Advances in Computational Mathematics (2024) 50:42
https://doi.org/10.1007/s10444-024-10133-8

# Dominant subspaces of high-fidelity polynomial structured parametric dynamical systems and model reduction

**Pawan Goyal**[1] · **Igor Pontes Duff**[1] · **Peter Benner**[1,2]

Received: 24 January 2023 / Accepted: 2 April 2024


## Abstract

In this work, we investigate a model order reduction scheme for high-fidelity nonlinear structured parametric dynamical systems. More specifically, we consider a class of nonlinear dynamical systems whose nonlinear terms are polynomial functions, and the linear part corresponds to a linear structured model, such as second-order, time-delay, or fractional-order systems. Our approach relies on the Volterra series representation of these dynamical systems. Using this representation, we identify the kernels and, thus, the generalized multivariate transfer functions associated with these systems. Consequently, we present results allowing the construction of reduced-order models whose generalized transfer functions interpolate these of the original system at predefined frequency points. For efficient calculations, we also need the concept of a symmetric Kronecker product representation of a tensor and derive particular properties of them. Moreover, we propose an algorithm that extracts dominant subspaces from the prescribed interpolation conditions. This allows the construction of reduced-order models that preserve the structure. We also extend these results to parametric systems and a special case (delay in input/output). We demonstrate the efficiency of the proposed method by means of various numerical benchmarks.

**Keywords** Model order reduction · Interpolation · Polynomial dynamical systems · Parametric systems · Structured systems · Tensor computation

---

Communicated by: Tobias Breiten

---

✉ Pawan Goyal
goyalp@mpi-magdeburg.mpg.de

Igor Pontes Duff
pontes@mpi-magdeburg.mpg.de

Peter Benner
benner@mpi-magdeburg.mpg.de; peter.benner@ovgu.de

1 Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany

2 Otto von Guericke University, Faculty of Mathematics, Universitätsplatz 2, 39106 Magdeburg, Germany

Published online: 03 May 2024

**Mathematics Subject Classification (2010)** 15A69 · 34C20 · 41A05 · 49M05 · 93A15 · 93C10 · 93C15

## 1 Introduction

Dynamical systems are the basic framework used for modeling, controlling, and analyzing a large variety of engineering processes. Due to the increasing use of dedicated computer-based modeling design software, numerical simulation is now used more frequently to understand the dynamics of a complex system and to shorten both development time and cost. However, the need for enhanced model accuracy inevitably leads to an increasing number of variables and resources, which entails a high computational cost. In this context, model order reduction (MOR) is a possible remedy for such complex simulations. Precisely, MOR aims to replace a complex high-fidelity model with a reduced-order model (ROM) that mimics a certain dynamical behavior of the original model and preserves its features. As a result, this alleviates the numerical burden and reduces the computational time.

Many MOR methods for (parametric) nonlinear systems are based on simulated data. This means that, for given inputs and parameters, *snapshots* of the state vector $\mathbf{x}$ are collected [18]. Then, a low-dimensional dominant subspace is determined by means of a singular value decomposition (SVD) of the matrix containing the collected snapshots as columns. Hence, a ROM is constructed via Galerkin projection. Among these methods, proper orthogonal decomposition is arguably the most favored method (see, e.g., [29] for more details). Additionally, for nonlinear systems, this approach is often combined with hyper-reduction methods, such as EIM [4], DEIM [24], and GNAT [21], allowing fast evaluation of nonlinear terms. Also, for parametric problems, reduced basis methods have been successfully applied to several nonlinear systems, see, e.g., [35]. Although these methods have been successful in several applications, they are *input-dependent*, i.e., the quality of the ROMs depends on the choices of input functions and parameters used to collect the snapshots. Hence, it may be harder to obtain a ROM independent of inputs, which are suitable, e.g., for control problems.

In this work, we focus on MOR methods that are *input-independent*, i.e., ROMs can approximate the high-fidelity model behavior for all admissible inputs. The reader is referred to [5] for an overview of input-independent MOR methods. These methods, broadly speaking, are divided into two classes: interpolation-based and balanced truncation approaches. In this work, we focus on the class of interpolation-based methods. For the class of nonlinear systems, interpolation-based MOR methods have been extended to certain classes of nonlinear systems, e.g., bilinear (see, e.g., [3, 20]), parametric bilinear (see, e.g., [37]), quadratic-bilinear (see, e.g., [1, 9, 27]) and, more recently, (parametric-)polynomial systems (see, e.g., [11, 12]).

Besides the nonlinearities, in many applications, dynamical systems possess a particular dynamical structure, e.g., second-order, time-delay, and fractional-order systems. We call these systems *structured*. There exist several MOR methods for structured systems, allowing to preserve such dynamical structures in a ROM. We refer to [23, 26, 36] for second-order systems and [32] for time-delay systems. Moreover, in [6], the authors propose a framework allowing interpolation-based MOR for a vast class of linear structured systems. This framework was extended to the class of parametric linear structured systems in [2], and a data-driven identification approach was proposed in [39]. Additionally, in [14], the authors have proposed an approach to determine the dominant subspaces of a given (parametric) linear structured model. It is worth noticing that a balanced truncation approach is proposed in [19] for this class of systems.

In this paper, we focus on MOR for structured systems with polynomial nonlinearities. To illustrate this class of systems, let us consider the bilinear time-delay system presented in [28], which is of the form:

$$\dot{\mathbf{x}}(t) - \mathbf{A}_1\mathbf{x}(t) - \mathbf{A}_2\mathbf{x}(t-\tau) = \mathbf{N}\mathbf{x}(t)\mathbf{u}(t) + \mathbf{B}\mathbf{u}(t), \tag{1a}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \tag{1b}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$ and $\mathbf{u}(t) \in \mathbb{R}^p$ are the state, input, and output vectors. The left-hand side of (1a) corresponds to the linear part of the dynamics, governed by a time-delay structure. On the right-hand side, the term $\mathbf{N}\mathbf{x}(t)\mathbf{u}(t)$ corresponds to a bilinear nonlinearity. Our main goal is to determine a ROM for the system (1) that preserves its structure as well, i.e., to search for a surrogate model of the form:

$$\dot{\widehat{\mathbf{x}}}(t) - \widehat{\mathbf{A}}_1\widehat{\mathbf{x}}(t) - \widehat{\mathbf{A}}_2\widehat{\mathbf{x}}(t-\tau) = \widehat{\mathbf{N}}\widehat{\mathbf{x}}(t)\mathbf{u}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \tag{2a}$$

$$\widehat{\mathbf{y}}(t) = \widehat{\mathbf{C}}\widehat{\mathbf{x}}(t), \tag{2b}$$

where $\widehat{\mathbf{x}}(t) \in \mathbb{R}^r$ is the reduced state vector and $r \ll n$. This ROM should have a similar input-output behavior as the original one for all admissible inputs. Note that interpolation-based MOR for (parametric-) bilinear structured systems has been proposed in [16, 17]. In this paper, we build upon the methodologies proposed in [15, 40] for structured bilinear and quadratic-bilinear systems, and in [12] for polynomial systems. Particularly, we focus on an interpolation-based MOR approach for structured systems with polynomial nonlinearities, which was not addressed before in the model reduction literature. To this aim, firstly, we derive a Volterra series expansion for this class of systems, which was not yet in the literature. Based on the Volterra kernels and the associated generalized transfer functions, we are able to state several new interpolatory results, in particular also for systems with parameter dependency. Moreover, to prove some of these results, we show that tensors appearing in the dynamical system representation can always have a matrix representation with particular symmetric properties. Additionally, by following the philosophy in [11, 12, 14], we propose an algorithm enabling us to determine the dominant subspace information via interpolation.

The remaining structure of the paper is as follows. In Section 2, we present the class of polynomial structured systems, and provide some results for symmetric tensor representations that play an important role in MOR. In Section 3, we discuss the Volterra series representation of polynomial structured systems. This representation allows us to identify the system kernels, thus enabling us to define the generalized transfer functions of the system. Based on these, we present results that yield ROMs whose generalized transfer functions interpolate those of the original model at predefined interpolation points. Then, in Section 4, these results are generalized to other classes, such as parametric and input-output delay systems. Moreover, in Section 5, we propose an algorithm to determine the dominant subspace information from a given large set of interpolation points. In Section 6, we illustrate the efficiency of the proposed algorithm by means of three benchmark problems and compare it with the state-of-the-art. We conclude the paper with a summary of our contributions and future perspectives.

We make use of the following notation in the paper:

- Matrices and vectors are denoted with bold symbols, e.g., $\mathbf{A}$, $\mathbf{B}$, $\mathbf{v}$, $\mathbf{w}$.
- The $i$th entry of a vector $\mathbf{q} \in \mathbb{R}^n$ is denoted by $\mathbf{q}_i$.
- The Kronecker product is denoted by '$\otimes$'.
- $\mathbf{I}_m$ is the identity matrix of size $m \times m$.
- $\mathbf{e}_i$ denotes the $i$th column of the identity matrix of appropriate size.
- $\mathcal{V}^{\textcircled{\xi}}$ is a short-hand notation for $\underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{\xi-\text{times}}$, where $\mathcal{V}$ is a vector/matrix.

## 2 Problem setting and tensor algebra

### 2.1 Problem setting

In this paper, we focus on structured dynamical systems with polynomial non-linearity. These systems can be written in the form:

$$(\mathcal{L}\mathbf{x})(t) = \mathcal{P}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{B}\mathbf{u}(t), \tag{3a}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \tag{3b}$$

with matrices $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$; the state, input, and output vectors are denoted by $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$ and $\mathbf{y}(t) \in \mathbb{R}^p$, respectively; $\mathcal{L}(\cdot)$ corresponds to a linear operator, while $\mathcal{P}(\cdot) : \mathbb{R}^{n+m} \to \mathbb{R}^n$ represents non-linear terms. Additionally, we assume the corresponding initial conditions for (3a) to be zero. We shall discuss its variants (e.g., the parametric version) in Section 4.

We consider linear operators $\mathcal{L}(\cdot)$ in the system (3), covering a large class of systems arising in various science and engineering applications, e.g., classical linear systems, second-order systems, time-delay systems, and integro-differential systems. We list some examples in Table 1. Furthermore, we assume that the non-linear function $\mathcal{P}(\cdot)$ corresponds to polynomial non-linearities as follows:

$$\mathcal{P}(\mathbf{x}(t), \mathbf{u}(t)) = \sum_{\xi=2}^{d} \mathbf{H}_\xi \mathbf{x}^{\textcircled{\xi}}(t) + \sum_{\eta=1}^{d-1} \mathbf{N}_\eta \left(\mathbf{u}(t) \otimes \mathbf{x}^{\textcircled{\eta}}(t)\right), \tag{4}$$

where $\mathbf{H}_\xi \in \mathbb{R}^{n\times n^\xi}$, $\xi \in \{2, \dots, d\}$, $\mathbf{N}_\eta \in \mathbb{R}^{n\times m\cdot n^\eta}$, $\eta \in \{1, \dots, d-1\}$. In this work, the system (3) is refereed to as *polynomial structured system*.

**Table 1** Examples of common linear operators in dynamical systems

| | Linear operator $\mathcal{L}\mathbf{x}(t)$ | Frequency-domain description $\mathcal{K}(s)\mathbf{X}(s)$ |
|---|---|---|
| first-order | $\mathbf{E}\dot{\mathbf{x}} - \mathbf{A}\mathbf{x}(t)$ | $(s\mathbf{E} - \mathbf{A})\,\mathbf{X}(s)$ |
| second-order | $\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t)$ | $\left(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K}\right)\mathbf{X}(s)$ |
| state delay | $\mathbf{E}\dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{A}_\tau\mathbf{x}(t-\tau)$ | $\left(s\mathbf{E} - \mathbf{A} - \mathrm{e}^{-\tau s}\mathbf{A}_\tau\right)\mathbf{X}(s)$ |
| fractional order | $\mathbf{E}\dot{\mathbf{x}}(t) - \frac{1}{\Gamma(\alpha)}\int_0^t s^{\alpha-1}\mathbf{A}x(t-s)ds$ | $\left(s\mathbf{E} - s^{-\alpha}\mathbf{A}\right)\mathbf{X}(s)$ |

