1 Introduction

Model order reduction (MOR) plays a vital role in numerical simulation of large-scale complex dynamical systems. These dynamical systems are governed by ordinary differential equations (ODEs), or partial differential equations (PDEs), or both. To capture the essential information about the dynamics of the systems, a fine semi-discretization of these governing equations in the spatial domain is required, leading to a large-scale system of ODEs or, in general, differential algebraic equations (DAEs). The simulation, control and optimization studies of such large-scale complex systems are numerically cumbersome and often not efficient. Thus, MOR provides a remedy to accelerate the simulation of such large-scale systems and seeks to determine low-dimensional surrogate systems with acceptable accuracy.

Model reduction for linear ODE systems has been studied for many years by now and is very well-established; see, e.g., [1, 3, 8, 14, 21]. However, there are many open and challenging problems with regard to model reduction for nonlinear systems. In this paper, we consider bilinear differential algebraic equations (DAEs) which are of the form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + \sum_{k=1}^{m} N_k x(t) u_k(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{1.1}$$

where $E, A, N_k \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$, and the matrix E is singular. It is assumed that the matrix pencil $(\lambda E - A)$ is stable, this means that all finite eigenvalues of the matrix pencil $(\lambda E - A)$ lie in the negative half plane. Such systems can be considered as weakly nonlinear systems [19]. They are linear in the state and input independently, but not jointly. Bilinear systems appear in various applications, for example, biology, nuclear fusion, PDE control problems, and electrical circuits [17, 18, 22]. Their applications can also be seen in stochastic control problems [15] and in parameter-varying linear systems [5]. Moreover, nonlinear systems can be approximated as bilinear systems via Carleman bilinearization [13, 20].

Many model reduction techniques for linear systems have been extended to bilinear systems with $E = I$ or E being invertible. For instance, Gramian-based model reduction techniques such as balanced truncation have been extended for bilinear systems [7] and interpolation-based model reduction techniques also have been successfully extended from the linear case to the bilinear case; see, e.g., [2, 9, 19], where interpolation of the leading k subsystems is considered. In [24], the Gramian-based Wilson conditions for $\mathcal{H}_2$ optimality were extended from linear systems [23] to bilinear systems.

Later, the analogue problem of determining an $\mathcal{H}_2$ optimal reduced-order system for bilinear systems was considered in [6], where the first-order necessary conditions for $\mathcal{H}_2$ optimality were derived by taking derivatives of the $\mathcal{H}_2$-norm of the error system with respect to the matrix entries of the realization of the reduced-order system. Based on these conditions, a bilinear iterative rational Krylov algorithm (*B-IRKA*) was proposed which on convergence leads to a locally $\mathcal{H}_2$ optimal reduced-order system. Moreover, recently, a new framework of interpolation for bilinear systems, the so-called multipoint interpolation, was considered which interpolates the whole underlying Volterra series at pre-defined frequency points [11] and therein also, the first-order necessary conditions for $\mathcal{H}_2$ optimality in terms of pole-residue formulation were proposed. It is also shown that the reduced-order system, satisfying these $\mathcal{H}_2$ optimality conditions in the pole-residues form also satisfies the optimality conditions derived in [6].

As has been noted, many model reduction techniques for linear systems have been extended to bilinear systems with $E = I$. But still, there are ample challenges when it comes to model reduction of bilinear descriptor systems with singular E, and it is necessary to study this case due to its omnipresence in applications [16]. In this paper, we focus on interpolatory model reduction techniques for bilinear descriptor systems with singular matrix E. The interpolation conditions for bilinear systems with $E = I$ can be readily extended to singular E by just replacing I by E. However, it was shown in [14] that directly extending the interpolation conditions for linear ODEs to linear DAEs leads to an unbounded error in the $\mathcal{H}_2$-norm due to the mismatch of the polynomial part of the system. This observation immediately holds for bilinear descriptor systems as well. As a consequence, we need to pay special attention to the polynomial part of the bilinear system along with interpolation.

Model reduction for a special family of bilinear descriptor systems, whose subsystems have constant polynomial parts, was recently considered in [12]. There, it was shown how to achieve the interpolation of the leading k subsystems together with retaining their polynomial parts. In contrast to this, in this paper we focus on extending the multipoint Volterra series interpolation to a similar special family of bilinear descriptor systems while paying attention to the polynomial part of the system. Secondly, we investigate the first-order necessary conditions for $\mathcal{H}_2$ optimality for a special family of bilinear descriptor systems and propose an iterative scheme to obtain an optimal reduced-order system.

The structure of the rest of the paper is as follows. We begin with giving a short overview on the multipoint interpolation framework for bilinear ODEs and visit the first-order necessary conditions for $\mathcal{H}_2$ optimality in Section 2. In Section 3, we show how to achieve the multipoint interpolation of the underlying Volterra series of bilinear descriptor systems along with retaining the constant polynomial part of each subsystem. Then, in Section 4, we extend the first-order necessary conditions for $\mathcal{H}_2$ optimality to bilinear descriptor systems and propose an iterative algorithm, the so-called *B-IRKA* for bilinear descriptor systems which on convergence gives rise to a locally $\mathcal{H}_2$ optimal reduced-order system. Finally in Section 5, we demonstrate the efficiency of the proposed methodology via several examples.

2 Multipoint Interpolation of the Volterra Series for Bilinear ODE Systems

In this section, we first briefly overview the multipoint interpolation of the Volterra series and the first-order necessary conditions for $\mathcal{H}_2$ optimality for bilinear ODE systems. For simplicity, we begin with considering a single-input single-output (SISO) bilinear system, i.e.,

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + Nx(t)u(t) + Bu(t), \\ y(t) = Cx(t), \quad x(0) = 0, \end{cases} \tag{2.1}$$

where the dimensions of A, B and C are the same as defined in (1.1) with $p = m = 1$, and $N \in \mathbb{R}^{n\times n}$. Assuming a stationary and causal bilinear system, the output $y(t)$ can be described by a nonlinear mapping of the input $u(t)$:

$$y(t) = \sum_{k=1}^{\infty} \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_k} g_k(t_1, t_2, \ldots, t_k) u(t - t_1 - t_2 \ldots - t_k) \cdots u(t - t_k) dt_1 \cdots dt_k,$$

where g_k is the regular Volterra kernel, whose corresponding multivariate transfer function can be given by

$$G_k(s_1, s_2, \ldots, s_k) = C(s_k I - A)^{-1} N \cdots (s_2 I - A)^{-1} N (s_1 I - A)^{-1} B.$$

The transfer function $G_k(s_1, s_2, \ldots, s_k)$ is also called the *kth* order multivariate transfer function associated with the bilinear system. Analogous to the linear case, the multivariate transfer function can be written in the pole-residue formulation which is given by the following proposition.

Proposition 2.1. *[11] Consider the multivariate transfer function* $G_k(s_1, s_2, \ldots, s_k) = C(s_k I - A)^{-1} N \cdots (s_2 I - A)^{-1} N (s_1 I - A)^{-1} B$ *and let* $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subset \mathbb{C}$ *be the* n *distinct zeros of* $det(sI - A)$*. Then the multivariate transfer function can also be written in the pole-residues form as follows:*

$$G_k(s_1, s_2, \ldots, s_k) = \sum_{l_1=1}^{n} \sum_{l_2=1}^{n} \cdots \sum_{l_k=1}^{n} \frac{\phi_{l_1,\ldots,l_k}}{\prod_{i=1}^{k} (s_i - \lambda_{l_i})},$$

where

$$\phi_{l_1,\ldots,l_k} = \lim_{s_k \to \lambda_{l_k}} (s_k - \lambda_{l_k}) \lim_{s_{k-1} \to \lambda_{l_{k-1}}} (s_{k-1} - \lambda_{l_{k-1}}) \cdots \lim_{s_1 \to \lambda_{l_1}} (s_1 - \lambda_{l_1}) G_k(s_1, \ldots, l_k). \tag{2.2}$$

Interpolatory model reduction techniques for bilinear systems have been studied widely in the literature; see, e.g., [2, 9, 19], where the leading k subsystems of the reduced-order system interpolate the corresponding original subsystem. However, recently in [11], multipoint interpolation for the whole Volterra series was considered at selected frequency points. We now outline the multipoint interpolation of the Volterra series problem statement for the bilinear system (2.1).

