2015 European Control Conference (ECC)
July 15-17, 2015. Linz, Austria

Model Reduction of Quadratic-Bilinear Descriptor Systems via Carleman Bilinearization*

Pawan Goyal†, Mian Ilyas Ahmad† and Peter Benner†

***Abstract*— We propose a model reduction technique for quadratic-bilinear descriptor systems. The approach involves approximating the system by a bilinear descriptor system using Carleman bilinearization [1]. It is shown that, by assuming a particular structure of the matrix pencil, the bilinearization process preserves the structure of the matrix pencil in the bilinearized system. Further, we extend the use of the bilinear iterative rational Krylov algorithm (B-IRKA) [2] to descriptor systems to identify a locally $\mathcal{H}_2$-optimal reduced-order system for the bilinearized system under the assumption that the $\mathcal{H}_2$ norm of the system exists. Applications to the simulation of a nonlinear RC circuit and a lid-driven cavity flow are presented to illustrate the proposed methodology.**

I. INTRODUCTION

We consider a single-input single-output (SISO) quadratic-bilinear descriptor system

$$E\dot{x}(t) = Ax(t) + H(x \otimes x) + Nx(t)u(t) + Bu(t), \quad \text{(1a)}$$
$$y(t) = Cx(t) + Du(t), \quad x(0) = 0, \quad \text{(1b)}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, and $u(t)$, $y(t)$ denote the input and output of the system, respectively. Also $E,\ A,\ N,\ B$ and C are state-space matrices with dimensions fixed by the state, input and output of the system, while H is a matricization of the Hesse tensor of the right-hand side of the system of differential-algebraic equations (DAEs) in (1a). All our considerations can also be extended to multi-input multi-output systems, however for simplicity of representation, we stick to the single-input single-output (SISO) case. There are many applications including the simulation of fluid flow and the dynamics of electrical circuits [3], which are naturally in quadratic-bilinear form. Also a large class of nonlinear systems can be transformed to quadratic-bilinear systems [4]. Often these models are of large scale, and thus, model order reduction (MOR) may become necessary before deriving observers or controllers.

MOR approximates these quadratic-bilinear systems by a reduced-order system in an efficient way. Various approaches have been proposed in the literature to tackle the problem of MOR for nonlinear systems. These include trajectory-based methods such as proper orthogonal decomposition (POD) [5], [6], [7] and its extensions [8], the reduced basis method [9], and the trajectory piecewise linear method (TPWL) [10]. All these methods are dependent on the trajectory of the original state vector for a given input. However if the input function is varied, which is common in control problems, the approximations obtained through these methods for the first input may not capture the state trajectory associated with the varied input.

An alternative method is also available. There, the reduced-order system is constructed such that it approximates the input-to-output behavior of the original system and thus, the reduced-order model generation is independent of the input variations. The approach extends the concept of interpolation or moment matching for linear systems to quadratic-bilinear systems [3], [11], [4]. The moments of the quadratic-bilinear system are defined by an infinite series of subsystems that are associated with the quadratic system. Each of the subsystems defines a moment at a particular frequency and the structure of these moments becomes more and more complex as the number of subsystems increases.

To avoid these issues, another approach has been used in the literature for reducing the quadratic-bilinear system. This involves first identifying an approximate bilinearized system using Carleman bilinearization and then applying the well-established bilinear model reduction techniques. Although the bilinearized system looses the quadratic-bilinear structure and increases the dimension of the system, it simplifies the MOR procedure. Note that just dropping the quadratic term in (1) also leads to a bilinear system, but completely ignoring the quadratic nonlinearity usually will yield a reduced-order model which may not capture the dynamics of the system.

The bilinearization allows us to utilize the bilinear iterative rational Krylov algorithm (B-IRKA) [2] for bilinear systems. Carleman bilinearization for nonlinear ODE systems has been discussed, e.g., in [1]. However, for descriptor systems, it still requires further research. In this paper our main focus is on the Carleman bilinearization of descriptor systems and on describing how a particular structure of the matrix pencil $\lambda E - A$ propagates with the bilinearization process.

We begin with Carleman bilinearization of the quadratic-bilinear system (1) in Section 2. Also in this section, a particular structure is considered allowing one to show that the bilinearization preserves the structure of the system. In Section 3, we briefly outline the B-IRKA method which is utilized on the bilinearized system to compute a reduced-order system. In Section 4, we present the numerical results, and we close with conclusions and future directions.

II. CARLEMAN APPROXIMATION

In this section, we discuss the Carleman bilinearization steps for descriptor systems and show how the properties of

*This research is supported by a research grant of the "International Max Planck Research School (IMPRS) for Advanced Methods in Process and System Engineering" (Magdeburg).

†Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, email: {goyalp,imahmad,benner}@mpi-magdeburg.mpg.de.

the matrix pencil $\lambda E - A$ propagate in the bilinearization process. Consider the application of Carleman bilinearization [1] to the quadratic-bilinear system (1) to obtain the following bilinear system, after neglecting the cubic term:

$$\begin{aligned} E^{\otimes}\dot{x}^{\otimes}(t) &= A^{\otimes}x^{\otimes}(t) + N^{\otimes}x^{\otimes}(t)u(t) + B^{\otimes}u(t), \\ y^{\otimes}(t) &= C^{\otimes}x^{\otimes}(t) + Du(t), \end{aligned} \tag{2}$$

where

$$E^{\otimes} = \begin{bmatrix} E & 0 \\ 0 & E\otimes E \end{bmatrix}, \quad N^{\otimes} = \begin{bmatrix} N & 0 \\ \mathcal{L}(E,B) & \mathcal{L}(E,N) \end{bmatrix},$$
$$A^{\otimes} = \begin{bmatrix} A & H \\ 0 & \mathcal{L}(E,A) \end{bmatrix}, \quad B^{\otimes} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad C^{\otimes} = \begin{bmatrix} C^T \\ 0 \end{bmatrix}^T$$

in which $x^{\otimes} = \begin{bmatrix} x^T & x^T \otimes x^T \end{bmatrix}^T$ and $\mathcal{L}(\mathcal{A},\mathcal{B}) = \mathcal{A}\otimes\mathcal{B} + \mathcal{B}\otimes\mathcal{A}$. In the following, we assume a special structure of the matrix pencil $\lambda E - A$ with

$$E = \begin{bmatrix} E_{11} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{3}$$

in which the n rows and columns of E and A are partitioned into n_1 and n_2, implying that $E_{11} \in \mathbb{R}^{n_1\times n_1}$ and $A_{22} \in \mathbb{R}^{n_2\times n_2}$. It is assumed that E_{11} and A_{22} are both invertible, ensuring that the pencil has a nilpotency index 1, see [12] for details on the index concept. Also, the other matrices in (1) are partitioned in a similar fashion:

$$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

The main purpose of assuming a particular structure is to be able to identify the nilpotency index of the matrix pencil $\lambda E^{\otimes} - A^{\otimes}$. For this, we do some simple transformations before presenting the main results.

Let the system in (2) be transformed by using $x^{\otimes} = Qz(t)$, where $Q = \left[\begin{smallmatrix} I & 0 \\ 0 & P \end{smallmatrix}\right]$ is the transformation matrix, in which $P \in \mathbb{R}^{n^2\times n^2}$ is an invertible matrix. Then

$$\begin{aligned} \underbrace{Q^TE^{\otimes}Q}_{\tilde{E}}\dot{z}(t) &= \underbrace{Q^TA^{\otimes}Q}_{\tilde{A}}z(t) + \underbrace{Q^TN^{\otimes}Q}_{\tilde{N}}z(t)u + B^{\otimes}u(t), \\ y^{\otimes}(t) &= C^{\otimes}z(t) + Du(t). \end{aligned} \tag{4}$$

Note that $Q^TB^{\otimes} = B^{\otimes}$ and $C^{\otimes}Q = C^{\otimes}$. The structure of the remaining block matrices in the transformed system becomes

$$\begin{aligned} \tilde{E} &= \begin{bmatrix} E & 0 \\ 0 & P^T(E\otimes E)P \end{bmatrix}, \quad z(t) = \begin{bmatrix} x(t) \\ P^{-1}(x\otimes x) \end{bmatrix}, \\ \tilde{A} &= \begin{bmatrix} A & HP \\ 0 & P^T\mathcal{L}(E,A)P \end{bmatrix}, \\ \tilde{N} &= \begin{bmatrix} N & 0 \\ P^T\mathcal{L}(E,B) & P^T\mathcal{L}(E,N)P \end{bmatrix}. \end{aligned} \tag{5}$$

Next, we want to exploit the structure of these matrices during Carleman bilinearization in order to determine the nilpotency index of the bilinearized system (2). The following theorem suggests a suitable choice of the matrix P so that the Kronecker products in (4) can be written as the corresponding Tracy-Singh product, see for example [13], [14] regarding details of these matrix products.

Theorem 2.1: Let $\mathcal{A},\mathcal{B} \in \mathbb{R}^{n\times m}$ be block matrices that are partitioned as

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} \quad \text{and} \quad \mathcal{B} = \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} \\ \mathcal{B}_{21} & \mathcal{B}_{22} \end{bmatrix}.$$

Then there exist permutation matrices $P_j \in \mathbb{R}^{n^j,n^j}$, $j = 1,2$, such that

$$P_1^T(\mathcal{A}\otimes\mathcal{B})P_2 = \mathcal{A}\circ\mathcal{B}, \tag{6}$$

where $\mathcal{A}\circ\mathcal{B}$ represents the Tracy-Singh product defined by

$$\mathcal{A}\circ\mathcal{B} = \begin{bmatrix} \mathcal{A}_{11}\circ\mathcal{B} & \mathcal{A}_{12}\circ\mathcal{B} \\ \mathcal{A}_{21}\circ\mathcal{B} & \mathcal{A}_{22}\circ\mathcal{B} \end{bmatrix},$$

where $\mathcal{A}_{ij}\circ\mathcal{B} = \begin{bmatrix} \mathcal{A}_{ij}\otimes\mathcal{B}_{11} & \mathcal{A}_{ij}\otimes\mathcal{B}_{12} \\ \mathcal{A}_{ij}\otimes\mathcal{B}_{21} & \mathcal{A}_{ij}\otimes\mathcal{B}_{22} \end{bmatrix}$.