In this paper, our aim is to construct ROMs of order $r$ that have the same structure as in (3):

$$\left(\widehat{\mathcal{L}}\widehat{\mathbf{x}}\right)(t) = \widehat{\mathcal{P}}(\widehat{\mathbf{x}}(t), \mathbf{u}(t)) + \widehat{\mathbf{B}}\mathbf{u}(t), \tag{5a}$$

$$\widehat{y}(t) = \widehat{\mathbf{C}}\widehat{x}(t), \tag{5b}$$

where $\widehat{\mathbf{B}} \in \mathbb{R}^{r\times m}$, $\widehat{\mathbf{C}} \in \mathbb{R}^{p\times r}$; the reduced state, input, and approximated output vectors are denoted by $\widehat{\mathbf{x}}(t) \in \mathbb{R}^r$, $\mathbf{u}(t) \in \mathbb{R}^m$ and $\widehat{\mathbf{y}}(t) \in \mathbb{R}^p$, respectively, with $r \ll n$, and $\widehat{\mathbf{y}}(t)$ approximates very well $\mathbf{y}(t)$ for all admissible inputs. Additionally, the reduced linear operator $\widehat{\mathcal{L}}$ has also the same structure as the linear operator $\mathcal{L}$. As an example, for a second order dynamical system, the original linear operator is represented by

$$\mathcal{L}(\mathbf{x}) = \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t),$$

where $\mathbf{M}, \mathbf{D}, \mathbf{K} \in \mathbb{R}^{n\times n}$. Hence, in order to preserve the system structure, the reduced linear operator has to possess the following form:

$$\widehat{\mathcal{L}}(\widehat{\mathbf{x}}) = \widehat{\mathbf{M}}\ddot{\widehat{\mathbf{x}}}(t) + \widehat{\mathbf{D}}\dot{\widehat{\mathbf{x}}}(t) + \widehat{\mathbf{K}}\widehat{\mathbf{x}}(t),$$

with $\widehat{\mathbf{M}}$, , $\widehat{\mathbf{D}}$, $\widehat{\mathbf{K}} \in \mathbb{R}^{r\times r}$. Also, the reduced non-linear function $\widehat{\mathcal{P}}$ has the same structure as the original non-linear term $\mathcal{P}$, i.e.,

$$\widehat{\mathcal{P}}(\widehat{\mathbf{x}}(t)) = \sum_{\xi=2}^{d} \widehat{\mathbf{H}}_\xi \widehat{\mathbf{x}}^{\textcircled{\scriptsize$\xi$}}(t) + \sum_{\eta=1}^{d-1} \widehat{\mathbf{N}}_\eta \left(\mathbf{u}(t) \otimes \widehat{\mathbf{x}}^{\textcircled{\scriptsize$\eta$}}(t)\right). \tag{6}$$

We aim at achieving this goal via a Petrov-Galerkin projection. This means, we require two matrices $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n\times r}$ such that the reduced operator $\widehat{\mathcal{L}}(\cdot)$ in (5) and matrices involved in determining the polynomial term in (6) can be given as follows:

$$\begin{aligned} \widehat{\mathcal{L}}(\cdot) &= \mathbf{W}^\top \mathcal{L}(\cdot)\mathbf{V}, \quad \widehat{\mathbf{H}}_\xi = \mathbf{W}^\top \mathbf{H}_\xi \mathbf{V}^{\textcircled{\scriptsize$\xi$}}, \quad \xi \in \{2, \ldots, d\}, \\ \widehat{\mathbf{N}}_\eta &= \mathbf{W}^\top \mathbf{N}_\eta \mathbf{V}^{\textcircled{\scriptsize$\eta$}}, \quad \eta \in \{1, \ldots, d-1\}. \end{aligned} \tag{7}$$

Clearly, the selection of the projection matrices $\mathbf{V}$ and $\mathbf{W}$ plays an important role in determining the desired ROMs. In this paper, we aim to determine these matrices such that the resulting ROM fulfills certain interpolation properties. We mention that for structured bilinear cases ($d = 1$), some interpolation-based results were developed in [17], which we generalize to the polynomial case in the next section.

### 2.2 Results on tensor algebra

In this subsection, we recall some tensor algebra concepts that will be useful later in the paper and discuss symmetric tensors. A motivation for that is that the matrices $\mathbf{H}_\xi$ and $\mathbf{N}_\eta$ appearing in (4) can be interpreted as unfoldings of higher-order tensors. The process of representing a tensor as a matrix is often referred to as *matricization*, see, e.g., [33, 34]. Typically, an Nth order tensor can be unfolded in N different ways, depending on the dimension along which the tensor is unfolded. So, we begin by recalling the definition of matricization.

**Definition 2.1 (e.g., [33])** Consider an Nth order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times\cdots\times n_\mathrm{N}}$. The mode-$n$ matricization of the tensor $\boldsymbol{\mathcal{X}}$, denoted by $\boldsymbol{X}_{(m)}$, is obtained by the following mapping:

$$\boldsymbol{X}_{(m)}(i_m, j) = \boldsymbol{\mathcal{X}}(i_1, \ldots, i_\mathrm{N}),$$

where $j = 1 + \sum_{k=1,k\neq m}^{\mathrm{N}} (i_k - 1) J_k$ with $J_k = \prod_{z=1,z\neq m}^{k-1} n_z$ and $i_m \in \{1, \ldots, n_m\}$.

Next, we recall a connection between the mode-$n$ matricization and Kronecker products from, e.g., [33]. For this, let the tensor-matrix product be:

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \cdots \times_N \mathbf{A}^{(\mathrm{N})},$$

where $\mathbf{A}^{(l)} \in \mathbb{R}^{J_l\times n_l}$, $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{J_1\times\cdots\times J_N}$, and $\times_i$ denotes the tensor contraction with respect to the $i$th dimension of the tensor. Then, the following relation between the unfolded tensors and Kronecker products holds ([33, Prop. 3.7]):

$$\mathbf{Y}_{(m)} = \mathbf{A}^{(m)}\mathbf{X}_{(m)} \left(\mathbf{A}^{(\mathrm{N})} \otimes \cdots \otimes \mathbf{A}^{(m+1)} \otimes \mathbf{A}^{(m-1)} \otimes \mathbf{A}^{(1)}\right)^\top, \quad m \in \{1, \ldots, \mathrm{N}\}. \tag{8}$$

Now, we discuss a special case—that is, if $\mathbf{A}^{(l)} = \mathbf{a}_l^\top$, where $\mathbf{a}_l$ is a column vector and $l \in \{1, \ldots, \mathrm{N}\}$. In this case, we observe that $\mathbf{Y}_{(m)}$ becomes a scalar, given as:

$$\mathbf{Y}_{(m)} = \mathbf{a}_m^\top\mathbf{X}_{(m)} \left(\mathbf{a}_\mathrm{N} \otimes \cdots \otimes \mathbf{a}_{m+1} \otimes \mathbf{a}_{m-1} \otimes \cdots \otimes \mathbf{a}_1\right), \quad m \in \{1, \ldots, \mathrm{N}\}. \tag{9}$$

In what follows, we provide a result in tensor calculus that is of particular interest for the setting considered in this paper.

Before we proceed further, we define some notations. Let $\mathcal{S}$ be a set $\{i_1, i_2, \ldots, i_n\}$, and denote the set of all permutations of $\mathcal{S}$ by $\mathcal{S}_\mathbf{i}$ and the number of elements in $\mathcal{S}_\mathbf{i}$ by $\alpha_\mathbf{i}$. For example, consider a set $\{1, 2, 3\}$. Then, its permutations are: $(1, 2, 3)$, $(2, 1, 3)$, $(1, 3, 2)$, $(2, 3, 1)$, $(3, 1, 2)$ and $(3, 2, 1)$, and the number of elements are six. Having said this, in the following, we provide a result on the symmetric representation of the tensors.

**Lemma 2.1** *Consider an* $(\mathrm{N}+1)$*st order tensor* $\mathcal{H} \in \mathbb{R}^{n\times\cdots\times n}$. *Let us consider a set* $\mathcal{S} \in \{i_1, \ldots, i_n\}$ *and denote the set containing all its permutations by* $\mathcal{S}_{\mathbf{i}}$ *and the number of elements in* $\mathcal{S}_{\mathbf{i}}$ *by* $\alpha_{\mathbf{i}}$. *Furthermore, let a tensor* $\widetilde{\mathcal{H}}$ *be defined such that the* $\omega := \left(i_1 + \sum_{l=2}^{\mathrm{N}} (i_l - 1)(n^{l-1})\right)$*th column of its mode-1 matricization (denoted by* $\widetilde{\mathbf{H}}_{(1)}$*) is given as follows:*

$$\widetilde{\mathbf{H}}_{(1)}(:, \omega) := \mathbf{H}_{(1)} \left( \sum_{(j_1,\ldots,j_n)\in\mathcal{S}_{\mathbf{i}}} \frac{1}{\alpha_{\mathbf{i}}} \left(\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2} \otimes \cdots \otimes \mathbf{e}_{j_n}\right) \right), \tag{10}$$

*where* $\mathbf{H}_{(1)}$ *is the mode-*1 *matricization of the tensor* $\mathcal{H}$. *Then, the following conditions are satisfied:*

*a).* $\mathbf{H}_{(1)}\,(\mathbf{x} \otimes \cdots \otimes \mathbf{x}) = \widetilde{\mathbf{H}}_{(1)}\,(\mathbf{x} \otimes \cdots \otimes \mathbf{x})$, *where* $\mathbf{x} \in \mathbb{R}^n$ *is a vector.*

*b).* $\widetilde{\mathbf{H}}_{(1)}\left(\mathbf{q}^{(1)} \otimes \cdots \otimes \mathbf{q}^{(\mathrm{N})}\right) = \widetilde{\mathbf{H}}_{(1)}\left(\widetilde{\mathbf{q}}^{(1)} \otimes \cdots \otimes \widetilde{\mathbf{q}}^{(\mathrm{N})}\right)$, *where* $\mathbf{q}^{(i)} \in \mathbb{R}$, $i \in \{1, \ldots, \mathrm{N}\}$, *and* $\left(\widetilde{\mathbf{q}}^{(1)}, \ldots, \widetilde{\mathbf{q}}^{(\mathrm{N})}\right)$ *belongs to the set of all permutations of* $\left\{\mathbf{q}^{(1)}, \ldots, \mathbf{q}^{(\mathrm{N})}\right\}$.

*c). Moreover, all mode-m matricizations of the tensor* $\widetilde{\mathcal{H}}$ *for* $m \geq 2$ *are the same, i.e.,* $\widetilde{\mathbf{H}}_{(2)} = \widetilde{\mathbf{H}}_{(3)} = \cdots = \widetilde{\mathbf{H}}_{(\mathrm{N})}$.

***Proof*** The proof is given in Appendix A.

For a better understanding, we illustrate Lemma 2.1 by an example. Consider a tensor $\mathcal{H}^{2\times2\times2\times2}$ such that its mode-1 matricization (denoted by $\mathbf{H}_{(1)}$) is given by

$$\mathbf{H}_{(1)} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & b_8 \end{bmatrix}. \tag{11}$$

Next, we write down explicitly the term $\mathbf{H}_{(1)}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$, where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top$– that is

$$\begin{bmatrix} a_1x_1^3 + a_2x_1^2x_2 + a_3x_1x_2x_1 + a_4x_1x_2^2 + a_5x_2x_1^2 + a_6x_2x_1x_2 + a_7x_2^2x_1 + a_8x_2^3 \\ b_1x_1^3 + b_2x_1^2x_2 + b_3x_1x_2x_1 + b_4x_1x_2^2 + b_5x_2x_1^2 + b_6x_2x_1x_2 + b_7x_2^2x_1 + b_8x_2^3 \end{bmatrix}.$$

Now, we define

$$\widetilde{a}_2 = \tfrac{(a_2+a_3+a_5)}{3}, \qquad \widetilde{a}_4 = \tfrac{(a_4+a_6+a_7)}{3} \qquad \widetilde{b}_2 = \tfrac{(b_2+b_3+b_5)}{3} \qquad \widetilde{b}_4 = \tfrac{(b_4+b_6+b_7)}{3}$$

Then, we can also write $\mathbf{H}_{(1)}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$ as:

$$\begin{bmatrix} a_1x_1^3 + \widetilde{a}_2x_1^2x_2 + \widetilde{a}_2x_1x_2x_1 + \widetilde{a}_4x_1x_2^2 + \widetilde{a}_2x_2x_1^2 + \widetilde{a}_4x_2x_1x_2 + \widetilde{a}_4x_2^2x_1 + a_8x_2^3 \\ b_1x_1^3 + \widetilde{b}_2x_1^2x_2 + \widetilde{b}_2x_1x_2x_1 + \widetilde{b}_4x_1x_2^2 + \widetilde{b}_4x_2x_1^2 + \widetilde{b}_4x_2x_1x_2 + \widetilde{b}_4x_2^2x_1 + b_8x_2^3 \end{bmatrix}.$$

Consequently, if we define a tensor $\widetilde{\mathcal{H}}$ such that its mode-1 matricization is given as

$$\widetilde{\mathbf{H}}_{(1)} = \begin{bmatrix} a_1 & \widetilde{a}_2 & \widetilde{a}_2 & \widetilde{a}_4 & \widetilde{a}_2 & \widetilde{a}_4 & \widetilde{a}_4 & a_8 \\ b_1 & \widetilde{b}_2 & \widetilde{b}_2 & \widetilde{b}_4 & \widetilde{b}_2 & \widetilde{b}_4 & \widetilde{b}_4 & b_8 \end{bmatrix}, \tag{12}$$

then $\widetilde{\mathbf{H}}_{(1)}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) = \mathbf{H}_{(1)}(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x})$. Moreover, it can also be observed that $\widetilde{\mathbf{H}}_{(1)}(\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}) = \widetilde{\mathbf{H}}_{(1)}(\mathbf{v} \otimes \mathbf{u} \otimes \mathbf{w}) = \cdots = \widetilde{\mathbf{H}}_{(1)}(\mathbf{w} \otimes \mathbf{v} \otimes \mathbf{u})$. Furthermore, if one aims at obtaining $\widetilde{\mathbf{H}}_{(1)}$ from Lemma 2.1, then we can write, e.g., its 2nd column as

$$\begin{aligned}\widetilde{\mathbf{H}}_{(1)}(:, 2) &= \mathbf{H}_{(1)}\left(\frac{1}{3}\left((\mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1) + (\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1) + (\mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2)\right)\right)\\ &= \frac{1}{3}\begin{bmatrix} a_2 + a_3 + a_5 \\ b_2 + b_3 + b_5 \end{bmatrix}.\end{aligned}$$

To sum up, according to Lemma 2.1, a tensor $\mathcal{H}$ can be symmetrized (denote the symmetrized tensor by $\widetilde{\mathcal{H}}$) without changing the quantity $\mathbf{H}_{(1)}(\mathbf{x} \otimes \cdots \otimes \mathbf{x})$ and the following commutation rule is also fulfilled:

$$\widetilde{\mathbf{H}}_{(1)}(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_n) = \widetilde{\mathbf{H}}_{(1)}(\widetilde{\mathbf{v}}_1 \otimes \cdots \otimes \widetilde{\mathbf{v}}_n), \tag{13}$$

for every permutation $(\widetilde{\mathbf{v}}_1, \ldots, \widetilde{\mathbf{v}}_n)$ of the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. This extends the discussion in [9] for 3rd order tensors to the general case. Therefore, in the rest of the paper, without loss of generality, we assume that all tensors associated with $\mathbf{H}_\xi$ and $\mathbf{N}_\eta$ are symmetric.