Consider two sets of interpolation points $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, for $j = 1, \ldots, r$, along with matrices $U, S \in \mathbb{C}^{r\times r}$, and define the weighted Volterra series

$$\zeta_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \sum_{l_2=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1,\ldots,l_{k-1},j} G_k(\sigma_{l_1}, \sigma_{l_2}, \ldots, \sigma_j) \tag{2.3}$$

and

$$\varphi_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \sum_{l_2=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \vartheta_{l_1,\ldots,l_{k-1},j} G_k(\mu_j, \mu_{l_1}, \ldots, \mu_{l_{k-1}}), \tag{2.4}$$

where $\eta_{l_1,\ldots,l_{k-1},j}$ and $\vartheta_{l_1,\ldots,l_{k-1},j}$ are the weights associated to each subsystem in the Volterra series, and are defined in terms of the elements of the matrices U and S as follows:

$$\begin{aligned} \eta_{l_1,\ldots,l_{k-1},j} &= u_{j,l_{k-1}} u_{l_{k-1},l_{k-2}} \cdots u_{l_2,l_1} \quad \text{for} \quad k \geq 2 \quad \text{and} \quad \eta_{l_1} = 1, \\ \vartheta_{l_1,\ldots,l_{k-1},j} &= s_{j,l_{k-1}} s_{l_{k-1},l_{k-2}} \cdots s_{l_2,l_1} \quad \text{for} \quad k \geq 2 \quad \text{and} \quad \vartheta_{l_1} = 1. \end{aligned} \tag{2.5}$$

The goal of the new interpolation framework is to construct a reduced-order system of dimension r:

$$\hat{\Sigma} : \begin{cases} \dot{\hat{x}}_r(t) = \hat{A}\hat{x}(t) + \hat{N}\hat{x}(t)u(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}\hat{x}(t), \quad \hat{x}(0) = 0, \end{cases} \tag{2.6}$$

where $\hat{A}, \hat{N} \in \mathbb{R}^{r\times r}$ and $\hat{B}, \hat{C}^T \in \mathbb{R}^r$, such that the following interpolation conditions are satisfied for each $j = 1, \ldots, r$:

$$\zeta_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \sum_{l_2=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1,\ldots,l_{k-1},j} \hat{G}_k(\sigma_{l_1}, \sigma_{l_2}, \ldots, \sigma_j) \tag{2.7}$$

and

$$\varphi_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \sum_{l_2=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \vartheta_{l_1,\ldots,l_{k-1},j} \hat{G}_k(\mu_j, \mu_{l_1}, \ldots, \mu_{l_{k-1}}), \tag{2.8}$$

where $\hat{G}_k(\mu_{l_1}, \ldots, \mu_k)$ is the kth order multivariate transfer function associated with the reduced-order bilinear system (2.6). Similar to the linear case, the reduced-order system matrices are constructed via projection matrices V and W, assuming $W^T V$ being invertible, as follows:

$$\begin{aligned} \hat{A} &= (W^T V)^{-1} W^T A V, & \hat{N} &= (W^T V)^{-1} W^T N V, \\ \hat{B} &= (W^T V)^{-1} W^T B, & \hat{C} &= CV. \end{aligned} \tag{2.9}$$

Then, the problem of identifying these projection matrices is considered in [11] which provides the reduced-order system such that the interpolation conditions are satisfied. The following theorem suggests the choice of the projection matrices.

Theorem 2.2. *[11] Consider a SISO bilinear system $\Sigma := (A, N, B, C)$ of dimension n and the interpolation points $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, $j = 1, \ldots, r$, along with matrices $U, S \in \mathbb{C}^{r\times r}$. Let the projection matrices V and W be the solutions of the following Sylvester equations*

$$V\Omega - AV - NVU^T = Be^T \tag{2.10}$$

and

$$W\Xi - A^TW - N^TWS^T = C^Te^T, \tag{2.11}$$

where $\Omega = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, $\Xi = \operatorname{diag}(\mu_1, \ldots, \mu_r)$, and e is the vector of ones in $\mathbb{R}^r$. Assume $W^TV \in \mathbb{R}^{r\times r}$ to be invertible and that the reduced-order system $\hat{\Sigma} := \{\hat{A}, \hat{N}, \hat{B}, \hat{C}\}$ of order r is computed using the projection matrices V and W as shown in (2.9). Then, the interpolation conditions (2.7) and (2.8) are fulfilled.

Furthermore, the $\mathcal{H}_2$-norm of the error system can be given in terms of the weighted sum of the multivariate transfer functions evaluated at all possible combinations of poles of the original and reduced-order system, see, [11]. Analogous to the linear case, the error in the $\mathcal{H}_2$-norm of the error system, due to the mismatch at the reduced-order system singularities, is eliminated, leading to the following first-order necessary conditions for optimality:

$$\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\hat{\phi}_{l_1,\ldots,l_k}\left(G_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k}) - \hat{G}_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k})\right) = 0 \tag{2.12}$$

and

$$\begin{aligned}&\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\hat{\phi}_{l_1,\ldots,l_k}\left(\sum_{j=1}^{k}\frac{\partial}{\partial s_j}G_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k})\right)\\ &=\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\hat{\phi}_{l_1,\ldots,l_{l_k}}\left(\sum_{j=1}^{k}\frac{\partial}{\partial s_j}\hat{G}_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k})\right),\end{aligned} \tag{2.13}$$

where the $\hat{\lambda}_i$'s are the zeros of $\det(s\hat{I} - \hat{A})$ and $\hat{\phi}_{l_1,\ldots,l_k}$ are the residues of the kth order transfer functions $\hat{G}_k(s_1, s_2, \ldots, s_k)$, as defined in (2.2). Here, the operator $\frac{\partial}{\partial s_j}G_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k})$ denotes the partial derivative of $G_k(s_1, \ldots, s_k)$ with respect to s_j, evaluated at $(s_1, \ldots, s_k) = (-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k})$.

It is also shown in [11] that the first-order necessary conditions for $\mathcal{H}_2$ optimality in terms of the pole-residues form are satisfied, if the projection matrices V and W are computed by setting the interpolation points as mirror image of the poles of the reduced-order system across the imaginary axis, i.e., $\Omega = \Xi = -\Theta$ in (2.10) and (2.11), respectively, where $\Theta = R^{-1}\hat{A}R$; the matrices U and S are given by the bilinear term $\hat{N}$ as $U = R^{-1}\hat{N}R$ and $S = R^T\hat{N}^TR^{-T}$; and the vector e in (2.10) and (2.11) is replaced with $R^{-1}\hat{B}$ and $\hat{C}R$, respectively. For details, we refer to [10, 11].

Remark 2.3. *The multipoint interpolation of the underlying Volterra series can be extended to bilinear descriptor systems by replacing I by E. This yields a reduced-order system which satisfies the interpolation conditions. However, directly extending the interpolation conditions to descriptor systems without any modifications, leads to a poor reduced-order systems with the $\mathcal{H}_2$-norm error blowing up, occuring due to the unmatched polynomial part of the system. This statement is based on the analysis in [14] for linear descriptor systems.*

Motivated by the work done in [14], we pay special attention to the polynomial part of the bilinear descriptor system in this paper along with interpolation. In the following section we show how to achieve multipoint interpolation of the underlying Volterra series along with matching the polynomial part of the multivariate transfer function of each subsystem.

3 Multipoint Interpolation of the Volterra Series for Bilinear Descriptor Systems

Here, we deal with a special family of bilinear descriptor systems (DAEs). The considered family consists of those systems with all associated multivariate transfer functions having constant polynomial part. For simplicity, we begin with a single-input single-output bilinear descriptor system, i.e.,

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Nx(t)u(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned} \tag{3.1}$$

where the dimensions of E, A, B and C are as defined in (1.1) with $p = m = 1$ and $N_1 = N \in \mathbb{R}^{n\times n}$. Similar to bilinear ODEs, the kth order multivariate transfer function of the bilinear descriptor systems (3.1) can be given by

$$H_k(s_1, s_2, \ldots, s_k) = C(s_kE - A)^{-1}N \cdots (s_2E - A)^{-1}N(s_1E - A)^{-1}B. \tag{3.2}$$

Our goal is to extend interpolation based model reduction techniques to bilinear DAEs. Therefore, we first intend to determine an explicit expression for the polynomial part of the multivariate transfer function $H_k(s_1, s_2, \ldots, s_k)$, by assuming a special structure of the matrices E and A in (3.1) as follows:

$$E = \begin{bmatrix} E_{11} & E_{12} \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{3.3}$$

where A_{22} and $E_{11} - E_{12}A_{22}^{-1}A_{21}$ are invertible. This means that the matrix pencil (A, E) has nilpotency index 1. It was shown in [12] that the kth order multivariate transfer function of the bilinear system, having the structure of the matrices as shown in (3.3), has a constant polynomial part which can be determined by the following lemma.