Remarks:

- If $\mathcal{A}$ and $\mathcal{B}$ are both square matrices of the same size, and $\mathcal{A}_{ii}$ and $\mathcal{B}_{kk}$ are square block matrices, then $A\circ B$ and $A\otimes B$ are permutation similar. That is, $P_1 = P_2$.
- Let $\mathcal{A}$ and/or $\mathcal{B}$ be partitioned as row block matrices. Then $\mathcal{A}\circ\mathcal{B} = P_1^T(\mathcal{A}\otimes\mathcal{B})I$, where I is an identity matrix of appropriate size.
- Since P_1 and P_2 are permutation matrices, $P_j^{-1} = P_j^T$ for $j = 1,2$.

This means that the Kronecker products in (5) can be represented by Tracy-Singh products, if one chooses P as in Theorem 2.1. That is,

$$\begin{aligned} &P^T(E\otimes E)P = E\circ E, \ P^T\mathcal{L}(A,E)P = A\circ E + E\circ A, \\ &P^T\mathcal{L}(B,E) = B\circ E + E\circ B, \qquad P^T(x\otimes x) = x\circ x, \\ &P^T\mathcal{L}(N,E)P = N\circ E + E\circ N. \end{aligned}$$

Using the special structure in (3) and Theorem 2.1, the permutation matrix P modifies the transformed state-space matrices in (5) to become

$$P^T\mathcal{L}(E,A)P = \begin{bmatrix} \mathcal{L}(A_{11},E_{11}) & E_{11}\otimes A_{12} & A_{12}\otimes E_{11} & 0 \\ E_{11}\otimes A_{21} & E_{11}\otimes A_{22} & 0 & 0 \\ A_{21}\otimes E_{11} & 0 & A_{22}\otimes E_{11} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$P^T\mathcal{L}(E,N)P = \begin{bmatrix} \mathcal{L}(N_{11},E_{11}) & E_{11}\otimes N_{12} & N_{12}\otimes E_{11} & 0 \\ E_{11}\otimes N_{21} & E_{11}\otimes N_{22} & 0 & 0 \\ N_{21}\otimes E_{11} & 0 & N_{22}\otimes E_{11} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$P^T(E\otimes E)P = \begin{bmatrix} E_{11}\otimes E_{11} & 0 \\ 0 & 0 \end{bmatrix}, \ P^T\mathcal{L}(B,E) = \begin{bmatrix} \mathcal{L}(B_1,E_{11}) \\ E_{11}\otimes B_2 \\ B_2\otimes E_{11} \\ 0 \end{bmatrix},$$

$$P^T(x\otimes x) = \begin{bmatrix} x_1\otimes x_1 \\ x_1\otimes x_2 \\ x_2\otimes x_1 \\ x_2\otimes x_2 \end{bmatrix}, \ HP = \begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{12} & \tilde{H}_{13} & \tilde{H}_{14} \\ \tilde{H}_{21} & \tilde{H}_{22} & \tilde{H}_{23} & \tilde{H}_{24} \end{bmatrix},$$

where HP is partitioned into 2 block rows of size n_1 and n_2 and 4 block columns of size n_1^2, n_1n_2, n_1n_2, and n_2^2, respectively. Let $x_i\otimes x_j$ be represented by x_{ij} for $i,j = 1,2$. Then, the following lemma shows how to represent x_{21} in terms of x_{12} or vice versa.

Lemma 2.1: For $A \in \mathbb{R}^{p\times q}$ and $B \in \mathbb{R}^{r\times s}$,

$$B \otimes A = S_{p,r}(A \otimes B)S_{q,s},$$

where $S_{m,n} = \sum_{i=1}^{m}(e_i^T \otimes I_n \otimes e_i) \in \mathbb{R}^{mn,mn}$, in which e_i is the i-th column of the $m \times m$ identity matrix.

This implies that $x_{21} = S_{n_1,n_2}x_{12}$, where $S_{n_1,n_2} \in \mathbb{R}^{n_1n_2,n_1n_2}$. The above observations show that some rows and columns of the transformed system for the special structure are zero and therefore, the system includes redundant states. The next theorem helps in identifying and eliminating these redundant states to obtain a trimmed bilinear system with a specific structure of the matrix pencil similar to (3).

Theorem 2.2: Let a quadratic-bilinear DAE system with the matrix pencil $\lambda E - A$ as in (3) be transformed to a bilinear DAE system of the form (4). Then, the transformed bilinear system can be represented by a trimmed bilinear DAE system

$$\begin{aligned} \mathcal{E}_t\dot{x}_t(t) &= \mathcal{A}_t x_t(t) + \mathcal{N}_t x_t(t)u(t) + \mathcal{B}_t u(t),\\ y_t(t) &= \mathcal{C}_t x_t(t) + \mathcal{D}_t u(t), \end{aligned} \tag{7}$$

where

$$\mathcal{A}_t = \begin{bmatrix} A_{11} & \tilde{H}_{11} & A_{12} & \tilde{H}_{12} + \tilde{H}_{13}S_{n_1,n_2}\\ 0 & \mathcal{L}(E_{11}, A_{11}) & 0 & S_{n_1}^I(E_{11}\otimes A_{12})\\ A_{21} & \tilde{H}_{21} & A_{22} & \tilde{H}_{22} + \tilde{H}_{23}S_{n_1,n_2}\\ 0 & E_{11}\otimes A_{21} & 0 & E_{11}\otimes A_{22} \end{bmatrix},$$