## 3 Volterra series and interpolation-based MOR

This section presents the Volterra series representation of nonlinear structured systems (3). For this, we extend the discussion on non-structured polynomial systems in [12] to the case of structured polynomial systems. We aim at identifying the kernels related to the system (3) that allow us to define generalized transfer functions. As a consequence, we intend to construct a ROM such that its generalized transfer functions interpolate these of the original system at pre-defined interpolation points. For simplicity, in this section, we assume the system to be single-input single-output (SISO), i.e., $m = p = 1$. We will also provide an extension to the multi–input multi–output case (MIMO).

### 3.1 Volterra series representation

First, let $\Phi$ be the fundamental solution associated with the linear operator $\mathcal{L}$. We detail the meaning of the fundamental solution in Appendix B. Consequently, the solution of an equation of the form:

$$(\mathcal{L}\mathbf{x})(t) = \mathbf{g}(t) \tag{14}$$

can be given as the convolution (using the fact that $\mathbf{x}(0) = 0$ by assumption)

$$\mathbf{x}(t) = \int_0^t \Phi(\sigma)\mathbf{g}(t - \sigma)d\sigma. \tag{15}$$

Now let us apply the convolution form (15) to the structured dynamical systems with polynomial nonlinearity in equation (3a), i.e., we choose $\mathbf{g}(t)$ as

$$\mathbf{g}(t) = \sum_{\xi=2}^{d} \mathbf{H}_\xi \mathbf{x}^{\textcircled{\xi}}(t) + \sum_{\eta=1}^{d-1} \mathbf{N}_\eta \left(\mathbf{u}(t) \otimes \mathbf{x}^{\textcircled{\eta}}(t)\right) + \mathbf{B}\mathbf{u}(t).$$

As a consequence from the linearity of the convolution operator, we obtain

$$\mathbf{x}(t) = \int_0^t \Phi(\sigma_1)\mathbf{B}\mathbf{u}(t_{\sigma_1})d\sigma_1 + \sum_{\xi=2}^{d} \int_0^t \Phi(\sigma_1)\mathbf{H}_\xi \mathbf{x}^{\textcircled{\xi}}(t_{\sigma_1})d\sigma_1 + \sum_{\eta=1}^{d-1} \int_0^t \Phi(\sigma_1)\mathbf{N}_\eta\, \mathbf{x}^{\textcircled{\eta}}(t_{\sigma_1})\, \mathbf{u}(t_{\sigma_1})d\sigma_1, \quad (16)$$

where $t_{\sigma_1} := t - \sigma_1$. Moreover, we can determine the expression for $\mathbf{x}(t_{\sigma_1})$ using the above equation – that is,

$$\mathbf{x}(t_{\sigma_1}) = \int_0^{t_{\sigma_1}} \Phi(\sigma_2)\mathbf{B}\mathbf{u}(t_{\sigma_1}-\sigma_2)d\sigma_2 + \sum_{\xi=2}^{d} \int_0^{t_{\sigma_1}} \Phi(\sigma_2)\mathbf{H}_\xi \mathbf{x}^{\textcircled{\xi}}(t_{\sigma_1}-\sigma_2)d\sigma_2 + \sum_{\eta=1}^{d-1} \int_0^{t_{\sigma_1}} \Phi(\sigma_2)\mathbf{N}_\eta\, \mathbf{x}^{\textcircled{\eta}}(t_{\sigma_1}-\sigma_2)\, \mathbf{u}(t_{\sigma_1}-\sigma_2)d\sigma_2. \quad (17)$$

We utilize the expression for $\mathbf{x}(t_{\sigma_1})$ in (16), multiplied by the matrix $\mathbf{C}$. Hence, we obtain

$$\begin{aligned}
\mathbf{y}(t) = &\int_0^t \mathbf{C}\Phi(\sigma_1)\mathbf{B}\mathbf{u}(t_{\sigma_1})d\sigma_1 \\
&+ \sum_{\xi=2}^{d} \int_0^t \underbrace{\int_0^{t_{\sigma_1}} \cdots \int_0^{t_{\sigma_1}}}_{\xi-\text{times}} \mathbf{C}\Phi(\sigma_1)\mathbf{H}_\xi \left(\Phi(\sigma_2)\mathbf{B} \otimes \cdots \otimes \Phi(\sigma_{\xi+1})\mathbf{B}\right) \\
&\times \left(\mathbf{u}(t_{\sigma_1} - \sigma_2) \cdots \mathbf{u}(t_{\sigma_1} - \sigma_{\xi+1})\right) d\sigma_1 d\sigma_2 \cdots d\sigma_{\xi+1} \\
&+ \sum_{\eta=1}^{d-1} \int_0^t \underbrace{\int_0^{t_{\sigma_1}} \cdots \int_0^{t_{\sigma_1}}}_{\eta-\text{times}} \mathbf{C}\Phi(\sigma_1)\mathbf{N}_\eta \left(\Phi(\sigma_2)\mathbf{B} \otimes \cdots \otimes \Phi(\sigma_{\eta+1})\mathbf{B}\right) \\
&\times \left(\mathbf{u}(t_{\sigma_1})\mathbf{u}(t_{\sigma_1} - \sigma_2) \cdots \mathbf{u}(t_{\sigma_1} - \sigma_{\eta+1})\right) d\sigma_1 d\sigma_2 \cdots d\sigma_{\eta+1} + \cdots,
\end{aligned}$$

with the assumption that the above series convergences. As a result, we have an analytical representation of the output $\mathbf{y}(t)$ as an infinite sum of multivariable convolutions. But in the above expression for $\mathbf{y}(t)$, we only note down the leading few *kernels* of the Volterra series for the system (3). In this paper, we focus only on those:

$$f_{\mathrm{L}}(\sigma_1) := \mathbf{C}\Phi(\sigma_1)\mathbf{B}, \quad (19a)$$

$$f_{\mathrm{H}}^{(\xi)}(\sigma_1,\ldots,\sigma_{\xi+1}) := \mathbf{C}\Phi(\sigma_1)\mathbf{H}_\xi\left(\Phi(\sigma_2)\mathbf{B}\otimes\cdots\otimes\Phi(\sigma_{\xi+1})\mathbf{B}\right), \tag{19b}$$

$$f_{\mathrm{N}}^{(\eta)}(\sigma_1,\ldots,\sigma_{\eta+1}) := \mathbf{C}\Phi(\sigma_1)\mathbf{N}_\eta\left(\Phi(\sigma_2)\mathbf{B}\otimes\cdots\otimes\Phi(\sigma_{\eta+1})\mathbf{B}\right), \tag{19c}$$

with $\xi \in \{2,\ldots,d\}$ and $\eta \in \{1,\ldots,d-1\}$. Furthermore, by taking the multivariate Laplace transform (see, e.g., [38]), we get the frequency-domain representations of the kernels as follows:

$$\mathbf{F}_{\mathrm{L}}(s_1) := \mathbf{C}\mathcal{K}^{-1}(s_1)\mathbf{B}, \tag{20a}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(s_1,\ldots,s_{\xi+1}) := \mathbf{C}\mathcal{K}^{-1}(s_{\xi+1})\mathbf{H}_\xi\left(\mathcal{K}^{-1}(s_\xi)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(s_1)\mathbf{B}\right), \tag{20b}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(s_1,\ldots,s_{\eta+1}) := \mathbf{C}\mathcal{K}^{-1}(s_{\eta+1})\mathbf{N}_\eta\left(\mathcal{K}^{-1}(s_\eta)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(s_1)\mathbf{B}\right), \tag{20c}$$

for $\xi \in \{2,\ldots,d\}$ and $\eta \in \{1,\ldots,d-1\}$, where $\mathcal{K}^{-1}(s)$ is the Laplace transform of the fundamental solution $\Phi$. We have listed in Table 1 some examples of the structure of $\mathcal{K}(s)$ for certain types of structured systems. In this paper, we refer to the functions in (20) as the *multivariate transfer functions* associated with the polynomial structured system (3). It is worth mentioning that one could have derived more Volterra kernels than those presented in (19) by following the same approach presented in this paper. However, for the brevity of this work, we decided not to include this here. Additionally, in practice, the leading kernels in (19) are those used in model reduction schemes since they typically contain the most dominant information for weakly nonlinear systems.

### 3.2 Interpolation-based MOR

In this subsection, we present the construction of projection matrices $\mathbf{V}$ and $\mathbf{W}$, yielding ROMs (6) such that the generalized transfer functions of the original model and ROM match at pre-defined interpolation points. The results presented here extend those presented in [12] from polynomial systems with $\mathcal{K}(s) = s\mathbf{E} - \mathbf{A}$ and those derived in [40] from quadratic-bilinear structured systems to the class of structured polynomial systems.

**Theorem 3.1** *Let a SISO polynomial structured system be given as in* (3)*. Assume $\sigma_i$ and $\mu_i$, $i \in \{1,\ldots,\widetilde{r}\}$, to be interpolation points such that $\mathcal{K}(s)$ is invertible for all $s = \{\sigma_i, \mu_i\}$, $i \in \{1,\ldots,\widetilde{r}\}$. Moreover, we define the projection matrices* $\mathbf{V}$ *and* $\mathbf{W}$ *as follows:*

$$\mathcal{V}_{\mathrm{L}} = \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_1)\mathbf{B},\ldots,\mathcal{K}^{-1}(\sigma_{\widetilde{r}})\mathbf{B}\right),$$

$$\mathcal{V}_{\mathrm{N}} = \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{N}_\eta\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\right)\right),$$

$$\mathcal{V}_{\mathrm{H}} = \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{H}_\xi\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\right)\right),$$

$$\mathcal{W}_{\mathrm{L}} = \operatorname{range}\left(\mathcal{K}^{-\top}(\mu_1)\mathbf{C}^\top,\ldots,\mathcal{K}^{-\top}(\mu_{\widetilde{r}})\mathbf{C}^\top\right),$$

$$\mathcal{W}_{\mathrm{N}} = \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-\top}(\sigma_i)\left(\mathbf{N}_\eta\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\mathcal{K}^{-\top}(\mu_i)\mathbf{C}^\top\right)\right),$$

$$\mathcal{W}_{\mathrm{H}} = \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\Big(\mathcal{K}^{-\top}(\sigma_i)\left(\mathbf{H}_\xi\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-\top}(\mu_i)\mathbf{C}^\top\right)\Big),$$

$$\operatorname{range}(\mathbf{V}) = \mathcal{V}_{\mathrm{L}} + \mathcal{V}_{\mathrm{N}} + \mathcal{V}_{\mathrm{H}},$$

$$\operatorname{range}(\mathbf{W}) = \mathcal{W}_{\mathrm{L}} + \mathcal{W}_{\mathrm{N}} + \mathcal{W}_{\mathrm{H}},$$

*where* $\left(\mathbf{H}_\xi\right)_{(2)} \in \mathbb{R}^{n\times n^\xi}$ *and* $\left(\mathbf{N}_\eta\right)_{(2)} \in \mathbb{R}^{n\times m\cdot n^\xi}$ *are, respectively, the mode-2 matricizations of the* $(\xi+1)$*-way symmetric tensor* $\boldsymbol{\mathcal{H}}_\xi \in \mathbb{R}^{n\times\cdots\times n}$ *and* $(\eta+2)$*-way symmetric tensor* $\boldsymbol{\mathcal{N}}_\eta \in \mathbb{R}^{n\times\cdots\times n}$ *whose mode-1 matricizations are* $\mathbf{H}_\xi$ *and* $\mathbf{N}_\eta$*, respectively. Assume* $\mathbf{V}$ *and* $\mathbf{W}$ *are of full column rank and* $(\mathbf{F}_{\mathrm{L}}, \mathbf{F}_{\mathrm{N}}, \mathbf{F}_{\mathrm{H}})$ *and* $(\widehat{\mathbf{F}}_{\mathrm{L}}, \widehat{\mathbf{F}}_{\mathrm{N}}, \widehat{\mathbf{F}}_{\mathrm{H}})$ *are the multivariate transfer functions of the original model and ROM, respectively. If the ROM is constructed by Petrov-Galerkin projections as in (7) using the matrices* $\mathbf{V}$ *and* $\mathbf{W}$*, then the ROM satisfies the following interpolation conditions:*

$$\mathbf{F}_{\mathrm{L}}(\sigma_i) = \widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i), \tag{22a}$$

$$\mathbf{F}_{\mathrm{L}}(\mu_i) = \widehat{\mathbf{F}}_{\mathrm{L}}(\mu_i), \tag{22b}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i) = \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i), \tag{22c}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mu_i) = \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mu_i), \tag{22d}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i) = \widehat{\mathbf{F}}_{H}^{(\xi)}(\sigma_i,\ldots,\sigma_i), \tag{22e}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mu_i) = \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mu_i), \tag{22f}$$

*provided the reduced matrix* $\widehat{\mathcal{K}}(s) := \mathbf{W}^\top\mathcal{K}(s)\mathbf{V}$ *is non-singular for* $s = \{\sigma_i, \mu_i\}$*,* $i \in \{1,\ldots,\widetilde{r}\}$.