Lemma 3.1. *[12] Let $H_k(s_1, s_2, \dots, s_k) = C(s_kE - A)^{-1}N \dots (s_2E - A)^{-1}N(s_1E - A)^{-1}B$ be the Laplace transform of the kth order subsystem. Then, the polynomial part of $H_k(s_1, s_2, \dots, s_k)$ is constant and can be given as*

$$D_k = C(MN)^{k-1}MB,$$

where $M = \begin{bmatrix} 0 & E_A^{-1}E_{12}A_{22}^{-1} \\ 0 & -A_{22}^{-1}\left(I + A_{21}E_A^{-1}E_{12}A_{22}^{-1}\right) \end{bmatrix}$ *and* $E_A = E_{11} - E_{12}A_{22}^{-1}A_{21}$.

Now, we discuss the interpolation based model reduction techniques for bilinear descriptor systems that also retain the explicitly computed constant polynomial part of the subsystems together with interpolation. Recently, the problem of the interpolation of the leading k subsystems of bilinear descriptor systems while retaining their polynomial parts was considered in [12], i.e.,

$$H_i(\sigma_1, \sigma_2, \dots, \sigma_i) = \hat{H}_i(\sigma_1, \sigma_2, \dots, \sigma_i), \quad \text{for} \quad i = 1, \dots, k, \tag{3.4}$$

where $\{\sigma_i\} \subset \mathbb{C}$ are the interpolation points and $\hat{H}_k(s_1, s_2, \dots, s_k)$ is the regular *kth* order multivariate transfer function of the reduced-order system, which is of the form

$$\hat{H}_k(s_1, s_2, \dots, s_k) = \hat{C}(s_k\hat{E} - \hat{A})^{-1}\hat{N} \cdots (s_2\hat{E} - \hat{A})^{-1}\hat{N}(s_1\hat{E} - \hat{A})^{-1}\hat{B} + D_k \tag{3.5}$$

with invertible $\hat{E}$. It can be easily seen that the polynomial parts of $H_k(s_1, s_2, \dots, s_k)$ and $\hat{H}_k(s_1, s_2, \dots, s_k)$ are equal to D_k.

In contrast to this, we focus on interpolating the underlying Volterra series and at the same retaining the polynomial part of each subsystem. Therefore, we revisit the following multipoint Volterra interpolation problem. We consider two sets of interpolation points $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, $j = 1, 2, \dots, r$, along with matrices $U, S \in \mathbb{C}^{r \times r}$, and define the weighted Volterra series as follows:

$$\nu_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1, l_2, \dots, l_{k-1}, j} H_k(\sigma_{l_1}, \sigma_{l_2}, \dots, \sigma_j) \tag{3.6}$$

and

$$\gamma_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \vartheta_{l_1, l_2, \dots, l_{k-1}, j} H_k(\mu_j, \mu_{l_1}, \dots, \mu_{l_{k-1}}), \tag{3.7}$$

where the weights $\eta_{l_1, l_2, \dots, l_{k-1}, j}$ are defined in (2.5) in terms of the elements of the matrix U and similarly for $\vartheta_{l_1, l_2, \dots, l_{k-1}, j}$. It is assumed that ν_j and γ_j converge for each $j = 1, 2, \dots, r$. The goal of the multipoint Volterra series interpolation is to determine a reduced-order system, with its *kth* order multivariate transfer function being of the form (3.5), so that the following are satisfied for each $j = 1, \dots, r$:

$$\nu_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1, l_2, \dots, l_{k-1}, j} \hat{H}_k(\sigma_{l_1}, \sigma_{l_2}, \dots, \sigma_j) \tag{3.8}$$

and

$$\gamma_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \vartheta_{l_1,l_2,\ldots,l_{k-1},j}\hat{H}_k(\mu_j,\mu_{l_1},\ldots,\mu_{l_{k-1}}). \tag{3.9}$$

As a first step in this direction, we establish the relation between the weighted Volterra series and the generalized Sylvester equation for the bilinear descriptor systems in the following lemma, similar to the case of bilinear ODEs in [11, Lemma 3.1].

Lemma 3.2. *Consider $\Sigma := \{E, A, N, B, C\}$ to be a SISO bilinear descriptor system and let $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, $j = 1, \ldots, r$, be two sets of interpolation points. Given matrices $U, S \in \mathbb{C}^{r\times r}$, and assume the following series*

$$v_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \eta_{l_1,l_2,\ldots,l_{k-1},j}(\sigma_j E - A)^{-1}N\cdots(\sigma_{l_2}E - A)^{-1}N(\sigma_{l_1}E - A)^{-1}B \tag{3.10}$$

and

$$w_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \vartheta_{l_1,l_2,\ldots,l_{k-1},j}(\sigma_j E - A)^{-T}N^T\cdots(\sigma_{l_2}E - A)^{-T}N^T(\sigma_{l_1}E - A)^{-T}C^T \tag{3.11}$$

converge for each $j = 1, 2, \ldots, r$. Then, the matrices V and W, whose jth columns are v_j and w_j, respectively, solve the following generalized Sylvester equations

$$EV\Omega - AV - NVU^T = Be^T \tag{3.12}$$

and

$$E^TW\Xi - A^TW - N^TWS^T = C^Te^T, \tag{3.13}$$

respectively, where $\Omega = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ and $\Xi = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_r)$.

The proof of the above lemma is analogous to [11, Lemma 3.1] where $E = I$ was considered. Nevertheless, it can be easily extended to $E \neq I$ in a similar fashion. Therefore, for brevity of the paper, we skip the proof.

Next, in the following theorem, we discuss the construction of a reduced-order system with required modifications so that (3.8) and (3.9) can be satisfied.

Theorem 3.3. *Consider a SISO bilinear descriptor system* (3.1) *of order n. Assume for some $r < n$ that two sets of interpolation points $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, $j = 1, 2, \ldots, r$, and matrices $U, S \in \mathbb{C}^{r\times r}$, such that $\Lambda(U) \cap \Lambda(S) = \emptyset$, where $\Lambda()$ denotes the spectrum of a matrix. Let the matrices V and W be the solutions of* (3.12) *and* (3.13)*, respectively, and L_A, L_N, L_B and L_c be the solutions to*

$$L_AV + L_NVU^T + L_Be^T = 0, \tag{3.14a}$$

$$L_A^TW + L_N^TWS^T + L_c^Te^T = 0, \tag{3.14b}$$

$$W^TL_B + [\alpha_1, \alpha_2, \ldots, \alpha_r]^T = 0, \tag{3.14c}$$

$$L_CV + [\beta_1, \beta_2, \ldots, \beta_r] = 0, \tag{3.14d}$$

where

$$\alpha_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \vartheta_{l_1,l_2,\ldots,l_{k-1},j} D_k \tag{3.15}$$

and

$$\beta_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1,l_2,\ldots,l_{k-1},j} D_k. \tag{3.16}$$

If the matrices of the reduced-order system are computed as

$$\begin{aligned} &\hat{E} = W^T E V, && \hat{A} = W^T (A + L_A) V, \quad \hat{N} = W^T (N + L_N) V, \\ &\hat{B} = W^T (B + L_B), && \hat{C} = (C + L_C) V, \end{aligned} \tag{3.17}$$

then, the interpolation conditions (3.8) *and* (3.9) *are satisfied for each* $j = 1, \ldots, r$. *Furthermore, if* $\hat{E}$ *is invertible, then the polynomial part of each subsystem is also matched.*