$$\mathcal{E}_t = \begin{bmatrix} E_{11} & 0 & 0 & 0\\ 0 & E_{11}\otimes E_{11} & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathcal{B}_t = \begin{bmatrix} B_1\\ 0\\ B_2\\ 0 \end{bmatrix}, \mathcal{C}_t = \begin{bmatrix} C_1^T\\ 0\\ C_2^T\\ 0 \end{bmatrix}^T,$$

$$\mathcal{N}_t = \begin{bmatrix} N_{11} & 0 & N_{12} & 0\\ \mathcal{L}(E_{11}, B_1) & \mathcal{L}(E_{11}, N_{11}) & 0 & S_{n_1}^I E_{11}\otimes N_{12}\\ N_{21} & 0 & N_{22} & 0\\ E_{11}\otimes B_2 & E_{11}\otimes N_{21} & 0 & E_{11}\otimes N_{22} \end{bmatrix},$$

$$x_t = \begin{bmatrix} x_1^T & x_{11}^T & x_2^T & x_{12}^T \end{bmatrix}^T, \ \mathcal{D}_t = D,$$

in which $S_{n_1}^I = I + \sum_{i=1}^{n_1}(e_i^T \otimes I_{n_1} \otimes e_i)$. Also, the nilpotency index of the matrix pencil (E, A) is inherited by $(\mathcal{E}_t, \mathcal{A}_t)$. That is, $(\mathcal{E}_t, \mathcal{A}_t)$ has nilpotency index 1.

Proof: Let the state-space equations in the transformed bilinear system (4) be re-arranged such that the state vector $z(t)$ becomes $\tilde{z}(t) = \begin{bmatrix} x_t^T & x_{21}^T & x_{22}^T \end{bmatrix}^T$. Thus, if we can show that the rows and columns corresponding to x_{21} and x_{22} are redundant and the remaining variables satisfy (7), then we are done. To see this, note that the columns in $E^{\otimes}$ and rows in all matrices corresponding to x_{22} in (4) are all zero for the special structured matrix pencil in (3). Thus, the complete row and all columns corresponding to x_{22} can be removed, since the initial condition is zero.

As the two algebraic block equations in the bilinear DAE system (4) corresponding to x_{12} and x_{21} are permutation equivalent, the following two equations are also permutation equivalent:

$$\begin{aligned} &(E_{11}\otimes A_{21})x_{11} + (E_{11}\otimes A_{22})x_{12} + (E_{11}\otimes N_{21})x_{11}u(t)\\ &\quad + (E_{11}\otimes N_{22})x_{12}u(t) + (E_{11}\otimes B_2)x_1u(t) = 0, \end{aligned} \tag{8}$$

$$\begin{aligned} &(A_{21}\otimes E_{11})x_{11} + (A_{22}\otimes E_{11})x_{21} + (A_{21}\otimes E_{11})x_{11}u(t)\\ &\quad + (N_{22}\otimes E_{11})x_{21}u(t) + (B_2\otimes E_{11})x_1u(t) = 0. \end{aligned} \tag{9}$$

To see this, we begin with the first equation,

$$\begin{aligned} &(E_{11}\otimes A_{21})(x_1\otimes x_1) + (E_{11}\otimes A_{22})(x_1\otimes x_2)\\ &\quad + (E_{11}\otimes N_{21})(x_1\otimes x_1)u(t)\\ &\quad + (E_{11}\otimes N_{22})(x_1\otimes x_2)u(t) + (E_{11}\otimes B_2)x_1u(t) = 0, \end{aligned}$$

which turns into

$$\begin{aligned} &(E_{11}x_1\otimes A_{21}x_1) + (E_{11}x_1\otimes A_{22}x_2)\\ &\quad + (E_{11}x_1\otimes N_{21}x_1)u(t) + (E_{11}x_1\otimes N_{22}x_2)u(t)\\ &\qquad + (E_{11}\otimes B_2)x_1u(t) = 0. \end{aligned}$$

Using Lemma 2.1, this can be re-written as

$$\begin{aligned} &S_{n_1,n_1}\Big(A_{21}x_1\otimes E_{11}x_1) + (A_{22}x_2\otimes E_{11}x_1)\\ &\quad + (N_{21}x_1\otimes E_{11}x_1)u(t) + (N_{22}x_2\otimes E_{11}x_1)u(t)\\ &\qquad + (B_2\otimes E_{11})x_1u(t)\Big) = 0, \end{aligned}$$

and, finally, we obtain

$$\begin{aligned} &S_{n_1,n_1}\Big((A_{21}\otimes E_{11})(x_1\otimes x_1) + (A_{22}\otimes E_{11})(x_2\otimes x_1)\\ &\quad + (N_{21}\otimes E_{11})(x_1\otimes x_1)u(t)\\ &\quad + (N_{22}\otimes E_{11})(x_2\otimes x_1)u(t) + (B_2\otimes E_{11})x_1u(t)\Big) = 0, \end{aligned}$$

which reduces to (9) because S_{n_1,n_1} is invertible. Thus, we can get rid of one of these constraints in the bilinear DAE system. Based on the above reasoning, we eliminate (9) from the system. Now by utilizing $x_{21} = S_{n_1,n_1}x_{12}$, we can combine the two columns associated with x_{12} and x_{21}. The above changes result in a trimmed bilinear DAE with structure as in (7) and the order of the system reduces from $n(n+1)$ to $n(n_1+1)$.