***Proof*** The theorem can be proven along the same lines as done in [12], where a special case with $\mathcal{K}(s) = s\mathbf{E} - \mathbf{A}$ is considered. However, we here provide the idea by providing the proof of the relation (22c) only.

Firstly notice that, from (7), we have $\widehat{\mathcal{K}}(s) = \mathbf{W}^\top\mathcal{K}(s)\mathbf{V}$. The interpolation conditions in (22a) and (22b) follow directly from the structured linear case, see [6]. Additionally, one can easily prove the following relations:

$$\mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}} = \mathcal{K}^{-1}(\sigma_i)\mathbf{B}, \qquad \sigma_i \in \{\sigma_1,\ldots,\sigma_{\widetilde{r}}\}, \tag{23a}$$

$$\widehat{\mathbf{C}}\widehat{\mathcal{K}}^{-1}(\mu_i)\mathbf{W}^\top = \mathbf{C}\mathcal{K}^{-1}(\mu_i), \qquad \mu_i \in \{\mu_1,\ldots,\mu_{\widetilde{r}}\}. \tag{23b}$$

Next, let us prove the interpolation condition in (22c). We begin with

$$\begin{aligned}
&\mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{N}}_\eta\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\otimes\cdots\otimes\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\right)\\
&\qquad = \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\mathbf{W}^\top\mathbf{N}_\eta\mathbf{V}^{\textcircled{\eta}}\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\otimes\cdots\otimes\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\right)\\
&\qquad = \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\mathbf{W}^\top\mathbf{N}_\eta\left(\mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\otimes\cdots\otimes\mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\mathbf{W}^\top\mathbf{N}_\eta\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\text{(using 23}a\text{)} \\
&= \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\mathbf{W}^\top\mathcal{K}(\sigma_i)\underbrace{\mathcal{K}^{-1}(\sigma_i)\mathbf{N}_\eta\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\right)}_{\in\text{range}\,\mathbf{V}(\therefore\,=:\mathbf{Vz})} \\
&= \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\mathbf{W}^\top\mathcal{K}(\sigma_i)\mathbf{Vz}, \\
&= \mathbf{V}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathcal{K}}(\sigma_i)\mathbf{z} = \mathbf{Vz},
\end{aligned} \tag{24a}
$$

where $\mathbf{z}$ is a vector such that $\mathbf{Vz} = \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i)\mathbf{B}$. Hence, pre-multiplying with $\mathbf{C}$ yields the interpolation condition (22c). The proof of the interpolation condition (22e) follows in a similar way. Moreover, the other interpolation conditions can be proven similarly. □

Theorem 3.1 allows to construct a ROM satisfying interpolation conditions of the multivariate system transfer functions (20). Additionally, the result, stated in the following remark, also holds for interpolation conditions involving different $\sigma_i$.

**Remark 3.1** Other combinations of $\sigma_i$ and $\mu_i$ are possible in the interpolation conditions if the subspaces are defined accordingly. As an example, under the assumptions of Theorem 3.1 and following the steps of the proof of Theorem 3.1, one can show the following result. Let

$$
\begin{aligned}
\mathcal{V}_\text{L} &= \text{range}\left(\mathcal{K}^{-1}(\sigma_1)\mathbf{B},\ldots,\mathcal{K}^{-1}(\sigma_{\widetilde{r}})\mathbf{B}\right), \\
\widetilde{\mathcal{V}}_\text{N} &= \text{range}\left(\mathcal{K}^{-1}(\sigma_{\eta+1})\mathbf{N}_\eta\left(\mathcal{K}^{-1}(\sigma_\eta)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_1)\mathbf{B}\right)\right), \\
\widetilde{\mathcal{V}}_\text{H} &= \text{range}\left(\mathcal{K}^{-1}(\sigma_{\eta+1})\mathbf{H}_\xi\left(\mathcal{K}^{-1}(\sigma_\eta)\mathbf{B}\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_1)\mathbf{B}\right)\right), \quad \text{and} \\
\text{range}(\mathbf{V}) &= \mathcal{V}_\text{L} + \widetilde{\mathcal{V}}_\text{N} + \widetilde{\mathcal{V}}_\text{H},
\end{aligned}
$$

and let $\mathbf{W}$ be such that $\widehat{\mathcal{K}}(s) := \mathbf{W}^\top\mathcal{K}(s)\mathbf{V}$ is non-singular for $s = \{\sigma_1,\ldots,\sigma_{\widetilde{r}}\}$. Then the ROM obtained by Petrov-Galerkin projection using the matrices $\mathbf{V}$ and $\mathbf{W}$ satisfies the following interpolation conditions in addition to (22a):

$$
\mathbf{F}_\text{N}^{(\eta)}(\sigma_1,\ldots,\sigma_{\eta+1}) = \widehat{\mathbf{F}}_\text{N}^{(\eta)}(\sigma_1,\ldots,\sigma_{\eta+1}), \tag{25}
$$

$$
\mathbf{F}_\text{H}^{(\xi)}(\sigma_1,\ldots,\sigma_{\eta+1}) = \widehat{\mathbf{F}}_\text{H}^{(\xi)}(\sigma_1,\ldots,\sigma_{\eta+1}). \tag{26}
$$

Furthermore, we can also choose an appropriate matrix $\mathbf{W}$ as shown in Theorem 3.1, yielding a ROM that would satisfy an increased number of interpolation conditions, but for the brevity of the paper, we avoid including these results in detail.

In what follows, we show that if the right and left interpolation points are equal, i.e., $\sigma_i = \mu_i$, $i \in \{1,\ldots,\widetilde{r}\}$, then Hermite interpolation conditions are also satisfied.

**Theorem 3.2** *Under the hypothesis of Theorem 3.1, assume that* $\sigma_i = \mu_i$, *for* $i \in \{1, \ldots, \widetilde{r}\}$. *Then,*

$$\frac{\partial}{\partial s_1}\mathbf{F}_{\mathrm{L}}(\sigma_i) = \frac{\partial}{\partial s_1}\widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i), \tag{27a}$$

$$\frac{\partial}{\partial s_j}\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i) = \frac{\partial}{\partial s_j}\widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i), \quad \textit{for } j \in \{1, \ldots, \eta+1\}, \tag{27b}$$

$$\frac{\partial}{\partial s_k}\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i, \ldots, \sigma_i) = \frac{\partial}{\partial s_k}\widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i, \ldots, \sigma_i), \quad \textit{for } k \in \{1, \ldots, \xi+1\}. \tag{27c}$$

***Proof*** First, note that the result (27a) is well-known in the literature, see, e.g., [6, Thm. 1]. Thus, we focus here on proving the interpolation condition (27b). Observe that

$$\frac{\partial}{\partial s}\mathcal{K}^{-1}(s) = -\mathcal{K}^{-1}(s)\left(\frac{\partial}{\partial s}\mathcal{K}(s)\right)\mathcal{K}^{-1}(s).$$

Hence,

$$\begin{aligned}
&\frac{\partial}{\partial s_1}\widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i) \\
&\quad = \widehat{\mathbf{C}}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{N}}_\eta\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}} \otimes \cdots \otimes \frac{\partial}{\partial s_1}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\right) \\
&\quad = -\widehat{\mathbf{C}}\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{N}}_\eta\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}} \otimes \cdots \otimes \widehat{\mathcal{K}}^{-1}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)\widehat{\mathbf{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}}\right) \\
&\qquad\qquad \text{(using (9))} \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\widehat{\mathbf{N}}_\eta\right)_{(2)}\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}} \otimes \cdots \otimes \widehat{\mathcal{K}}^{-T}(\sigma_i)\widehat{\mathbf{C}}^\top\right) \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\mathbf{V}^\top\left(\mathbf{N}_\eta\right)_{(2)}(\mathbf{V} \otimes \cdots \otimes \mathbf{V} \otimes \mathbf{W}) \\
&\qquad\qquad \times\left(\widehat{\mathcal{K}}^{-1}(\sigma_i)\widehat{\mathbf{B}} \otimes \cdots \otimes \widehat{\mathcal{K}}^{-\top}(\sigma_i)\widehat{\mathbf{C}}^\top\right) \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\mathbf{V}^\top\left(\mathbf{N}_\eta\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B} \otimes \cdots \otimes \mathcal{K}^{-\top}(\sigma_i)\mathbf{C}^\top\right) \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\mathbf{V}^\top\mathcal{K}(\sigma_i)^\top \\
&\qquad\qquad \times\underbrace{\mathcal{K}(\sigma_i)^{-\top}\left(\mathbf{N}_\eta\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B} \otimes \cdots \otimes \mathcal{K}^{-\top}(\sigma_i)\mathbf{C}^\top\right)}_{\in \operatorname{range} \mathbf{W}(\therefore =:\mathbf{W}z)} \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\mathbf{V}\mathcal{K}(\sigma_i)^\top\mathbf{W}z \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\left(\frac{\partial}{\partial s_1}\widehat{\mathcal{K}}(\sigma_i)\right)^\top z \\
&\quad = -\widehat{\mathbf{B}}^\top\widehat{\mathcal{K}}^{-\top}(\sigma_i)\mathbf{V}^\top\left(\frac{\partial}{\partial s_1}\mathcal{K}(\sigma_i)\right)^\top\mathbf{W}z
\end{aligned}$$

$$
\begin{aligned}
&= -\mathbf{B}^\top \mathcal{K}^{-\top}(\sigma_i) \left( \frac{\partial}{\partial s_1} \mathcal{K}(\sigma_i) \right)^\top \mathbf{W} z \\
&= -\mathbf{B}^\top \mathcal{K}^{-\top}(\sigma_i) \left( \frac{\partial}{\partial s_1} \mathcal{K}(\sigma_i) \right)^\top \mathcal{K}(\sigma_i)^{-\top} \left(\mathbf{N}_\eta\right)_{(2)} \left( \mathcal{K}^{-1}(\sigma_i)\mathbf{B} \otimes \cdots \otimes \mathcal{K}^{-\top}(\sigma_i)\mathbf{C}^\top \right) \\
&= \mathbf{C}\mathcal{K}^{-1}(\sigma_i)\mathbf{N}_\eta \left( \mathcal{K}^{-1}(\sigma_i)\mathbf{B} \otimes \cdots \otimes \frac{\partial}{\partial s_1} \mathcal{K}^{-1}(\sigma_i)\mathbf{B} \right) \\
&= \frac{\partial}{\partial s_1} \mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i).
\end{aligned}
$$

Furthermore, due to the symmetric tensors and Lemma 2.1, we have

$$
\frac{\partial}{\partial s_1} \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i) = \frac{\partial}{\partial s_j} \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i), \quad j \in \{2, \ldots, \eta\}.
$$

Moreover, along the same lines as above, we can also show that

$$
\frac{d}{ds_{\eta+1}} \mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i) = \frac{d}{ds_{\eta+1}} \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i). \tag{29}
$$

Hence, the relations (27b) are proven, and similarly, the relations in (27c) can also be proven.

From Theorems 3.1 and 3.2, we are able to determine Petrov-Galerkin matrices $\mathbf{V}$ and $\mathbf{W}$, allowing to construct a desired interpolatory ROM. These results are derived up to now for the SISO case. The results generalize [17, Thms. 5 and 6]. Moreover, they can be easily generalized for the MIMO case using the idea of the so-called tangential interpolation. We discuss the MIMO result in Appendix C.

## 4 Extension to parametric and special cases

In this section, we discuss extensions of the result presented in the previous section to parametric and some special cases. We begin with the parametric case.