Proof. We begin with the Sylvester equation, determining the projection matrix V

$$EV\Omega - AV - NVU^T - Be^T = 0. \tag{3.18}$$

Subtracting (3.18) and (3.14a) yields

$$EV\Omega - (A + L_A)V - (N + L_N)VU^T - (B + L_B)e^T = 0.$$

Premultiplying the above equation by W^T, we obtain

$$W^T \left(EV\Omega - (A + L_A)V - NVU^T - (B + L_B)e^T \right) = 0.$$

This implies

$$\hat{E}\Omega - \hat{A} - \hat{N}U^T - \hat{B}e^T = 0.$$

From the above equation, it follows that $\Psi = \hat{I}$ solves the following projected Sylvester equation:

$$\hat{E}\Psi\Omega - \hat{A}\Psi - \hat{N}\Psi U^T - \hat{B}e^T = 0, \tag{3.19}$$

where $\hat{I}$ is an identity matrix of appropriate dimension. The above projected Sylvester equation has a structure similar to the one in Lemma 3.2. So, using Lemma 3.2, the jth column of Ψ, denoted by ψ_j, can be given as

$$\psi_j = \sum_{k=1}^{\infty} \sum_{l_1=1}^{r} \cdots \sum_{l_{k-1}=1}^{r} \eta_{l_1,l_2,\ldots,l_{k-1},j} (\sigma_j \hat{E} - \hat{A})^{-1} \hat{N} \cdots (\sigma_{l_2} \hat{E} - \hat{A})^{-1} \hat{N} (\sigma_{l_1} \hat{E} - \hat{A})^{-1} \hat{B}. \tag{3.20}$$

Now, we multiply ψ_j by $\hat{C}$ to obtain

$$\hat{C}\psi_j = (C + L_C)V\psi_j = CV\psi_j + L_C V\psi_j. \tag{3.21}$$

The vector ψ_j is the jth column of the identity matrix. Therefore, $V\psi_j$ gives the jth column of the matrix V, given in (3.10) and multiplication with C gives

$$CV\psi_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\ldots,l_{k-1},j}H_k(\sigma_{l_1},\sigma_{l_2},\ldots,\sigma_j) = \nu_j. \tag{3.22}$$

By (3.14d), we get

$$L_cV\psi_j = -\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\ldots,l_{k-1},j}D_k. \tag{3.23}$$

Finally, we substitute (3.22), (3.23) and the expression for ψ_j from (3.20) in (3.21) to have

$$\begin{aligned}\nu_j &= \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\ldots,l_{k-1},j}\hat{C}(\sigma_j\hat{E}-\hat{A})^{-1}\hat{N}\cdots(\sigma_{l_2}\hat{E}-\hat{A})^{-1}\hat{N}(\sigma_{l_1}\hat{E}-\hat{A})^{-1}\hat{B}\\ &\quad+\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\ldots,l_{k-1},j}D_k\\ &= \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\ldots,l_{k-1},j}\hat{H}_k(\sigma_{l_1},\sigma_{l_2},\ldots,\sigma_j).\end{aligned}$$

Using a similar argument, we can prove

$$\gamma_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\vartheta_{l_1,l_2,\ldots,l_{k-1},j}\hat{H}_k(\mu_j,\mu_{l_1},\ldots,\mu_{l_{k-1}}).$$

Since, we have assumed the form of the kth order multivariate transfer function of the reduced-order system as shown in (3.5) and $\hat{E}$ being invertible, this means that the polynomial parts of each subsystem of the original and reduced-order system are equal to D_k. This concludes the proof. □

Remark 3.4. *Theorem 3.3 extends the interpolation for linear systems with $D \neq D_r$ [4, Thm. 3] to bilinear systems.*

Remark 3.5. *In Theorem 3.3, it is assumed that the matrices U and S do not have any common eigenvalue in order to have simultaneous solutions of the set of equations* (3.14a)–(3.14d) *for the matrices L_A, L_N, L_B and L_C. If the matrices U and S have common eigenvalues, then this leads to numerical issues which we discuss in the next section.*

Theorem 3.3 shows how to choose the projection matrices and to obtain a reduced-order system with the required modifications which not only interpolates the underlying Volterra series but also retains the polynomial part of each subsystem. Meanwhile, we also like to highlight an important aspect that the reduced-order system matrices obtained from Theorem 3.3 are not obtained via projection of the original system matrices (3.1). They are rather obtained via projection of another bilinear system (intermediate bilinear system) of order n whose kth order multivariate transfer function is given by

$$\tilde{H}(s_1, s_2, \dots, s_k) = \tilde{C}(s_k\tilde{E} - \tilde{A})^{-1}\tilde{N} \cdots (s_2\tilde{E} - \tilde{A})^{-1}\tilde{N}(s_1\tilde{E} - \tilde{A})^{-1}\tilde{B} + D_k, \tag{3.24}$$

where

$$\begin{aligned} \tilde{E} &= E, & \tilde{A} &= A + L_A, & \tilde{N} &= N + L_N, \\ \tilde{B} &= B + L_B, & \tilde{C} &= C + L_C. \end{aligned} \tag{3.25}$$

Interestingly, we project the intermediate bilinear system using the projection matrices V and W which depend on the original bilinear system matrices, as opposed to the intermediate bilinear system matrices. So next, to resolve this discrepancy, we show the formulation of the reduced-order system, obtained in Theorem 3.3, in a standard projection framework using the intermediate bilinear system. We reveal that the projection matrices obtained using the original and intermediate bilinear system matrices are exactly the same.

Proposition 3.6. *For some $r < n$, we consider two sets of interpolation points $\sigma_j \in \mathbb{C}$ and $\mu_j \in \mathbb{C}$, $j = 1, \dots, r$, and matrices $U, S \in \mathbb{C}^{r\times r}$ such that $\Lambda(U) \cap \Lambda(S) = \emptyset$. Let the matrices V and W be the solutions of* (3.12) *and* (3.13)*, respectively, and let the projection matrices $\tilde{V}$ and $\tilde{W}$ be the solutions to*

$$\tilde{E}\tilde{V}\Omega - \tilde{A}\tilde{V} - \tilde{N}\tilde{V}U^T = \tilde{B}e^T \tag{3.26}$$

and

$$\tilde{E}^T\tilde{W}\Omega - \tilde{A}\tilde{W} - \tilde{N}^T\tilde{W}S^T = \tilde{C}^Te^T, \tag{3.27}$$

respectively. Then, $\tilde{V} = V$ and $\tilde{W} = W$ also solve (3.26) *and* (3.27)*, respectively.*

Proof. We begin with proving that the matrix V also satisfies (3.26). We start with

$$\begin{aligned} &\tilde{E}V\Omega - \tilde{A}V - \tilde{N}VU^T \\ &\quad = EV\Omega - AV - L_AV - NVU^T - L_NVU^T \quad \text{(substituting for } \tilde{A} \text{ and } \tilde{N} \text{ from (3.25))} \\ &\quad = (EV\Omega - AV - NVU^T) - (L_AV + L_NVU^T) \end{aligned}$$

From (3.12), $EV\Omega - AV - NVU^T = Be^T$ and using the relation between L_A, L_N and L_B from (3.14a), we get

$$\begin{aligned} \tilde{E}V\Omega - \tilde{A}V - \tilde{N}VU^T &= Be^T + L_Be^T \\ &= (B + L_B)e^T = \tilde{B}e^T. \end{aligned}$$

An analogous argument can be given for (3.27) as well. This proves the assertion. □

Based on this investigation, we propose the following corollary.

Corollary 3.7. *The reduced-order system, determined in Theorem 3.3, coincides with the reduced system obtained from the intermediate bilinear system, whose kth order multivariate transfer function is given in* (3.24)*, via the projection subspaces* $\tilde{V}$ *and* $\tilde{W}$ *in a standard projection framework.*

4 $\mathcal{H}_2$-Model Order Reduction for Bilinear Descriptor Systems

So far, we have shown how to determine a reduced-order system with the appropriate modifications such that the multipoint interpolation of the underlying Volterra series can be achieved together with retaining the polynomial part of each subsystem. In this section, we discuss the first-order necessary conditions for $\mathcal{H}_2$ optimality of the special structured bilinear descriptor systems. The first-oder necessary conditions, in terms of the pole-residues of the multivariate transfer functions, for bilinear ODEs were derived in [11] by minimizing the error in the $\mathcal{H}_2$-norm of the error system. In this paper, we consider the analogous first-order necessary conditions for optimality for bilinear descriptor systems which are as follows:

$$\begin{aligned}\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}&\hat{\phi}_{l_1,l_2,\dots,l_{k-1},j}H_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k})\\&=\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\hat{\phi}_{l_1,l_2,\dots,l_{k-1},j}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k})\end{aligned}\tag{4.1}$$

and

$$\begin{aligned}\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}&\hat{\phi}_{l_1,\dots,l_k}\left(\sum_{j=1}^{k}\frac{\partial}{\partial s_j}H_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k})\right)\\&=\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\hat{\phi}_{l_1,\dots,l_k}\left(\sum_{j=1}^{k}\frac{\partial}{\partial s_j}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k})\right),\end{aligned}\tag{4.2}$$

where $\hat{\phi}_{l_1,\dots,l_k}$ and $\hat{\lambda}_{l_i}$ are the residues and poles, respectively, of the transfer functions $\hat{H}_k(s_1,s_2,\dots,s_k)$. In this regard, we first establish the connection between the multipoint interpolation of the Volterra series interpolation conditions and the pole-residues of the *kth* order multivariate transfer function of the reduced-order system.