Note that the matrix pencil of the trimmed bilinear DAE (7) can be partitioned as

$$\mathcal{E}_t = \begin{bmatrix} \mathcal{E}_{t_{11}} & 0\\ 0 & 0 \end{bmatrix}, \quad \mathcal{A}_t = \begin{bmatrix} \mathcal{A}_{t_{11}} & \mathcal{A}_{t_{12}}\\ \mathcal{A}_{t_{21}} & \mathcal{A}_{t_{22}} \end{bmatrix},$$

where $\mathcal{E}_{t_{11}}$ and $\mathcal{A}_{t_{22}}$ are the square and block triangular matrices of size $n_1(n_1+1)$ and $n_2(n_1+1)$, respectively, and has a structure as in (7). Clearly $\mathcal{E}_{t_{11}}$ and $\mathcal{A}_{t_{22}}$ are invertible and thus the pencil has nilpotency index 1. ■

The above results are useful for reducing a quadratic-bilinear DAE system of order $n = n_1 + n_2$ to a trimmed bilinear DAE system of order $\tilde{n} = n(n_1+1)$. Although the dimension of the bilinear system is still much higher than the dimension of the original system (1), the hope is to get a satisfactory reduced-order bilinear system of dimension much smaller than n. Thus our goal is to identify a reduced system of order $r \ll n$ that approximates the bilinearized DAE system which is expected to capture the dynamics of the original quadratic-bilinear DAE system as well.

III. Model Reduction Approach

In this section, we begin with a brief review of the bilinear iterative rational Krylov algorithm (B-IRKA) [2], [15] and see how B-IRKA can be utilized to reduce the trimmed bilinear DAE system (7). We first consider a bilinear system of the form

$$\Sigma_B = \begin{cases} E\dot{x} = Ax + Nxu + Bu, \\ \quad y = Cx, \quad x(0) = 0, \end{cases} \tag{10}$$

where E is a non-singular matrix. The aim is to construct a reduced-order system of the form

$$\hat{\Sigma}_B = \begin{cases} \hat{E}\dot{\hat{x}} = \hat{A}\hat{x} + \hat{N}\hat{x}u + \hat{B}u, \\ \quad \hat{y} = \hat{C}x, \quad \hat{x}(0) = 0, \end{cases} \tag{11}$$

such that $\hat{y} \approx y$ for all admissible inputs $u(t)$. We construct such a reduced system by applying B-IRKA to the system in (10). B-IRKA will converge to a locally $\mathcal{H}_2$ optimal reduced system satisfying the necessary conditions derived in [2] for $\mathcal{H}_2$ optimality, if it converges. This means that the resulting reduced-order bilinear system of dimension r locally minimizes the $\mathcal{H}_2$ norm of the error system.

The $\mathcal{H}_2$ norm for bilinear systems is introduced in [16] and can be expressed in terms of the reachability Gramian, P, or the observability Gramian, Q, as

$$\|\Sigma_B\|_{\mathcal{H}_2} = \sqrt{CPC^T} = \sqrt{B^TQB},$$

where the Gramians P and Q are solutions of the following generalized Lyapunov equations:

$$\begin{aligned} APE^T + EPA^T + NPN^T + BB^T = 0, \\ A^TQE + E^TQA + N^TQN + C^TC = 0. \end{aligned}$$

In the following, we outline the main steps of B-IRKA for constructing the required $\mathcal{H}_2$-optimal reduced bilinear system. (Here, $\mathrm{vect}(M)$ denotes the reshaping operation of stacking the columns of M on top of each other.)

Algorithm 1 B-IRKA for Bilinear Descriptor Systems

Input: E, A, N, B, C.
Choose an initial guess for $\hat{E}, \hat{A}, \hat{N}, \hat{B}, \hat{C}$.
while not converged **do**
  Choose S, R such that $S\hat{A}R = \Lambda$ and $S\hat{E}R = I_r$.
  Define $\tilde{B} = \hat{B}^TS^T, \tilde{C} = \hat{C}R, \tilde{N} = R^T\hat{N}^TS^T$ and $\mathcal{P} = \Lambda \otimes E + \hat{E} \otimes A + \tilde{N}^T \otimes N$.
  $\mathrm{vect}(X) = -\mathcal{P}^{-1}(\tilde{B}^T \otimes B)\mathcal{I}_m, \quad \mathcal{I}_q = \mathrm{vect}(I_q)$.
  $\mathrm{vect}(Y) = -\mathcal{P}^{-T}(\tilde{C}^T \otimes C^T)\mathcal{I}_p$.
  $V = \mathrm{orth}(X), \quad W = \mathrm{orth}(Y)$.
  $\hat{E} = W^TEV,\ \hat{A} = W^TAV,\ \hat{N} = W^TNV,$ $\hat{B} = W^TB, \quad \hat{C} = CV$.
end while
Output: $\hat{E}, \hat{A}, \hat{N}, \hat{B}, \hat{C}$.