### 4.1 Parametric case

Here, we consider parametric systems of the form:

$$
(\mathcal{L}(\mathbf{p})\mathbf{x})\,(t, \mathbf{p}) = \mathcal{P}(\mathbf{x}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p}) + \mathbf{B}(\mathbf{p})\mathbf{u}(t), \tag{30a}
$$

$$
\mathbf{y}(t, \mathbf{p}) = \mathbf{C}(\mathbf{p})\mathbf{x}(t, \mathbf{p}), \tag{30b}
$$

where $\mathbf{p} \in \Omega \subset \mathbb{R}^q$ contains the system parameters; $\mathcal{L}(\mathbf{p})$ is a parameterized linear operator; $\mathbf{B}(\mathbf{p}) \in \mathbb{R}^{n \times m}$, $\mathbf{C}(\mathbf{p}) \in \mathbb{R}^{q \times n}$ are parameter-dependent matrices, and the nonlinear term $\mathcal{P}(\mathbf{x}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p})$ takes the form:

$$
\mathcal{P}(\mathbf{x}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p}) = \sum_{\xi=2}^{d} \mathbf{H}_\xi(\mathbf{p}) x^{\textcircled{\xi}}(t, \mathbf{p}) + \sum_{\eta=1}^{d-1} \mathbf{N}_\eta(\mathbf{p}) \left( \mathbf{u}(t) \otimes \mathbf{x}^{\textcircled{\eta}}(t, \mathbf{p}) \right). \tag{31}
$$

The parametric case results can be obtained by following the same lines as for the non-parametric case and discussion in [11] for non-structured systems; thus, here we only briefly sketch the ideas. Here again, for simplicity, we present the results for the SISO case. In the MIMO case, we can use the idea of tangential interpolation as discussed in Appendix C.

Similar to the non-parametric case, we can derive generalized transfer functions for the parametric system, which are given as follows:

$$\mathbf{F}_{\mathrm{L}}(s_1,\mathbf{p}) := \mathbf{C}(\mathbf{p})\mathcal{K}^{-1}(s_1,\mathbf{p})\mathbf{B}(\mathbf{p}), \tag{32a}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(s_1,\ldots,s_{\xi+1},\mathbf{p}) := \mathbf{C}(\mathbf{p})\mathcal{K}^{-1}(s_{\xi+1},\mathbf{p})\mathbf{H}_{\xi}(\mathbf{p})\left(\mathcal{K}^{-1}(s_{\xi},\mathbf{p})\mathbf{B}(\mathbf{p})\otimes\cdots\otimes\mathcal{K}^{-1}(s_1,\mathbf{p})\mathbf{B}(\mathbf{p})\right), \tag{32b}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(s_1,\ldots,s_{\eta+1},\mathbf{p}) := \mathbf{C}(\mathbf{p})\mathcal{K}^{-1}(s_{\eta+1},\mathbf{p})\mathbf{N}_{\eta}(\mathbf{p})\left(\mathcal{K}^{-1}(s_{\eta},\mathbf{p})\mathbf{B}(\mathbf{p})\otimes\cdots\otimes\mathcal{K}^{-1}(s_1,\mathbf{p})\mathbf{B}(\mathbf{p})\right), \tag{32c}$$

where $\mathcal{K}^{-1}(s,\mathbf{p})$ is the Laplace transform of the fundamental solution of $\mathcal{L}(\mathbf{p})$. If a ROM is computed using the projection matrices $\mathbf{V}$ and $\mathbf{W}$, assuming $\mathbf{V}$ and $\mathbf{W}$ are full rank matrices, the reduced operator and matrices are given as follows:

$$\begin{aligned}\widehat{\mathcal{K}}(s,\mathbf{p}) &= \mathbf{W}^{\top}\widehat{\mathcal{K}}(s,\mathbf{p})\mathbf{V}, \quad \widehat{\mathbf{N}}_{\eta}(\mathbf{p}) = \mathbf{W}^{\top}\mathbf{N}_{\eta}(\mathbf{p})\mathbf{V}^{\textcircled{\scriptsize\eta}}, \quad \eta\in\{1,\ldots,d-1\},\\ \widehat{\mathbf{B}}(\mathbf{p}) &= \mathbf{W}^{\top}\mathbf{B}(\mathbf{p}), \quad \widehat{\mathbf{C}}(\mathbf{p}) = \mathbf{C}(\mathbf{p})\mathbf{V}, \quad \widehat{\mathbf{H}}_{\xi}(\mathbf{p}) = \mathbf{W}^{\top}\mathbf{H}_{\xi}(\mathbf{p})\mathbf{V}^{\textcircled{\scriptsize\xi}}, \quad \xi\in\{2,\ldots,d\},\end{aligned} \tag{33}$$

and the inverse Laplace transform of $\widehat{\mathcal{K}}^{-1}(s,\mathbf{p})$ is the fundamental solution related to the reduced operator $\widehat{\mathcal{L}}(\mathbf{p})$. In the following, an extension of Theorem 3.1 to the parametric case is presented that allows to construct an interpolatory ROM. Interpolation-based MOR for structured parametric bilinear systems has been investigated in [16], which we generalize to more general nonlinear systems.

**Theorem 4.1** *Consider the original polynomial parametric system* (30)*, together with its multivariate transfer functions given in* (32)*. Let* $\sigma_i$, $\mathbf{p}_i$ *and* $\mu_i$, $i\in\{1,\ldots,\widetilde{r}\}$, *be interpolation points such that* $\mathcal{K}(s,\mathbf{p})$ *is invertible for all* $s\in\{\sigma_i,\mu_i\}$, $i\in\{1,\ldots,\widetilde{r}\}$, $\mathbf{p}\in\{\mathbf{p}_1,\ldots,\mathbf{p}_{\widetilde{r}}\}$. *Moreover, let* $\mathbf{V}$ *and* $\mathbf{W}$ *be defined as follows:*

$$\mathcal{V}_{\mathrm{L}} = \bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\right),$$

$$\mathcal{V}_{\mathrm{N}} = \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{N}_{\eta}(\mathbf{p}_i)\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\right)\right),$$

$$\mathcal{V}_{\mathrm{H}} = \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{H}_{\xi}(\mathbf{p}_i)\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\cdots\otimes\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\right)\right),$$

$$\mathcal{W}_{\mathrm{L}} = \bigcup_{i=1}^{\widetilde{r}}\operatorname{range}\left(\mathcal{K}^{-\top}(\mu_i,\mathbf{p}_i)\mathbf{C}(\mathbf{p}_i)^{\top}\right)$$

$$\mathcal{W}_{\mathrm{N}} = \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\left(\mathbf{N}_\eta(\mathbf{p}_i)\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\cdots\right.\right.$$
$$\left.\left.\otimes\,\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\mathcal{K}^{-\top}(\mu_i,\mathbf{p}_i)\mathbf{C}(\mathbf{p}_i)^\top\right)\right),$$

$$\mathcal{W}_{\mathrm{H}} = \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\left(\mathbf{H}_\xi(\mathbf{p}_i)\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\cdots\right.\right.$$
$$\left.\left.\otimes\,\mathcal{K}^{-1}(\sigma_i,\mathbf{p}_i)\mathbf{B}(\mathbf{p}_i)\otimes\mathcal{K}^{-\top}(\mu_i,\mathbf{p}_i)C(\mathbf{p}_i)^\top\right)\right),$$

$$\operatorname{range}(\mathbf{V}) = \mathcal{V}_{\mathrm{L}} + \mathcal{V}_{\mathrm{N}} + \mathcal{V}_{\mathrm{H}},$$
$$\operatorname{range}(\mathbf{W}) = \mathcal{W}_{\mathrm{L}} + \mathcal{W}_{\mathrm{N}} + \mathcal{W}_{\mathrm{H}},$$

*where* $\left(\mathbf{H}_\xi(\mathbf{p})\right)_{(2)} \in \mathbb{R}^{n\times n^\xi}$ *and* $\left(\mathbf{N}_\eta(\mathbf{p})\right)_{(2)} \in \mathbb{R}^{n\times m\cdot n^\xi}$ *are, respectively, the mode-2 matricizations of the* $(\xi+1)$*-way symmetric tensor* $\boldsymbol{\mathcal{H}}_\xi(\mathbf{p}) \in \mathbb{R}^{n\times\cdots\times n}$ *and* $(\eta+2)$*-way symmetric tensor* $\boldsymbol{\mathcal{N}}_\eta(\mathbf{p}) \in \mathbb{R}^{n\times\cdots\times n}$ *whose mode-1 matricizations are* $\mathbf{H}_\xi(\mathbf{p})$ *and* $\mathbf{N}_\eta(\mathbf{p})$*, respectively. Then, the following interpolation conditions are fulfilled:*

$$\begin{aligned}
\mathbf{F}_{\mathrm{L}}(\sigma_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i,\mathbf{p}_i),\\
\mathbf{F}_{\mathrm{L}}(\mu_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{L}}(\mu_i,\mathbf{p}_i),\\
\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i),\\
\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mu_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mu_i,\mathbf{p}_i)\\
\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i),\\
\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mu_i,\mathbf{p}_i) &= \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mu_i,\mathbf{p}_i),
\end{aligned}$$

*provided the reduced matrix* $\mathbf{W}^\top\mathcal{K}(s,\mathbf{p})\mathbf{V}$ *is non-singular for* $s \in \{\sigma_i,\mu_i\}$ *and* $\mathbf{p} \in \{\mathbf{p}_i\}$*,* $i \in \{1,\ldots,\widetilde{r}\}$*.*

***Proof*** The proof is analogous to the proof of Theorem 3.1 and extends the proof of [16, Thm. 2]. Therefore, for the brevity of the paper, we skip the proof.

Note that we have assumed a general parametric structure for the system matrices, e.g., $\mathcal{L}(\mathbf{p})$ (or $\mathcal{K}(s,\mathbf{p})$), $\mathbf{H}_\xi(\mathbf{p})$. Then, the corresponding ROM can be computed as shown in (33), but it may be required to compute a ROM for each parameter setting. However, if we assume an affine parametric structure of the system matrices as follows:

$$\mathcal{K}(s,\mathbf{p}) = \sum_{i=1}^{t_A}\kappa_i(s,\mathbf{p})\mathbf{A}^{(i)},\qquad \mathbf{B}(\mathbf{p}) = \sum_{i=1}^{t_b}\alpha_b^{(i)}(\mathbf{p})\mathbf{B}^{(i)},\quad \mathbf{C}(\mathbf{p}) = \sum_{i=1}^{t_c}\alpha_c^{(i)}(\mathbf{p})\mathbf{C}^{(i)},$$
$$\mathbf{N}_\eta(\mathbf{p}) = \sum_{i=1}^{t_{n_\eta}}\alpha_{n_\eta}^{(i)}(\mathbf{p})\mathbf{N}_\eta^{(i)},\qquad \mathbf{H}_\xi(\mathbf{p}) = \sum_{i=1}^{t_{h_\xi}}\alpha_{h_\xi}^{(i)}(\mathbf{p})\mathbf{H}_\xi^{(i)}, \tag{36}$$

then the resulting ROM with the same structure can be determined using

$$
\begin{aligned}
\mathcal{K}(s,\mathbf{p}) &= \sum_{i=1}^{t_A} \kappa_i(s,\mathbf{p})\widehat{\mathbf{A}}^{(i)}, \qquad \widehat{\mathbf{B}}(\mathbf{p}) = \sum_{i=1}^{t_b} \alpha_b^{(i)}(\mathbf{p})\widehat{\mathbf{B}}^{(i)}, \quad \widehat{\mathbf{C}}(\mathbf{p}) = \sum_{i=1}^{t_c} \alpha_c^{(i)}(\mathbf{p})\widehat{\mathbf{C}}^{(i)},\\
\widehat{\mathbf{N}}_\eta(\mathbf{p}) &= \sum_{i=1}^{t_{n_\eta}} \alpha_{n_\eta}^{(i)}(\mathbf{p})\widehat{\mathbf{N}}_\eta^{(i)}, \qquad \widehat{\mathbf{H}}_\xi(\mathbf{p}) = \sum_{i=1}^{t_{h_\xi}} \alpha_{h_\xi}^{(i)}(\mathbf{p})\widehat{\mathbf{H}}_\xi^{(i)},
\end{aligned}
\tag{37}
$$

where the original matrices are reduced by the standard projections

$$
\begin{aligned}
\widehat{\mathbf{A}}^{(i)} &= \mathbf{W}^\top \mathbf{A}^{(i)}\mathbf{V}, \qquad \widehat{\mathbf{B}}^{(i)} = \mathbf{W}^\top \mathbf{B}^{(i)}, \qquad \widehat{\mathbf{C}}^{(1)} = \mathbf{C}^{(i)}\mathbf{V},\\
\widehat{\mathbf{N}}_\eta^{(i)} &= \mathbf{W}^\top \mathbf{N}_\eta^{(i)}\mathbf{V}^{\textcircled{\scriptsize$\eta$}}, \quad \widehat{\mathbf{H}}_\xi^{(i)} = \mathbf{W}^\top \mathbf{H}_\xi^{(i)}\mathbf{V}^{\textcircled{\scriptsize$\xi$}},
\end{aligned}
\tag{38}
$$

for $\eta \in \{1,\ldots,d-1\}$ and $\xi \in \{2,\ldots,d\}$. Consequently, we can pre-compute these reduced matrices. Hence, the computation of a ROM for each parameter becomes numerically cheaper.