Lemma 4.1. *Let* $H_k(s_1,s_2,\dots,s_k)$ *and* $\hat{H}_k(s_1,s_2,\dots,s_k)$ *be the kth order multivariate transfer functions of the original and reduced-order systems as shown in* (3.2) *and in* (3.5)*, respectively. Decompose* $Y\hat{A}Z=\Omega=\mathrm{diag}(\hat{\lambda}_1,\hat{\lambda}_2,\dots,\hat{\lambda}_r)$ *and* $Y\hat{E}Z=\hat{I}$*, where* $\{\hat{\lambda}_1,\hat{\lambda}_2,\dots,\hat{\lambda}_r\}$ *are the eigenvalues matrix pencil* (A,E) *and the columns of* $Z=[z_1,z_2,\dots,z_r]$ *and* $Y=[y_1,y_2,\dots,y_r]$ *are the right and left eigenvectors, respectively.*

Moreover, define $\mathcal{B} = Y\hat{B}$, $\mathcal{N} = Y\hat{N}Z$ and $\mathcal{C} = \hat{C}Z$, and let $\hat{\phi}_{l_1,l_2,\ldots,l_k}$ be the residues corresponding to the kth order multivariate transfer function $\hat{H}_k(s_1,\ldots,s_k)$. Assume that the projection matrices V and W solve

$$EV(-\Omega) - AV - NV\mathcal{N}^T = B\mathcal{B}^T, \tag{4.3}$$

$$E^TW(-\Omega) - A^TW - N^TW\mathcal{N} = C^T\mathcal{C}, \tag{4.4}$$

respectively. Then,

$$\mathcal{C}\,(CV)^T = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\hat{\phi}_{l_1,l_2,\ldots,l_k}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k}). \tag{4.5}$$

Proof. We begin by comparing (4.3) and (3.12) which readily shows that these two equations are equivalent after setting

$$U = \mathcal{N}, \quad e = \mathcal{B} \quad \text{and} \quad \sigma_j = -\hat{\lambda}_j, \quad j = 1,\ldots,r.$$

By applying Lemma 3.2, we can write the jth column of V, v_j, as follows:

$$\begin{aligned} v_j = &\sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,\ldots,l_{k-1},j}\mathcal{B}_{l_1}(\hat{\lambda}_jE-A)^{-1}N\cdots(\hat{\lambda}_{l_2}E-A)^{-1}N(\hat{\lambda}_{l_1}E-A)^{-1}B \\ &+ \mathcal{B}_j(\hat{\lambda}_jE-A)^{-1}B, \end{aligned} \tag{4.6}$$

where $\eta_{l_1,\ldots,l_{k-1},j} = \mathcal{N}(j,l_{k-1})\mathcal{N}(l_{k-1},l_{k-2})\cdots\mathcal{N}(l_2,l_1)$ for $k \geq 2$ by the definition of $\eta_{l_1,\ldots,l_{k-1}}$ in (2.5), and $\mathcal{B}_i$ is the ith element of $\mathcal{B}$. Multiplying (4.6) by C yields

$$Cv_j = \sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,\ldots,l_{k-1},j}\mathcal{B}_{l_1}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_j) + \mathcal{B}_jH_1(-\hat{\lambda}_j). \tag{4.7}$$

Hence,

$$(CV)^T = \begin{bmatrix} \sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,\ldots,l_{k-1},1}\mathcal{B}_{l_1}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_1) + \mathcal{B}_1H_1(-\lambda_1) \\ \sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,\ldots,l_{k-1},2}\mathcal{B}_{l_1}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_2) + \mathcal{B}_2H_1(-\lambda_2) \\ \vdots \\ \sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,\ldots,l_{k-1},r}\mathcal{B}_{l_1}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_r) + \mathcal{B}_rH_1(-\lambda_r) \end{bmatrix}. \tag{4.8}$$

Next, we premultiply the above equation by $\mathcal{C} = [\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_r]$, where $\mathcal{C}_i$ is the ith element of $\mathcal{C}$. This yields

$$\mathcal{C}(CV)^T = \sum_{k=2}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_k=1}^{r}\eta_{l_1,\ldots,l_{k-1},l_k}\mathcal{C}_{l_k}\mathcal{B}_{l_1}H_k(-\hat{\lambda}_{l_1},\ldots,-\hat{\lambda}_{l_k}) + \sum_{l_k=1}^{r}\mathcal{C}_{l_k}\mathcal{B}_{l_k}H_1(-\lambda_{l_k}). \tag{4.9}$$

Now, we recall the expression for the residues $\hat{\phi}_{l_1,\dots,l_k}$ of the kth order multivariate transfer function of the reduced-order system which are given as

$$\begin{aligned}\hat{\phi}_{l_k} &= \mathcal{C}_{l_k}\mathcal{B}_{l_k},\\ \hat{\phi}_{l_1,\dots,l_k} &= \mathcal{C}_{l_k}\eta_{l_1,\dots,l_{k-1},l_k}\mathcal{B}_{l_1}, \quad \text{for} \quad k \geq 2.\end{aligned}$$

Lastly, we substitute the above relation in (4.9) which leads to the desired result. □

Our next task is to obtain a reduced-order system which satisfies the necessary conditions for optimality (4.1) and (4.2). The following theorem reveals the choice of a reduced-order system ensuring the first-order necessary conditions for $\mathcal{H}_2$ optimality.

Theorem 4.2. *Let $H_k(s_1, s_2, \dots, s_k)$ and $\hat{H}_k(s_1, s_2, \dots, s_k)$ be the kth order multivariate transfer functions of the original and reduced-order bilinear systems, respectively, and assume the projection matrices V and W are given by* (4.3) *and* (4.4)*, respectively. Also, assume that L_A, L_N, L_B and L_C satisfy the following set of equations:*

$$L_A V + L_N V \mathcal{N}^T + L_B \mathcal{B}^T = 0, \tag{4.10a}$$
$$L_A^T W + L_N^T W \mathcal{N} + L_c^T \mathcal{C} = 0, \tag{4.10b}$$
$$W^T L_B + [\alpha_1, \alpha_2, \dots, \alpha_r]^T = 0, \tag{4.10c}$$
$$L_C V + [\beta_1, \beta_2, \dots, \beta_r] = 0, \tag{4.10d}$$

where

$$\alpha_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \vartheta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{C}_{l_1}D_k \tag{4.11}$$

and

$$\beta_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \eta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{B}_{l_1}D_k \tag{4.12}$$

with

$$\begin{aligned}\eta_{l_1,\dots,l_{k-1},j} &= \mathcal{N}(j, l_{k-1})\mathcal{N}(l_{k-1}, l_{k-2})\cdots\mathcal{N}(l_2, l_1) \quad \textit{for} \quad k \geq 2,\\ \vartheta_{l_1,\dots,l_{k-1},j} &= \mathcal{N}(l_{k-1}, j)\mathcal{N}(l_{k-2}, l_{k-1})\cdots\mathcal{N}(l_1, l_2) \quad \textit{for} \quad k \geq 2.\end{aligned}$$

If the reduced-order system matrices are computed as shown in (3.17)*, then the first-order necessary conditions for $\mathcal{H}_2$ optimality* (4.1) *and* (4.2) *are satisfied along with retaining the polynomial part of each subsystem.*

Proof. We begin by recalling Lemma 3.2 which provides us the formulation of the jth column of the identity matrix, ψ_j, see equation (3.20),

$$\psi_j = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r} \eta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{B}_{l_1}(\sigma_j\hat{E} - \hat{A})^{-1}\hat{N}\cdots(\sigma_{l_2}\hat{E} - \hat{A})^{-1}\hat{N}(\sigma_{l_1}\hat{E} - \hat{A})^{-1}\hat{B},$$