In case of descriptor systems, i.e., singular E, one can still apply the above extended version of B-IRKA to obtain a reduced-order system as long as the reduced matrix $\hat{E}$ is nonsingular and each subsystem of the original bilinear descriptor system has zero polynomial part. In other words,

$$\lim_{s_1, s_2, \ldots, s_k \to \infty} H_k(s_1, s_2, \ldots, s_k) = 0,$$

where

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

represents the k-dimensional multivariate transfer function associated with the k-th subsystem of the bilinear system. For details, we refer to [1], [17]. It can be seen that Algorithm 1 extends IRKA for index-1 linear descriptor systems [18] to bilinear descriptor system in case of zero polynomial parts. Its use in practice is demonstrated by the examples in the next section, while its theoretical properties require further investigation.

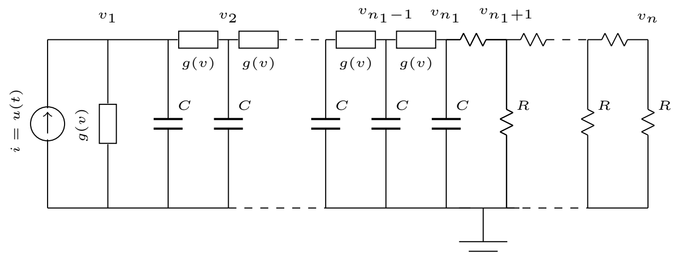

Fig. 1. RC Circuit Diagram.

IV. Numerical Results

In this section, we present our results for Carleman bilinearization and model reduction of quadratic-bilinear descriptor systems. The stopping criterion for Algorithm 1 is based on the relative change of the norm of the poles of the reduced-order system. The algorithm stops when the relative change is below a tolerance value, $tol = \sqrt{\epsilon}$, where ϵ denotes machine precision.

A. Nonlinear RC Circuit

As a first example, we consider a simple extension of the nonlinear RC circuit [4] as shown in Figure 1. All the linear resistors R and the capacitors C are set to the fixed values of 1 and 500, respectively. The nonlinearity in the system is due to the diode I-V characteristics, given by $g(v_D) = e^{40v_D} - 1$, where v_D is the voltage difference between the two nodes. The input and output of the system are current (i) and voltage at node 1, $v_1(t)$, respectively. Using similar steps as in [4], we model the RC circuit as a quadratic-bilinear descriptor system with matrices E and A of structure as shown in (3). Due to space limitations, we skip the modeling details of the circuit in Figure 1. We choose $n_1 = 20$ and $n = 30$, which gives a quadratic-bilinear descriptor system of order 50.

The proposed Carleman bilinearization is used to obtain a bilinearized descriptor system of order $\tilde{n} = 2050$, and is reduced by utilizing the version of *B-IRKA* given in Algorithm 1, with the order of the resulting reduced-order system set to $r = 5$. To show the accuracy of the bilinearized system and the reduced-order bilinear system, we plot the time-domain simulation of the systems in Figure 2 and

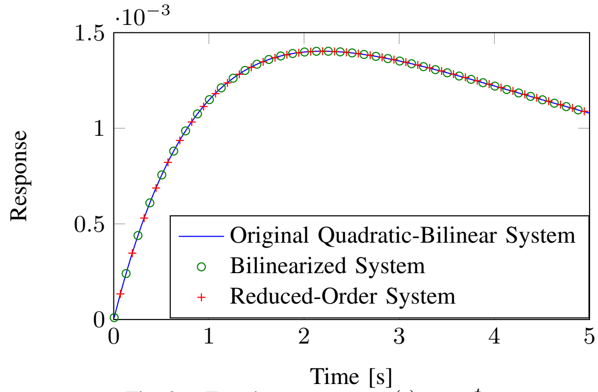

Fig. 2. Transient response, $u(t) = e^{-t}$.

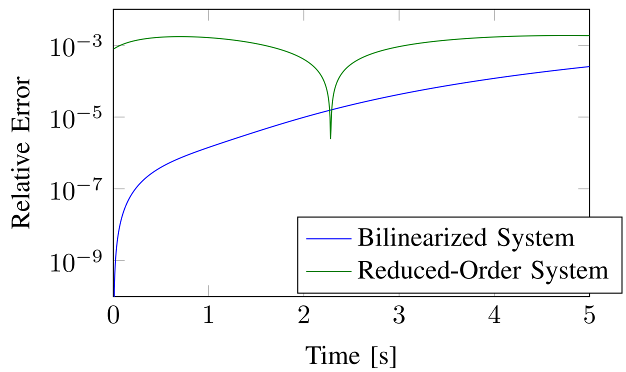

Fig. 3. Relative errors between the transient response for $u(t) = e^{-t}$.

the relative error associated with the approximate systems in Figure 3. Here, we choose an exponentially decaying input function. Similar results are obtained using other input functions. These results show that the important dynamics of the quadratic-bilinear system are captured by the bilinearized system and are approximated by the reduced-order system.