**Remark 4.1** As in Theorem 3.2, if $\sigma_i = \mu_i$, Hermite interpolation conditions with respect to the parameter are satisfied, i.e.,

$$
\begin{aligned}
\nabla_{\mathbf{p}}\, \mathbf{F}_{\mathrm{L}}(\sigma_i,\mathbf{p}_i) &= \nabla_{\mathbf{p}}\, \widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i,\mathbf{p}_i),\\
\nabla_{\mathbf{p}}\, \mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i) &= \nabla_{\mathbf{p}}\, \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i),\\
\nabla_{\mathbf{p}}\, \mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i) &= \nabla_{\mathbf{p}}\, \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i,\mathbf{p}_i).
\end{aligned}
$$

Thus, the sensitivity of the multivariate transfer functions with respect to the parameter is preserved, which may be useful in, e.g., parameter optimization. Since this result can be proven in a similar way as the one in Theorem 3.2 and [16, Thm. 3], we decide to omit it for the brevity of the paper.

### 4.2 Structured input and output matrices

One may also consider the case where the input and output matrices $\mathbf{B}$ and $\mathbf{C}$ depend on the frequency $s$. This happens to be the case if the system is subject to input or/and output delays, e.g, $\mathcal{B}(s) = \mathbf{B}_1 + \mathbf{B}_2 e^{-\tau s}$. In such scenarios, we can employ the results of Theorem 3.1 by replacing $\mathbf{B}$ by $\mathcal{B}(\sigma_i)$ and $\mathbf{C}$ by $\mathcal{C}(\mu_i)$ everywhere, e.g., the matrices $\mathbf{V}_{\mathrm{L}}$ and $\mathbf{W}_{\mathrm{L}}$ in Theorem 3.1 are given by

$$
\begin{aligned}
\mathbf{V}_{\mathrm{L}} &= \mathrm{range}\left(\mathcal{K}^{-1}(\sigma_1)\mathcal{B}(\sigma_1),\ldots,\mathcal{K}^{-1}(\sigma_{\tilde r})\mathcal{B}(\sigma_{\tilde r})\right),\\
\mathbf{W}_{\mathrm{L}} &= \mathrm{range}\left(\mathcal{K}^{-\top}(\mu_1)\mathcal{C}^\top(\mu_1),\ldots,\mathcal{K}^{-\top}(\mu_{\tilde r})\mathcal{C}^\top(\mu_{\tilde r})\right).
\end{aligned}
$$

The interpolation results presented in this paper also hold in this context. Since their proof is a straightforward extension of the results presented up-to-now, we refrain from providing details.

## 5 Determining lower-order approximate interpolatory models

In the previous sections, we have stated results for constructing interpolatory ROMs by a Petrov-Galerkin projection. In the proposed methodology, the quality of the ROMs highly depends on the choice of the interpolation points. It is an open question of how to select these interpolation points optimally, and this remains an important problem to be investigated in the future.

However, inspired by the discussions in [3, 12, 27, 31], we propose a scheme based on a oversampling as follows. In this scheme, we first compute the projection matrices **V** and **W** by considering several interpolation points in a given domain. After that, we aim to determine the dominant subspaces that not only allow us to determine lower-order models but also approximately interpolate all the considered interpolation points. To that end, let us assume to have a structured polynomial system as in (3) and the projection matrices **V** and **W** as defined in Theorem 3.1. For this section, we assume that **V** and **W** contain the columns of the Krylov basis exactly for the given interpolation points. Furthermore, let us construct the matrices

$$\begin{aligned} \widetilde{\mathcal{K}}(s) &= \mathbf{W}^\top \mathcal{K}(s)\mathbf{V}, \quad \widetilde{\mathbf{B}} = \mathbf{W}^\top \mathbf{B}, \quad \widetilde{\mathbf{C}} = \mathbf{C}\mathbf{V}, \\ \widetilde{\mathbf{H}}_\xi &= \mathbf{W}^\top \mathbf{H}_\xi \mathbf{V}^{\textcircled{\xi}}, \quad \widetilde{\mathbf{N}}_\eta = \mathbf{W}^\top \mathbf{N}_\eta \mathbf{V}^{\textcircled{\eta}}. \end{aligned} \tag{40}$$

Next, we can extend the observation from [3, 12, 27]. This means that if the pencil $\widetilde{\mathcal{K}}(s)$ is regular, then the realization (40) is a realization interpolating the original model.

However, when we consider many interpolation points, then often the pencil $\widetilde{\mathcal{K}}(s)$ becomes singular. In this case, there exists a lower-order realization that can interpolate at all the given interpolation points. To obtain such a realization, we follow the idea proposed in [3, 12, 14, 27]. For this, we first consider the form of $\mathcal{K}(s)$ to be

$$\mathcal{K}(s) = \alpha_1(s)\mathbf{A}^{(1)} + \cdots + \alpha_l(s)\mathbf{A}^{(l)}, \tag{41}$$

and assume that

$$\operatorname{rank}\left[\mathbf{W}^\top\mathbf{A}^{(1)}\mathbf{V}, \ldots, \mathbf{W}^\top\mathbf{A}^{(l)}\mathbf{V}\right] = \operatorname{rank}\begin{bmatrix}\mathbf{W}^\top\mathbf{A}^{(1)}\mathbf{V}\\ \vdots \\ \mathbf{W}^\top\mathbf{A}^{(l)}\mathbf{V}\end{bmatrix} = \widehat{r}. \tag{42}$$

Then, according to [3, 12, 14, 27], there exists a structured system of order $\widehat{r} \leq r$, realizing the model whose generalized transfer functions also interpolate at the predefined interpolation points. Consequently, using (42), we can estimate the complexity of the underlying dynamical system. Additionally, an SVD procedure based on the matrices in (42) allows us to construct a ROM using appropriate subspaces. Moreover, if small singular values are neglected while estimating the order $\widehat{r}$ in (42), the resulting subspaces will lead to a reduced-order model satisfying approximately the interpolation conditions for all selected interpolation points.

**Algorithm 1** Construction of ROMs for Structured Polynomial Systems.

1: **Input:** The polynomial structured system matrices as in (36), and order of a ROM $r$.
2: Choose interpolation points to construct $\mathbf{V}$ and $\mathbf{W}$.
3: Compute $\mathbf{V}$ and $\mathbf{W}$ using the interpolation points as in Theorem 3.1.
4: Determine SVDs

$$\left[\mathbf{W}^\top\mathbf{A}^{(1)}\mathbf{V},\ldots,\mathbf{W}^\top\mathbf{A}^{(l)}\mathbf{V}\right] = \mathbf{W}_1\Sigma_l\widetilde{\mathbf{V}}^\top \quad \text{and} \quad \begin{bmatrix}\mathbf{W}^\top\mathbf{A}^{(1)}\mathbf{V}\\ \vdots\\ \mathbf{W}^\top\mathbf{A}^{(l)}\mathbf{V}\end{bmatrix} = \widetilde{\mathbf{W}}\Sigma_r\mathbf{V}_1^\top \tag{43}$$

5: Compute projection matrices: $\mathbf{V}_\mathrm{e} = \mathbf{V}\mathbf{V}_1(:, 1{:}r)$ and $\mathbf{W}_\mathrm{e} = \mathbf{W}\mathbf{W}_1(:, 1{:}r)$.
6: Compute reduced matrices:

$$\widehat{\mathcal{K}}(s) = \mathbf{W}_\mathrm{e}^\top\mathcal{K}(s)\mathbf{V}_\mathrm{e}, \quad \widehat{\mathbf{B}} = \mathbf{W}_\mathrm{e}^\top\mathbf{B}, \quad \widehat{\mathbf{C}} = \mathbf{C}\mathbf{V}_\mathrm{e},$$
$$\widehat{\mathbf{H}}_\xi = \mathbf{W}_\mathrm{e}^\top\mathbf{H}_\xi\mathbf{V}_\mathrm{e}^{\textcircled{\xi}}, \quad \widehat{\mathbf{N}}_\eta = \mathbf{W}_\mathrm{e}^\top\mathbf{N}_\eta\mathbf{V}_\mathrm{e}^{\textcircled{\eta}},$$

7: **Output:** The reduced-order matrices: $\widehat{\mathcal{K}}(s)$, $\widehat{\mathbf{B}}$, $\widehat{\mathbf{C}}$, $\widehat{\mathbf{H}}_\xi$ and $\widehat{\mathbf{N}}_\eta$.

To obtain the corresponding subspaces and the ROM, we propose Algorithm 1, enabling construction of ROMs for polynomial structured systems (3). This can be seen as an extension of [12, Algo. 3] to structured polynomial systems. The procedure consists in selecting interpolation points $\sigma_i$, $i \in \{1,\ldots,\widetilde{r}\}$, and constructing the matrices $\mathbf{V}$ and $\mathbf{W}$ as in Theorem 3.1 (steps 2 and 3) containing the original Krylov basis. The interpolation points ideally should be selected in the frequency range of interest, which often depends on the application. In our numerical experiments, they are typically selected using a logarithmically spaced grid in the suitable frequency range. Then, in step 4, we compute the SVDs of the matrices in (43). As discussed earlier, the numerical rank of these matrices indicates the order of a good ROM, and the left and right singular vectors allow us to determine dominant subspaces. Hence, in step 5, the projection matrices $\mathbf{V}_\mathrm{e}$ and $\mathbf{W}_\mathrm{e}$ are constructed. Finally, in step 6, the ROM is computed in the framework of Petrov-Galerkin projection.

Algorithm 1 can easily be adapted to parametric and MIMO cases by computing the projection matrices $\mathbf{V}$ and $\mathbf{W}$ appropriately in step 3 of the algorithm. In this case, we need to specify parameters within the range of interest and tangential directions along with the interpolation points for the frequency variable.

**Remark 5.1** In many applications, such as dynamical systems with symmetric matrices, or systems with a dissipative realization, it is desirable to apply Galerkin projections, i.e., to enforce $\mathbf{W} = \mathbf{V}$ (see, e.g., [22]). As a result, the reduced systems would potentially preserve system properties, e.g., dissipativity, symmetry. Although Algorithm 1 is designed for Petrov-Galerkin projection, one can still adapt it to allow only Galerkin projection. Indeed, in step 3 one just needs to compute $\mathbf{V}$ and set $\mathbf{W} = \mathbf{V}$. Hence, in step 5, $\mathbf{W}_\mathrm{e} = \mathbf{V}_\mathrm{e}$.

**Remark 5.2** It is worth mentioning that one does not need to use any hyper reduction method for the fast computation of the nonlinear terms. Indeed, as shown in (7), the nonlinear terms are already directly computed using the Petrov-Galerkin projection.

Additionally, at a first glance, the computation of the reduced matrix $\widehat{\mathbf{H}}_\xi = \mathbf{W}^\top \mathbf{H}_\xi \mathbf{V}^{\textcircled{\xi}}$ seems to be numerically expensive as one needs to evaluate $\mathbf{V}^{\textcircled{\xi}}$. However, the authors in [11] proposed a procedure based on Hadamard product form of the term $\mathbf{H}_\xi x^{\textcircled{\xi}} = \mathcal{A}_1 x \circ \cdots \circ \mathcal{A}_\eta x$. As a consequence, the computation of $\widehat{\mathbf{H}}_\xi$ can be performed without the explicit computation of $\mathbf{V}^{\textcircled{\xi}}$.

## 6 Numerical examples

In this section, we illustrate the efficiency of the proposed method by means of three examples. We also compare with the recent interpolation method proposed in [16, 17] for two of the examples. To integrate nonlinear structured systems, we use an explicit Euler scheme, and for non-structured systems, we use the function `ode15s` in MATLAB®. The interpolation points are selected using a logarithmically spaced grid in the suitable frequency range. For the random number generator, we have used the seed '0'. All experiments were performed using MATLAB (2020b) running on a Macbook Pro with 2,3 GHz 8-Core Intel® Core™i9 CPU, 16GB of RAM, and Mac OS X v10.15.6.

### 6.1 Parametric Chafee-Infante equation

In our first example, we consider the one-dimensional parametric Chafee-Infante system governed by the following partial differential equation

$$
\begin{aligned}
\dot{v}(t) &= v_{xx} + v(\mathbf{p} - v^2), \quad x \in (0,1) \times (0,T), & v(0,t) &= u(t), \quad t \in (0,T),\\
v_x(L,t) &= 0, \quad t \in (0,T), & v(x,0) &= 0, \quad x \in (0,1),
\end{aligned}
$$

where the parameter is assumed to lie in the interval $\mathbf{p} \in [0.25, 2]$. After a spatial discretization using a finite-difference method for a uniform grid with $k = 500$ points, one obtains a high-fidelity cubic parametric model of the form:

$$
\begin{aligned}
\dot{\mathbf{v}}(t) &= \mathbf{A}_1\mathbf{v}(t) + \mathbf{p}\mathbf{A}_p + \mathbf{H}_3\left(\mathbf{x}(t) \otimes \mathbf{x}(t) \otimes \mathbf{x}(t)\right) + \mathbf{B}\mathbf{u}(t),\\
\mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t).
\end{aligned} \tag{44}
$$

The MOR problem for its non-parametric variant (for $\mathbf{p} = 1$) has been considered in [12], where the polynomial non-linearity in the ROMs is preserved. Also, authors in, e.g., [9, 10, 13] have constructed ROMs of the model (44), but they do not preserve the cubic structure. These methods rewrite the model into quadratic-bilinear form, followed by employing MOR techniques for quadratic-bilinear systems.