Now, we multiply the above equation by $\hat{C}$ to get

$$\hat{C}\Psi = (C + L_C)V = CV + L_C V. \tag{4.13}$$

Transposing (4.13) and premultiplying by $\mathcal{C}$ leads to

$$\mathcal{C}(\hat{C}\Psi)^T = \mathcal{C}(CV)^T + \mathcal{C}(L_C V)^T.$$

Next, we substitute $L_C V$ given in (4.10d) and employ (4.12) which on simplification yields

$$\mathcal{C}(CV)^T = \mathcal{C}\begin{bmatrix} \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{B}_{l_1}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_1)) \\ \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{B}_{l_1}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_2)) \\ \vdots \\ \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\eta_{l_1,l_2,\dots,l_{k-1},j}\mathcal{B}_{l_1}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_r)) \end{bmatrix}.$$

Using Lemma 4.1 and simple algebra gives us

$$\begin{aligned} &\sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\hat{\phi}_{l_1,l_2,\dots,l_{k-1},j}H_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k}) \\ &\qquad = \sum_{k=1}^{\infty}\sum_{l_1=1}^{r}\cdots\sum_{l_{k-1}=1}^{r}\hat{\phi}_{l_1,l_2,\dots,l_{k-1},j}\hat{H}_k(-\hat{\lambda}_{l_1},\dots,-\hat{\lambda}_{l_k}). \end{aligned} \tag{4.14}$$

The second necessary condition (4.2) can be easily obtained in a similar fashion as shown in [11, Thm. 4.2] by tracing the terms corresponding to $W(:,j)^T V(:,j)$, for $j = 1, 2, \dots, r$. □

Clearly, still the computation of the reduced-order system matrices involves the matrices L_A, L_N, L_B and L_c which are not readily available. In what follows we show how to compute the reduced-order system without explicitly computing these matrices and related computational issues.

Computational issues

Now, we discuss the computational issues regarding determining the reduced-order system. It is interesting to note that we do not need the matrices L_A, L_N, L_B and L_C explicitly, but we rather require expressions for $W^T L_A V, W^T L_N V, W^T L_B$ and $L_C V$ to determine the reduced-order system. The expressions for $W^T L_B$ and $L_C V$ are given in (4.10c) and (4.10d), respectively which are

$$\begin{aligned} W^T L_B &= -[\mathcal{C}^T D_1 + \mathcal{N}^T\mathcal{C}^T D_2 + (\mathcal{N}^T)^2\mathcal{C}^T D_3 + \cdots], \\ L_C V &= -[D_1\mathcal{B}^T + D_2\mathcal{B}^T\mathcal{N}^T + D_3\mathcal{B}^T(\mathcal{N}^T)^2 + \cdots]. \end{aligned}$$

In order to determine the expressions for $W^T L_A V$ and $W^T L_N V$, we premultiply (4.10a) and (4.10b) by W^T and V^T, respectively, and obtain

$$W^T L_A V + W^T L_N V \mathcal{N}^T + W^T L_B \mathcal{B}^T = 0, \tag{4.15}$$

$$V^T L_A^T W + V^T L_N^T W \mathcal{N} + V^T L_C^T \mathcal{C} = 0. \tag{4.16}$$

Now, we subtract (4.15) from the transpose of (4.16), leading to the following Sylvester equation in $W^T L_N V$:

$$\mathcal{N}^T (W^T L_N V^T) - (W^T L_N V)\mathcal{N}^T + \mathcal{C}^T L_C V - W^T L_B \mathcal{B}^T = 0. \tag{4.17}$$

In order to have a unique solution of the above Sylvester equation, the matrix $(\hat{I} \otimes \mathcal{N}^T - \mathcal{N} \otimes \hat{I})$ should be invertible. But, it is easy to see that the matrix contains zero eigenvalues. Therefore, the Sylvester equation (4.17) either does not have a unique solution or has no solution. However, if one assumes $D_k = 0$ for $k \geq 3$, then at-least one solution for $W^T L_N V$ can be easily computed which satisfies (4.17). In this scenario, the equation (4.17) boils down to

$$\mathcal{N}^T (W^T L_N V - \mathcal{C}^T D_2 \mathcal{B}^T) - (W^T L_N V - \mathcal{C}^T D_2 \mathcal{B}^T)\mathcal{N}^T = 0.$$

This implies $W^T L_N V = \mathcal{C}^T D_2 \mathcal{B}^T$ satisfies the above Sylvester equation, although it is not unique. The expression for $W^T L_A V$ can be simply computed by inserting the expressions for $W^T L_B$ and $W^T L_N V$ in (4.15).

Remark 4.3. *As we have noted above, the Sylvester equation* (4.17) *either does not have unique solution or even has no solution. However, it is possible to determine the solution if $D_k = 0 \ \forall \ k \geq 3$.*

In case of $D_k \neq 0$ for some $k \geq 3$, the equation (4.17)*, in general, does not have any solution. This implies that it is not possible to obtain a reduced-order system, satisfying necessary conditions for optimality. Nevertheless, here we set $W^T L_N V$ equal to $\mathcal{C}^T D_2 \mathcal{B}^T$ which often is a good choice as D_k generally decreases fast.*

Remark 4.4. *It is shown in [12] that if the bilinear term has the following structure:*

$$N = \begin{bmatrix} N_{11} & N_{12} \\ 0 & 0 \end{bmatrix},$$

then the higher order systems with $k \geq 2$, all have zero polynomial parts, i.e., $D_k = 0 \ \forall \ k \geq 2$.

Here now, we sketch an iterative algorithm based on our theoretical discussions for the special class of bilinear descriptor systems considered here.

Remark 4.5. *Algorithm 1 extends the algorithm proposed in [13] for bilinear descriptor system for which the polynomial part of each subsystem was assumed to be zero.*

Remark 4.6. *The expressions for $\mathcal{R}_B$ and $\mathcal{R}_C$ require the summation of the infinite series. However, D_i generally decreases fast, therefore one can consider only the leading terms which may approximate the infinite summation very well. In all practical applications we consider in the next section, $D_k = 0 \ \forall \ k \geq 2$.*

Algorithm 1 B-IRKA for bilinear descriptor systems with index-1 matrix pencil.

Input: E, A, N, B, C.
Make an initial guess of $\Omega, \mathcal{B}, \mathcal{N}$ and $\mathcal{C}$.
while no convergence **do**
  Solve for V and W
  $EV(-\Omega) + AV + NV\mathcal{N}^T + B\mathcal{B}^T = 0,$
  $E^T W(-\Omega) + A^T W + N^T W\mathcal{N} + C^T\mathcal{C} = 0.$
  Compute the expression for
  $$W^T L_B = -\sum_{k=1}^{\infty} (\mathcal{N}^T)^{k-1}\mathcal{C}^T D_k =: \mathcal{R}_B\ ,$$
  $$L_C V \ = -\sum_{k=1}^{\infty} D_k \mathcal{B}^T (\mathcal{N}^T)^{k-1} =: \mathcal{R}_C.$$
  Determine the expression for $W^T L_N V = \mathcal{C}^T D_2 \mathcal{B}^T =: \mathcal{R}_N$.
  Determine the expression for $W^T L_A V =: \mathcal{R}_A$,
  $\mathcal{R}_A = -\mathcal{R}_N\mathcal{N}^T - \mathcal{R}_B\mathcal{B}^T$.
  Compute the reduced-order system matrices:
  $\hat{E} = W^T EV, \quad \hat{A} = W^T AV + \mathcal{R}_A, \quad \hat{N} = W^T NV + \mathcal{R}_N,$
  $\hat{B} = W^T B + \mathcal{R}_B, \quad \hat{C} = CV + \mathcal{R}_C.$
  Determine Y and Z such that $Y\hat{A}Z = \Omega$, $Y\hat{E}Z = \hat{I}$.
  Compute $\mathcal{N} = Y\hat{N}Z$, $\mathcal{B} = Y\hat{B}$ and $\mathcal{C} = \hat{C}Z$.
end while
Output: $\hat{E}, \hat{A}, \hat{N}, \hat{B}, \hat{C}$.