B. Lid-Driven Cavity Flow

We consider a lid-driven cavity problem which is a well-known benchmark example for incompressible fluid flow [19], [20]. The system includes a square cavity with three rigid walls and a moving lid with unit velocity as shown in Figure 4. The flow dynamics of the system are given by ("vorticity-stream function formulation")

$$\frac{\partial \omega}{\partial t} = \frac{\partial \psi}{\partial x}\frac{\partial \omega}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial \omega}{\partial x} + \frac{1}{\mathrm{Re}}\nabla^2 \omega, \tag{12}$$

$$\nabla^2 \psi = -\omega, \tag{13}$$

where ω, ψ and Re represent vorticity, stream function and Reynolds number, respectively, and ∇^2 denotes the Laplacian operator. The boundary conditions on the system are set to mixed Dirichlet and Neumann conditions

$$\psi = 0, \quad \omega = -\frac{\partial^2 \psi}{\partial y^2} \quad \text{on} \quad \Gamma_1 \text{ and } \Gamma_2,$$

$$\psi = 0, \quad \omega = -\frac{\partial^2 \psi}{\partial x^2} \quad \text{on} \quad \Gamma_3 \text{ and } \Gamma_4.$$

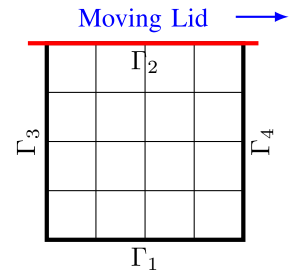

Fig. 4. Lid-driven cavity with boundary conditions.

The moving lid conditions can be captured by the boundary conditions for ψ and ω, since these are related to the velocity of the lid as

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial x},$$

where u and v are the x and y components of the velocity, respectively. To get a mathematical model from the above governing equations together with the boundary conditions, we use a centered finite-difference discretization on an equidistant $k \times k$ mesh on the unit square with nodes ω_{ij} and ψ_{ij}. Let A_p be the Poisson matrix and A_x, A_y the matrices obtained from the operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$, respectively. Then, if $\omega = \mathrm{vect}(\omega_{ij})$, $\psi = \mathrm{vect}(\psi_{ij})$, we have

$$\begin{aligned} \dot{\omega} &= A_x\psi.^*A_y\omega - A_y\psi.^*A_x\omega + A_p\omega, \\ A_p\psi &= \omega, \end{aligned} \tag{14}$$

where $a.^*b$ denotes the element by element product of a and b. We have a Dirichlet condition in ψ which can be easily imposed in the system (14), while Neumann boundary conditions on ω can be written in terms of ψ_{in}, where ψ_{in} represents the stream function value of interior nodes. After simple manipulations, the system (14) leads to the quadratic-bilinear uncontrolled problem with constant lid velocity

$$\begin{aligned} \dot{\omega}_{in} &= A_{11}\omega_{in} + A_{12}\psi_{in} + H\ \omega_{in} \otimes \psi_{in} + B, \\ 0 &= \omega_{in} + A_{22}\psi_{in}, \end{aligned} \tag{15}$$

where H is obtained by re-writing the element-by-element (Hadamard) product of two vectors into Kronecker form. Also, it is easy to see that

$$E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ I & A_{22} \end{bmatrix} \quad \text{and} \quad N = 0.$$

Now let $k = 12$, so that the discretized quadratic descriptor system has order $2k^2 = 288$, with one input and k outputs as stream function variables in the full domain. Also, the Reynolds number (Re) of the system is set to 10. We apply the proposed bilinearization procedure which yields a bilinearized system of dimension $n = 41760$. The dimension of the reduced system is set to $r = 5$. The results are shown in Figure 5, where the implicit Euler method is used to get the time domain representations of the original and the reduced-order systems. The error contours between the original and the bilinearized systems, and between the original and the reduced-order systems, are also given in Figure 6.

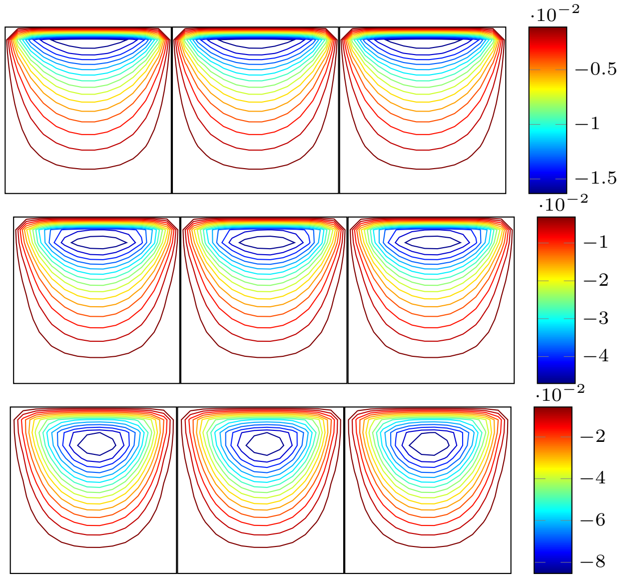

Fig. 5. Stream function contour generated with full-order system (left side), bilinearized system (middle) and reduced-order system (right side) at $t = 0.01s, 0.05s$ and steady state, respectively (from top to bottom).