Here, we aim at constructing a reduced cubic parametric system using Algorithm 1. For this, we take 200 points in the frequency range $\left[10^{-3}, 10^{3}\right]$ and the equal number of points for the parameter in the considered interval.

First, in Fig. 1, we plot the decay of the singular values based on the matrices in (42) and we observe a very fast decay. Subsequently, we determine a reduced parametric system of order $r = 5$. To compare the quality of the ROM, we simulate

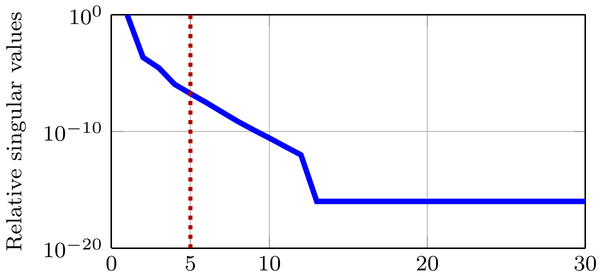



**Fig. 1** Parametric Chafee-Infante equation: relative decay of singular values based on the Loewner pencil

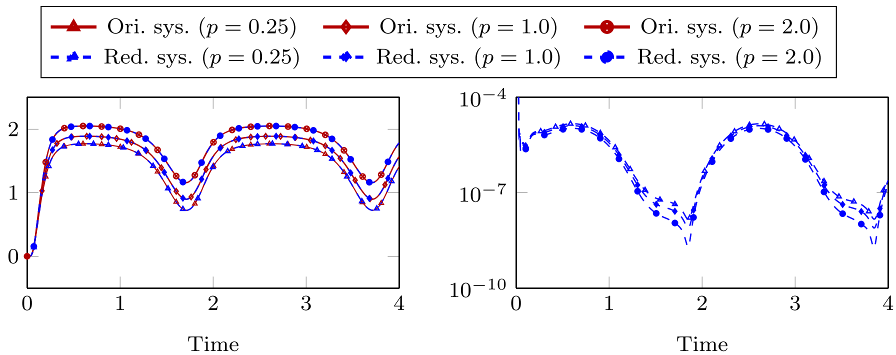
for the same inputs $u^{(1)}(t) = 10(\sin(\pi t) + 1)$ and $u^{(2)}(t) = 5(te^{-t})$ and parameters $\mathbf{p} = \{0.25, 1, 2\}$. We plot the transient response and relative errors in Figs. 2 and 3, illustrating that the reduced parametric system can capture the dynamics of the high-fidelity model very well for different inputs and different parameters.

## 6.2 Mechanical system

In our second example, we consider a damped mass-spring system from [17]. The dynamics of the system is governed by a second-order bilinear system of the form:

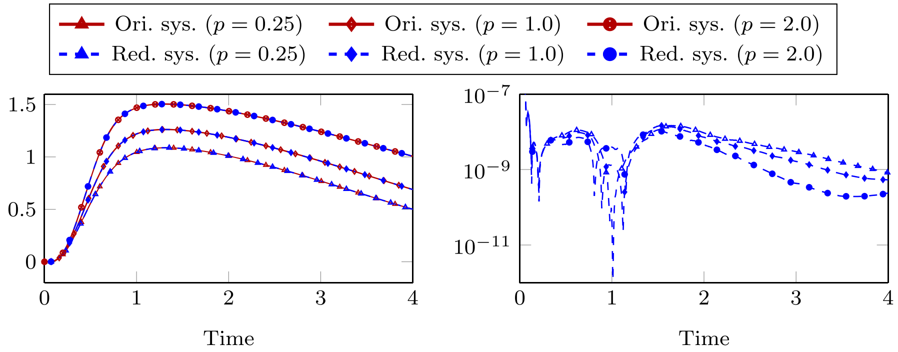
$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{N}_1\mathbf{x}(t)\mathbf{u}_1(t) + \mathbf{N}_2\mathbf{x}(t)\mathbf{u}_2(t) + \mathbf{B}\mathbf{u}(t),$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)\mathbf{x}(t).$$

We consider the same setting as provided in [17, Sec. 4.2]. The order of the full model is $n = 1\,000$, and the model has two inputs and two outputs. Mechanical systems can have many structural properties such as passivity, and positive definiteness of mass, damping, and stiffness matrices, see, e.g., the survey paper [7]. Therefore, it is desired to preserve these properties in the ROM. This can be achieved if the ROM




(a) Transient response. (b) Relative error.

**Fig. 2** Parametric Chafee-Infante equation: a comparison of the original and ROM for the input $\mathbf{u}^{(1)} = 10\,(\sin(\pi t) + 1)$ and for different parameter values




(a) Transient response.

(b) Relative error.

**Fig. 3** Parametric Chafee-Infante equation: a comparison of the original and ROM for the input $\mathbf{u}^{(2)} = 5\left(e^{-t}t\right)$ and for different parameter values

is determined via Galerkin (one-sided) projection instead of Petrov-Galerkin (two-sided) one. Hence, in this example, we construct a ROM using Galerkin projection via Algorithm 1 as discussed in Remark 5.1.

Next, to employ Algorithm 1, we consider 1 000 logarithmically distributed frequency points on the imaginary axis in the range $\left[10^{-3}, 10^{3}\right]$. Since the system is MIMO, we also choose tangential directions which are taken randomly. We then observe the singular values obtained from Algorithm 1 in Fig. 4, indicating a sharp decay of the singular values. From the figure, we note that the singular values after 33 are at the level of machine precision.

We determine ROMs of order $r = \{10, 20, 30\}$ (denoted by `StrDsp_SO`). We compare the quality of the ROMs `StrDsp_SO` with the method proposed in [17]. The ROM in [17] (denoted by `StrInt_SO`) is computed based on interpolation. We choose the same interpolation points as in [17], which yields 36 basis vectors. To determine ROMs of order $r = \{10, 20, 30\}$, we take the same number of dominant basis vectors as the order out of 36 basis vectors, and these dominant basis vectors are determined based on the QR decomposition of the basis vectors.

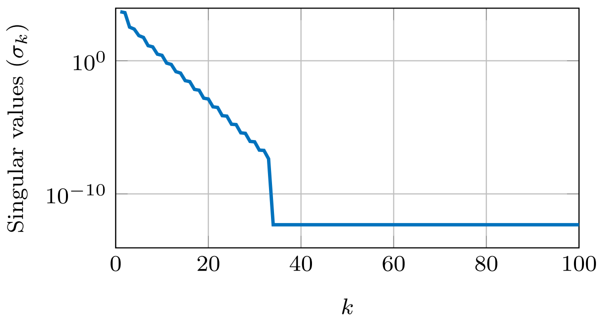
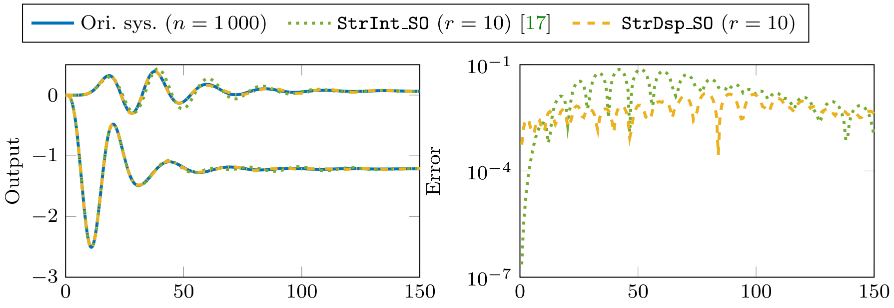



**Fig. 4** Mechanical example: The decay of singular values obtained from Algorithm 1




(a) Transient response.

(b) Mean absolute error.

**Fig. 5** Mechanical example: A comparison of the transient responses the original and ROMs of order $r = 10$ for the control input $\mathbf{u}(t) = 50 \cdot \begin{bmatrix} \sin(20t) + 1 \\ \sin(t)e^{-0.1t} \end{bmatrix}$

Next, to assess the quality of both ROMs, we compare transient responses of them with the full model using a control input $\mathbf{u}(t) = 50 \cdot \begin{bmatrix} \sin(20t) + 1 \\ \sin(t)e^{-0.1t} \end{bmatrix}$. This is shown in Fig. 5. Moreover, Table 2 shows a comparison of the $L_2$ and $L_\infty$-errors these three different ROMs. We observe that our method constructs ROMs which are consistently better both in $L_2$ and $L_\infty$-norm for all orders.

Finally, we would like to remark on computational aspects. Although we could obtain a better reduced-order model using `StrDsp_SO` (approximately two-order of magnitude better for smaller orders), it comes with computational expenses. The philosophy of `StrDsp_SO` relies on considering many interpolation points and, after that, on compression to determine global dominant subspaces. On the other hand, `StrInt_SO` considers carefully choosing a few interpolation points. Consequently, `StrDsp_SO` becomes much more computationally expensive as compared to `StrDsp_SO`; for this example, 1 000 interpolation points are considered for `StrDsp_SO`, where only 3 points interpolation points are considered for `StrInt_SO`. In the future, we will investigate active-learning-based approaches to sample a few points for `StrDsp_SO` so that only relevant interpolation points (e.g., from 1 000) are considered to determine the global dominant subspaces.

**Table 2** Mechanical example: The $L_2$ and $L_\infty$-errors of the outputs between the original model and ROMs

| | $L_2$-error | | $L_\infty$-error | |
|---|---|---|---|---|
| | `StrInt_SO` | `StrDsp_SO` | `StrInt_SO` | `StrDsp_SO` |
| $r = 10$ | $4.68 \cdot 10^{-4}$ | $1.01 \cdot 10^{-4}$ | $7.48 \cdot 10^{-2}$ | $1.59 \cdot 10^{-2}$ |
| $r = 20$ | $1.06 \cdot 10^{-5}$ | $6.31 \cdot 10^{-6}$ | $1.63 \cdot 10^{-3}$ | $7.65 \cdot 10^{-4}$ |
| $r = 30$ | $8.59 \cdot 10^{-8}$ | $3.83 \cdot 10^{-8}$ | $2.16 \cdot 10^{-5}$ | $3.98 \cdot 10^{-6}$ |

### 6.3 Parametric bilinear time-delay system

In our last experiment, we consider an example from [16, 28] that models a time-delayed heated rod using a one-dimensional parametric heat equation:

$$\partial_t v(x,t) = \partial_x^2 v(x,t) - \mathbf{p}\sin(x)v(x,t) + \mathbf{p}\sin(x)v(x,t-1) + u(t), \tag{45}$$

with homogeneous Dirichlet boundary conditions and $\mathbf{p} \in [1, 10]$. Spatial discretization yields a parametric bilinear systems of the form:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}(\mathbf{p})\mathbf{x}(t) + \mathbf{p}\mathbf{A}_\mathrm{d}\mathbf{x}(t-1) + \mathbf{N}\mathbf{x}(t)\mathbf{u}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t), \end{aligned} \tag{46}$$

where $\mathbf{A}(\mathbf{p}) = \mathbf{A}_0 - \mathbf{p}\mathbf{A}_\mathrm{d}$. We have employed 5 000 grid points to discretize the PDEs (45), thus leading to the state-space model (46) of order $n = 5\,000$. Additionally, this model is single-input single-output. The example also completely fits in our set-up, discussed in Section 2. In this case,

$$\begin{aligned} \mathcal{K}(s,\mathbf{p}) = s\mathbf{I}_n - (\mathbf{A}_0 - \mathbf{p}\mathbf{A}_\mathrm{d}) - \mathbf{p}e^{-s}\mathbf{A}_\mathrm{d}, \quad \mathbf{B}(\mathbf{p}) = \mathbf{B}, \\ \mathbf{N}(\mathbf{p}) = \mathbf{N}, \quad \mathbf{H}(\mathbf{p}) = 0, \quad \text{and } \mathbf{C}(\mathbf{p}) = \mathbf{C}. \end{aligned}$$

Next, we aim at constructing a ROM using Algorithm 1. In order to employ the algorithm, we consider 1 000 logarithmically distributed frequency points in the range $[10^{-2}, 10^2]$, and for the parameter $\mathbf{p}$, we randomly take the parameter in the considered parameter range. Consequently, we have tuples $\{\sigma_i, \mathbf{p}_i\}$, $i \in \{1, \ldots, 1\,000\}$.