Remark 4.7. *The application of Algorithm 1 is not only restricted to bilinear descriptor systems with index-1 matrix pencil as shown in* (3.3)*, but also can be applied to all bilinear descriptor systems, whose subsystems all have constant polynomial parts. For instance, all subsystems of the bilinear descriptor system with the following structure of matrices:*

$$E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix},$$
$$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \quad \textit{and} \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix},$$

have constant polynomial parts. Here, the matrix pencil (A, E) has a nilpotency index-2. Theoretically, Algorithm 1 can be employed to determine a $\mathcal{H}_2$ optimal reduced-order system. But numerically, we have experienced that as the nilpotency index of the matrix pencil (A, E) increases, the convergence of Algorithm 1 becomes more and more difficult.

Thus far, we have presented how to obtain the realization of the reduced-order system which satisfy it's the first-order necessary conditions for $\mathcal{H}_2$ optimality together with retaining the polynomial part of each subsystem, by assuming the structure of the *kth* order multivariate transfer function of the reduced-order system as in (3.5).

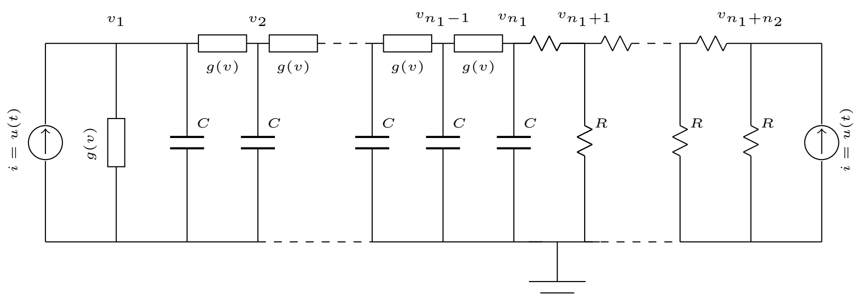


Figure 1: Nonlinear transmission line circuit.

However, the corresponding time-domain bilinear system can be given by

$$\begin{aligned}\hat{E}\dot{\hat{x}}(t) &= \hat{A}\hat{x}(t) + \hat{N}\hat{x}(t)u(t) + \hat{B}u(t),\\ \hat{y}(t) &= \hat{C}\hat{x}(t) + \sum_{k=1}^{\infty} D_k u^k(t).\end{aligned}$$

For a detailed proof, we refer to [12]. Also, therein, the computational issue of $\sum\limits_{k=1}^{\infty} D_k u^k(t)$ is also discussed and shown how to deal with this summation cheaply.

5 Numerical Results

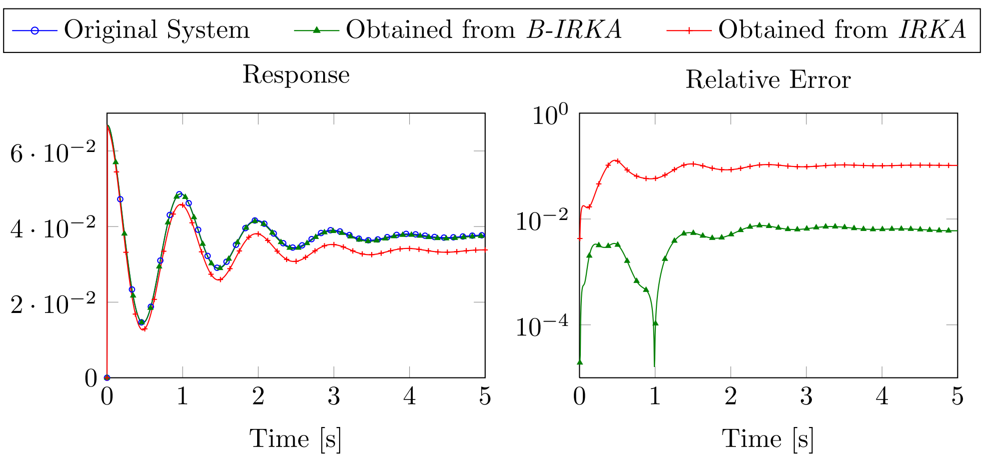
In this section, we illustrate the performance of the proposed *B-IRKA* for bilinear descriptor systems using various numerical examples. We also compare them with the reduced bilinear system, whose matrices $\hat{E}, \hat{A}, \hat{B}$ and $\hat{C}$ are obtained by linear *IRKA* [14, Algo. 5.2], and the reduced bilinear terms, $\hat{N}$, are determined by simply projecting the bilinear terms by using the same projection matrices. The stopping criterion for Algorithm 1 is chosen based on the relative change of the norm of the poles of the reduced-order system. If the relative change becomes smaller than *tol* then we stop the iterations, where *tol* is chosen as the square-root of machine precision. Moreover, the initialization of the algorithm is done by choosing arbitrary interpolation points and tangential directions. We also consider a scaling factor for smooth convergence of *B-IRKA* as discussed in [6, 10]. All the simulations are carried out in MATLAB® version 7.11.0.584(R2010b)64-bit(glnza64) on an Intel(R) Core(TM)2 Quad CPU Q9550 @2.83GHz, 6MB cache, 4GB RAM, openSUSE 12.1 (x86-64).

5.1 Nonlinear RC Circuit

As a first example, we study the nonlinear transmission line circuit whose circuit diagram is shown in Figure 1. The nonlinearity in the system appears due to the diode I-V characteristic $g(v_D) = e^{40v_D} + v_D - 1$, where v_D is the voltage across the diode. As discussed in [12], the system can be modelled as a quadratic-bilinear descriptor system

(QBDAE) of dimension $(2n_1 + n_2)$, where n_1 and n_2 are the numbers of capacitors (C) and linear resistors (R), respectively. The output of the system is the average voltage over all nodes. We set $n_1 = 10$ and $n_2 = 20$, and all electrical component equal to 1, leading a QBDAE of order $n = 40$ with the structure of the matrices E and A as shown in (3.3). However, Carleman bilinearization for descriptor systems [13] is employed on the QBDAE which gives us a bilinearized system of order 840. The polynomial part of the first subsystem of the bilinearized system is $D_1 = 0.0333$ and higher order subsystems all have zero polynomial parts.

We determine the reduced-order system of order $r = 5$ by employing Algorithm 1. We choose the scaling factor $\gamma = 0.5$. We also determine the reduced bilinear system by employing *IRKA*. To illustrate the accuracy of the reduced-order systems, we plot the time-domain response for the input $u(t) = \cos(2\pi t)e^{-t} + 1$ in Figure 2a, and the relative errors are shown in Figure 2b. Also, we compare the reduced bilinear system obtained by applying Algorithm 1 with the reduced bilinear system determined by the linear *IRKA*.



(a) Transient response for an input $u(t) = \cos(2\pi t)e^{-t} + 1$.

(b) Relative error.

Figure 2: Comparison of reduced-order systems obtained by employing *B-IRKA* and *IRKA* with the original system.

Evidently, the reduced-order system obtained by using *B-IRKA* replicates the input-output behaviour of the original system much better as compared to the reduced-order system obtained by using *IRKA*. We also observe that the higher accuracy of the reduced-order system, determined by *B-IRKA*, can be achieved by increasing the order of the reduced system. On the other hand, the reduced bilinear system determined by using *IRKA* produces an unstable reduced-order system, as the order of the reduced system increases.

5.2 Parametric RLC Circuit

Next, we consider an RLC circuit as shown in Figure 3 whose first node has three branches, connected to the voltage source $\mathcal{V}$ via a constant resistance R_c, to a variable resistance, and to ground via a capacitor. The last, nth, node of the circuit is connected to the ground via a capacitor. All other nodes also have three branches; the first one is grounded via a capacitor; the second one is connected to an inductor and the third one is connected to a variable resistor as shown in Figure 3.

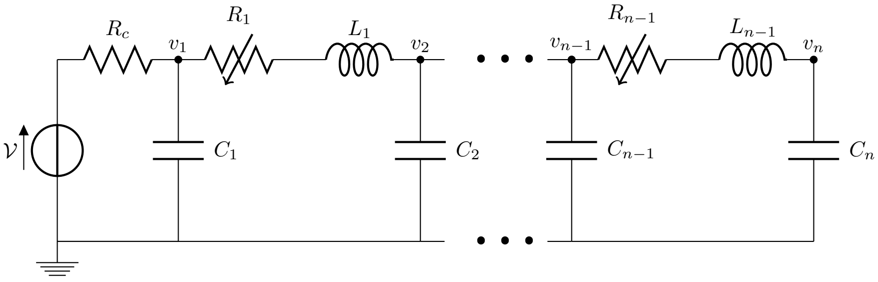


Figure 3: RLC circuit diagram.