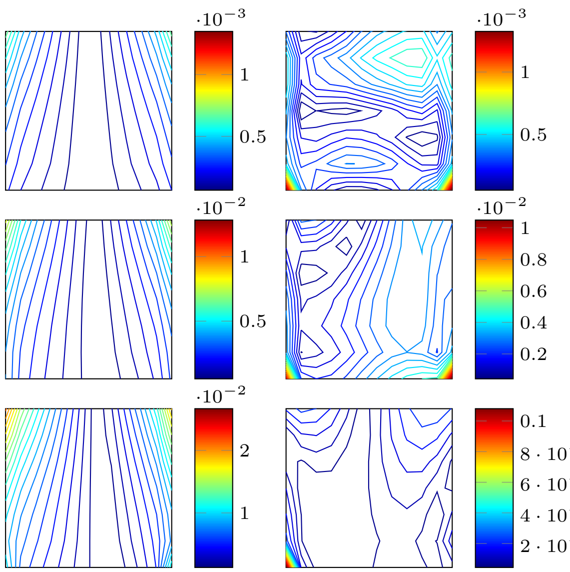

Fig. 6. The relative errors between the stream function contours generated from the bilinearized system (left) and reduced-order system (right) at $t = 0.01s, 0.05s$ and steady state, respectively (from top to bottom).

Clearly the reduced model has captured most of the dynamics of the quadratic-bilinear system very well.

V. Conclusions

In this paper, we have shown the use of Carleman bilinearization for quadratic-bilinear descriptor systems to obtain an approximate bilinear descriptor system. Also we showed that the bilinearization process preserves the index-1 (nilpotency index) property of the linear matrix pencil part of the bilinearized system. Assuming that the $\mathcal{H}_2$ norm of the system exists, we pointed out that the recently developed B-IRKA method can be used to identify a locally $\mathcal{H}_2$-optimal reduced bilinear system.

As further research, it would be interesting to extend the B-IRKA approach for general descriptor systems. In this regard, one possibility would be to identify a strictly proper and polynomial part of the original bilinear system and utilize B-IRKA only on the strictly proper part as is done in the linear case [18]. Another issue is to see how a higher nilpotency index of the matrix pencil corresponding to the nonlinear descriptor systems propagates in the bilinearized system obtained from Carleman bilinearization.

References

[1] W. J. Rugh, *Nonlinear System Theory*. Baltimore, MD: The Johns Hopkins University Press, 1981.

[2] P. Benner and T. Breiten, "Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 859–885, 2012.

[3] ——, "Two-sided moment matching methods for nonlinear model reduction," *SIAM Journal on Scientific Computing*, vol. 37, no. 2, pp. B239–B260, 2015.

[4] 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*, vol. 30, no. 9, pp. 1307–1320, 2011.

[5] K. Kunisch and S. Volkwein, "Proper orthogonal decomposition for optimality systems," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 42, no. 1, pp. 1–23, 2008.

[6] ——, "Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition," *Journal of Optimization Theory and Applications*, vol. 102, no. 2, pp. 345–371, 1999.

[7] T. Bui-Thanh, M. Damodaran, and K. Willcox, "Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition," *AIAA Journal*, vol. 42, no. 8, pp. 1505–1516, 2004.

[8] S. Chaturantabut and D. C. Sorensen, "Nonlinear model reduction via discrete empirical interpolation," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2737–2764, 2010.

[9] B. Haasdonk and M. Ohlberger, "Reduced basis method for finite volume approximations of parametrized linear evolution equations," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 42, no. 2, pp. 277–302, 2008.

[10] M. J. Rewieński, "A trajectory piecewise-linear approach to model order reduction of nonlinear dynamical systems," PhD Thesis, Massachusetts Institute of Technology, 2003.

[11] P. Benner and T. Breiten, "Krylov-subspace based model reduction of nonlinear circuit models using bilinear and quadratic-linear approximations," in *Progress in Industrial Mathematics at ECMI 2010, Mathematics in Industry*, vol. 17. Springer-Verlag, Berlin, 2012, pp. 153–159.

[12] G. W. Stewart and J. Sun, *Matrix perturbation theory*. New York: Academic press, 1990.

[13] Z. Al Zhour and A. Kilicman, "Matrix equalities and inequalities involving Khatri-Rao and Tracy-Singh sums," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 7, no. 1, pp. 496–513, 2006.

[14] S. Liu, "Several inequalities involving Khatri–Rao products of positive semidefinite matrices," *Linear Algebra and Its Applications*, vol. 354, no. 1, pp. 175–186, 2002.

[15] A. Bruns and P. Benner, "Parametric model order reduction of thermal models using the bilinear interpolatory rational Krylov algorithm," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 21, no. 2, pp. 103–129, 2015.

[16] L. Zhang and J. Lam, "On H_2 model reduction of bilinear systems," *Automatica*, vol. 38, no. 2, pp. 205–216, 2002.

[17] T. Breiten and T. Damm, "Krylov subspace methods for model order reduction of bilinear control systems," *Systems & Control Letters*, vol. 59, no. 10, pp. 443–450, 2010.

[18] S. Gugercin, T. Stykel, and S. Wyatt, "Model reduction of descriptor systems by interpolatory projection methods," *SIAM Journal on Scientific Computing*, vol. 35, no. 5, pp. B1010–B1033, 2013.

[19] W. Cazemier, R. Verstappen, and A. Veldman, "Proper orthogonal decomposition and low-dimensional models for driven cavity flows," *Physics of Fluids*, vol. 10, no. 7, pp. 1685–1699, 1998.

[20] F. Terragni, E. Valero, and J. Vega, "Local POD plus Galerkin projection in the unsteady lid-driven cavity problem," *SIAM Journal on Scientific Computing*, vol. 33, no. 6, pp. 3538–3561, 2011.