Using Kirchhoff's voltage law at each node, we obtain the following system of equations:

$$\begin{aligned}
C_j \frac{v_j(t)}{dt} &= i_j - i_{j+1}, && j = 1, 2, \ldots, n-1, \\
L_j \frac{i_{j+1}}{dt} &= -R_j i_{j+1} + v_{j+1} - v_j, && j = 1, 2, \ldots, n-1, \\
0 &= v_1 + i_1 R_c - \mathcal{V}, \\
C_n \tfrac{v_n}{dt} &= i_n.
\end{aligned}$$

Here, we set all the capacitors C, inductors L, and the resistance R_C equal to 1. We also assume that the variable resistances vary linearly with the parameter p as follows:

$$R_j = \mathcal{R}_j(1 + p). \tag{5.1}$$

Also, we consider $\mathcal{R}_j = 1$. Combining all these equations and utilizing the parametric relation of the variable resistance, we obtain the following parametric linear system:

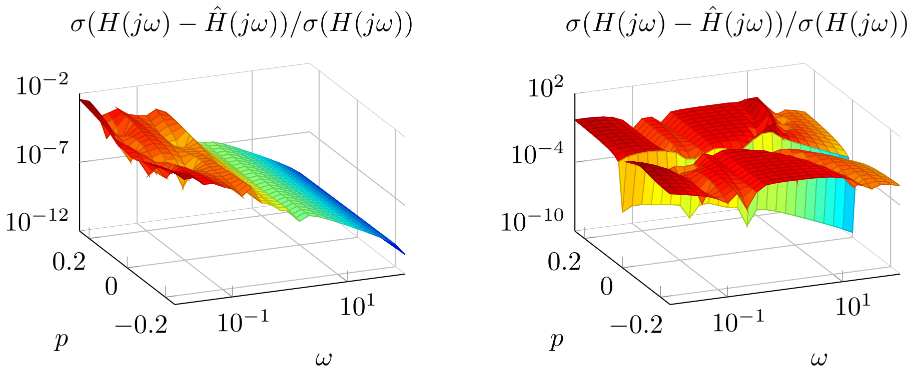
$$\begin{aligned}
E\dot{x}(t) &= Ax(t) + pA_1x(t) + Bu(t), \\
y(t) &= Cx(t),
\end{aligned} \tag{5.2}$$

where $x(t)$ is the state vector containing the voltage at each node and current through resistances. The input $u(t)$ is the voltage source, and the quantity of interest $y(t)$ is the current through the voltage source. It has been shown in [5] that a special class of linear parametric systems can be treated as bilinear systems, by re-writing parameters

p as inputs to the system. Therefore, we can write the system 5.2 as a bilinear system with two inputs $\tilde{u}(t) = [u(t), p]^T$ as follows:

$$\begin{aligned} E\dot{x}(t) &= A(t) + \sum_{i=1}^{2} N_i x(t)\tilde{u}(t) + \tilde{B}\tilde{u}(t), \\ y(t) &= Cx(t), \end{aligned} \tag{5.3}$$

where $N_1 = 0, N_2 = A_1$, and $\tilde{B} = [B, \mathbf{0}]$. We determine reduced bilinear systems of order $r = 15$ using *B-IRKA*. We choose the scaling factor $\gamma = 0.1$ for smooth convergence of *B-IRKA*. We also determine the reduced bilinear system by employing *IRKA* of the same order. The computed reduced bilinear systems can be again re-written as reduced parametric linear systems. In Figure 4, we show the comparison of the transfer functions of the original and reduced-order systems with respect to the parameter p and the frequency ω. In Figure 5, we plot the relative $\mathcal{H}_\infty$ error with respect to the parameter p.



(a) Using *B-IRKA*. (b) Using *IRKA*.

Figure 4: Comparison of the reduced-order systems obtained by using *B-IRKA* and *IRKA* with the original system.

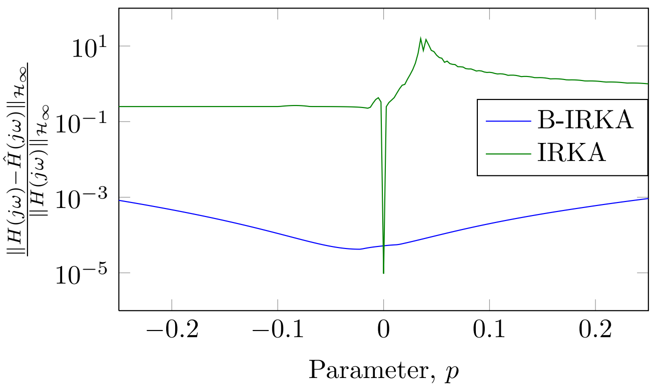


Figure 5: Comparison of relative $\mathcal{H}_\infty$-norm.

These figures clearly show that the reduced-order system obtained by using *B-IRKA* outperforms the one obtained by using *IRKA* for a wide range of the parameter. However, the projection matrices computed by using *IRKA* capture the dynamics of the original system very well in the vicinity of the parameter $p = 0$. This is why one can see a sharp drop in the relative errors in Figure 4 and 5 around the parameter $p = 0$.

6 Conclusions

In this paper, we have extended the multipoint Volterra series interpolation to a family of bilinear descriptor systems with the polynomial part of its *kth* order multivariate transfer function being constant. We have presented the modified interpolation conditions which not only achieve multipoint interpolation of the underlying Volterra series, but also retain the polynomial part of each subsystem. Based on the first-order necessary conditions for $\mathcal{H}_2$ optimality, we have proposed an iterative rational Krylov algorithm, the so-called *B-IRKA* for bilinear descriptor systems, which converges to a locally $\mathcal{H}_2$ optimal reduced-order system, if it converges. Using two numerical examples, we have demonstrated the efficiency of the proposed methodology and compared it with reduced-order systems obtained by using *IRKA*.

References

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

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

[3] U. Baur, P. Benner, and L. Feng. Model order reduction for linear and nonlinear systems: A system-theoretic perspective. 21(4):331–358, 2014.

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

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

[6] 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.

[7] 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.

[8] P. Benner, V. Mehrmann, and D.C. Sorensen. *Dimension Reduction of Large-Scale Systems.* Springer-Verlag, Berlin/Heidelberg, 2005.

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

[10] G. M. Flagg. *Interpolation Methods for the Model Reduction of Bilinear Systems.* PhD thesis, Virginia Polytechnic Institute and State University, 2012.

[11] G. M. Flagg and S. Gugercin. Multipoint Volterra series interpolation and $\mathcal{H}_2$ optimal model reduction of bilinear systems. *SIAM J. Matrix Anal. Appl.*, 36(2):549–579, 2015.

[12] P. Goyal, M. I. Ahmad, and P. Benner. Krylov subspace-based model reduction for a class of bilinear descriptor systems. Preprint MPIMD/15-07, Max Planck Institute Magdeburg, 2015. Available from `http://www.mpi-magdeburg.mpg.de/preprints/`.

[13] P. Goyal, M. I. Ahmad, and P. Benner. Model reduction of quadratic-bilinear descriptor systems via Carleman bilinearization. *Proceedings of European Control Conference*, pages 1171–1176, 2015.

[14] 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.

[15] C. Hartmann, A. Zueva, and B. Schäfer-Bung. Balanced model reduction of bilinear systems with applications to positive systems. *SIAM J. Cont. Optim.*, 51(3):2356–2378, 2013.

[16] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution.* European Mathematical Society, 2006.

[17] R. R. Mohler. Natural bilinear control processes. *IEEE Transactions on Systems Science and Cybernetics*, 6(3):192–197, 1970.

[18] R. R. Mohler. *Nonlinear Systems (Vol. 2): Applications to Bilinear Control.* Prentice-Hall, Inc., NJ, 1991.

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

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

[21] 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.

[22] D. D. Weiner and J. F. Spina. *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits: With Application to Nonlinear Interference Effects.* New York: Van Nostrand Reinhold, 1980.

[23] D. A. Wilson. Optimum solution of model-reduction problem. *Proc. Inst. Ele. Eng.*, 117(6):1161–1165, 1970.

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