Next, we plot the decay of the singular values obtained from Algorithm 1 in Fig. 6, indicating a rapid decay. Hence, we can expect a good quality ROM of small order. Then, we determine ROMs of order $r = \{5, 10, 20\}$ using Algorithm 1 (denoted by (`StrDsp_delay`)). We compare the quality of the (`StrDsp_delay`) with a ROM obtained using the interpolation-based methodology proposed in [16] (denoted by `StrInt_delay`). We use the interpolation points as in [16], which yield 24 basis

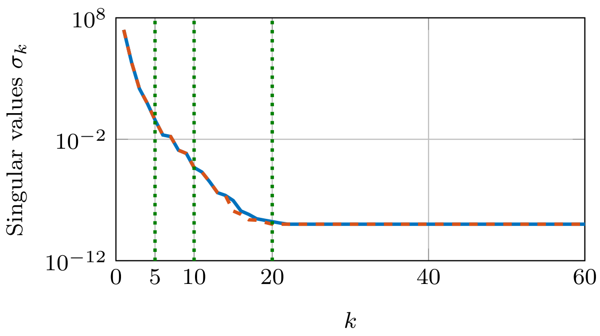
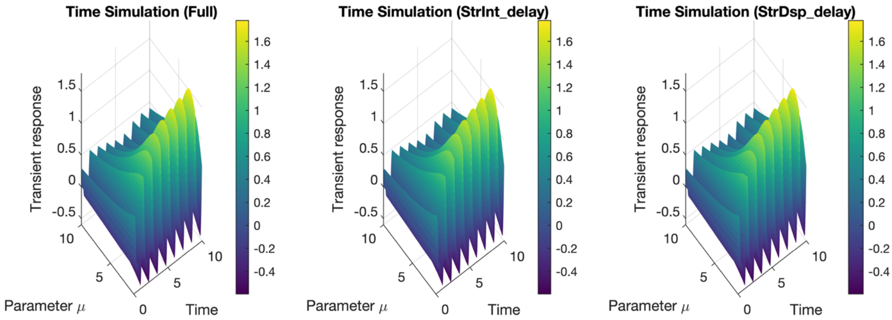
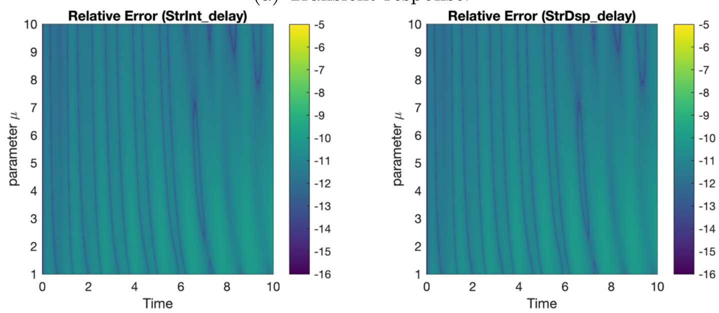



**Fig. 6** Delay parametric example: The decay of singular values obtained from Algorithm 1




(a) Transient response.




(b) Relative error.

**Fig. 7** Delay parametric example: A comparison of transient response for an input $u(t) = 0.05\,(\cos(10t) + \cos(5t))$

vectors. We construct `StrInt_delay` ROMs of order $r = \{5, 10, 20\}$ using the dominant basis vectors of the 24 basis vectors, which is done by taking QR decomposition of the basis vectors.

Next, we compare the time-domain simulation of both ROMs with the high-fidelity model for a wide range of parameters. We use the same control input as in [16, 28]; that is $u(t) = 0.05\,(\cos(10t) + \cos(5t))$, and vary the parameter in the range of interest. To perform time-domain simulation, we employ the explicit Euler method with time stepping $dt = 2 \cdot 10^{-2}$. To measure the quality of the ROMs, we consider the following error function:

$$\mathcal{E}(t, \mathbf{p}) := \frac{\|\mathbf{y}(t; \mathbf{p}) - \widehat{\mathbf{y}}(t; \mathbf{p})\|}{\max\limits_{t} \max\limits_{\mathbf{p}\in[1,10]} \|\mathbf{y}(t; \mathbf{p})\|}. \tag{47}$$

**Table 3** Delay parametric example: Error between the full model and ROMs, namely `StrInt_delay` and `StrDsp_delay` of different orders

| | `StrInt_SO` | `StrDsp_SO` |
|---|---|---|
| $r = 5$ | $8.45 \cdot 10^{-6}$ | $1.00 \cdot 10^{-6}$ |
| $r = 10$ | $1.87 \cdot 10^{-8}$ | $3.25 \cdot 10^{-9}$ |
| $r = 20$ | $1.38 \cdot 10^{-10}$ | $1.39 \cdot 10^{-10}$ |

Then, we plot the transient responses and errors (as defined in (47)) for different parameters in Fig. 7. Furthermore, we compute the maximum error in the time and parameter domain as follows:

$$\mathcal{E}_{\max} := \max_{\mathbf{p}\in[1,10]} \left( \max_{t\in[0,10]} \mathcal{E}(t, \mathbf{p}) \right),$$

which is reported in Table 3 (see the first columns of the table).

We notice that both ROMs perform equally good, but `StrDsp_delay` is consistently better than `StrInt_delay`. In particular, the proposed methodology can produce high-quality parametric ROMs of small orders as it determines the dominant subspace jointly for parameter and frequency in the given domains.

## 7 Conclusions

In this paper, we have studied a model order reduction problem for structured nonlinear systems. To that aim, we have defined generalized transfer functions for structured systems based on the associated Volterra series. We have then shown how to construct a reduced-order model such that its generalized transfer functions interpolate those of the original model at the pre-defined frequency points. Subsequently, we have proposed an algorithm that allows determining the dominant subspaces to construct minimal interpolatory reduced-order models. Moreover, we have discussed extensions of those results to parametric nonlinear structured systems, as well as special structured systems such as input-delay systems. Finally, we have illustrated the performance of the proposed methodology based on a couple of examples and compared it with the existing methodologies. In future work, we will focus on choosing interpolation points adaptively, which is essential to reduce computational efforts and to capture all the important dynamics of the system. To achieve this goal, one may tailor the ideas presented in [25].

## A Proof of Lemma 2.1

***Proof*** a). First note that

$$\mathbf{x}^{\circledN} = \sum_{i_1=1}^{n} \cdots \sum_{i_{\mathrm{N}}=1}^{n} \left(x_{i_1} x_{i_2} \cdots x_{i_n}\right) \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \cdots \otimes \mathbf{e}_{i_n}\right). \tag{48}$$

Since $x_{i_1} x_{i_2} \cdots x_{i_n} = x_{j_1} x_{j_2} \cdots x_{j_n}$, for every $(j_1, \ldots, j_n) \in \mathcal{S}_{\mathbf{i}}$, we can write

$$\mathbf{x}^{\circledN} = \sum_{i_1=1}^{n} \cdots \sum_{i_{\mathrm{N}}=1}^{n} \left(x_{i_1} x_{i_2} \cdots x_{i_n}\right) \left( \frac{1}{\alpha_{\mathbf{i}}} \sum_{\substack{(j_1,\ldots,j_n) \\ \in \mathcal{S}_{\mathbf{i}}}} \left(\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2} \otimes \cdots \otimes \mathbf{e}_{j_n}\right) \right).$$

Next, we have

$$\begin{aligned}
\mathbf{H}_{(1)}\mathbf{x}^{\circledN} &= \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} \left(x_{i_1}x_{i_2}\cdots x_{i_n}\right) \mathbf{H}_{(1)} \left( \frac{1}{\alpha_{\mathbf{i}}} \sum_{\substack{(j_1,\ldots,j_n)\\ \in \mathcal{S}_{\mathbf{i}}}} \left(\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2} \otimes \cdots \otimes \mathbf{e}_{j_n}\right) \right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} \left(x_{i_1}x_{i_2}\cdots x_{i_n}\right) \left( \widetilde{\mathbf{H}}_{(1)} \left( :, i_1 + \sum_{l=2}^{N} (i_l - 1) \left(n^{l-1}\right) \right) \right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} \left(x_{i_1}x_{i_2}\cdots x_{i_n}\right) \widetilde{\mathbf{H}}_{(1)} \left(\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \cdots \otimes \mathbf{e}_{i_n}\right) \\
&= \widetilde{\mathbf{H}}_{(1)}\mathbf{x}^{\circledN},
\end{aligned}$$

which proves part (a).

b). We begin with

$$\begin{aligned}
&\widetilde{\mathbf{H}}^{(1)} \left(\mathbf{q}^{(1)} \otimes \cdots \otimes \mathbf{q}^{(N)}\right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} \widetilde{\mathbf{H}}_{(1)} \left( :, \left( i_1 + \sum_{l=2}^{N} (i_l - 1) n^l \right) \right) \left( q_{i_1}^{(1)} \cdots q_{i_N}^{(N)} \right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_N=1}^{n} \mathbf{H}_{(1)} \left( \sum_{\substack{(j_1,\ldots,j_n)\\ \in \mathcal{S}_{\mathbf{i}}}} \frac{1}{\alpha_{\mathbf{i}}} \left(\mathbf{e}_{j_1} \otimes \mathbf{e}_{j_2} \otimes \cdots \otimes \mathbf{e}_{j_n}\right) \right) \\
&\qquad \left( q_{i_1}^{(1)} \cdots q_{i_N}^{(N)} \right).
\end{aligned} \tag{49}$$

Since $(j_1, \ldots, j_n) \in \mathcal{S}_{\mathbf{i}}$, the above equation is invariant to the interchange of the indices $i_k$. Therefore, the Kronecker product of the $q_i$'s can appear in any order that would yield the same $\widetilde{\mathbf{H}}_{(1)}\ (\widetilde{\mathbf{q}}_1 \otimes \cdots \otimes \widetilde{\mathbf{q}}_N)$, where $(\widetilde{\mathbf{q}}_1, \ldots, \widetilde{\mathbf{q}}_N)$ belongs to the set of all permutations of the set $\{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$. This proves the result.

c). Assuming $l_j \in \{1, \ldots, n\}$ for $j \in \{1, \ldots, N+1\}$, we have

$$\begin{aligned}
&\mathbf{e}_{l_2}\widetilde{\mathbf{H}}_{(2)} \left(\mathbf{e}_{l_{N+1}} \otimes \mathbf{e}_{l_N} \otimes \cdots \otimes \mathbf{e}_{l_3} \otimes \mathbf{e}_{l_1}\right) \\
&\quad \text{using (9)} \\
&\quad = \mathbf{e}_{l_1}\widetilde{\mathbf{H}}_{(1)} \left(\mathbf{e}_{l_{N+1}} \otimes \mathbf{e}_{l_N} \otimes \cdots \otimes \mathbf{e}_{l_3} \otimes \mathbf{e}_{l_2}\right) \\
&\quad \text{using (49)} \\
&\quad = \mathbf{e}_{l_1}\widetilde{\mathbf{H}}_{(1)} \left(\mathbf{e}_{l_{N+1}} \otimes \mathbf{e}_{l_N} \otimes \cdots \otimes \mathbf{e}_{l_{m+1}} \otimes \mathbf{e}_{l_2} \otimes \mathbf{e}_{l_m} \otimes \mathbf{e}_{l_{m-1}} \otimes \mathbf{e}_{l_3}\right) \\
&\quad \text{using (9)} \\
&\quad = \mathbf{e}_{l_2}\widetilde{\mathbf{H}}_{(m)} \left(\mathbf{e}_{l_{N+1}} \otimes \mathbf{e}_{l_N} \otimes \cdots \otimes \mathbf{e}_{l_{m+1}} \otimes \mathbf{e}_{l_m} \otimes \mathbf{e}_{l_{m-1}} \otimes \mathbf{e}_{l_3} \otimes \mathbf{e}_{l_1}\right).
\end{aligned}$$

This shows that the entries in $\widetilde{\mathbf{H}}_{(2)}$ and $\widetilde{\mathbf{H}}_{(m)}$ $(m \geq 2)$ are equal, implying the result.

## B Fundamental solution using frequency domain methods

The fundamental solution of a linear operator can be defined in different ways. In this work, we follow the approaches that use the frequency domain representation of the linear operator. We refer the reader to [8, 30] for more details.

Let us start by defining the unilateral Laplace transform. Given a function $\mathbf{g}(\cdot)$, its unilateral Laplace transform is given by

$$\mathscr{L}(\mathbf{g}(\cdot)) = \mathbf{G}(s) = \int_0^\infty e^{-st}\mathbf{g}(t)dt.$$

where $\mathbf{G}(\cdot)$ corresponds to the frequency domain representation of $\mathbf{g}(\cdot)$.

Now, let us consider a linear operator $\mathcal{L}$ such as shown in Table 1. By means of the Laplace transform, we obtain the frequency domain representation of $(\mathcal{L}\mathbf{x})(t)$ as follows:

$$\mathscr{L}((\mathcal{L}\mathbf{x})(t)) = \mathcal{K}(s)\mathbf{X}(s), \tag{50}$$

where $\mathbf{X}(s)$ corresponds to the Laplace transform of $\mathbf{x}(t)$ and $\mathcal{K}(s)$ corresponds to the frequency domain representation of the operator $\mathcal{L}$. Let us assume that the inverse of the Laplace transform of $\mathcal{K}^{-1}(s)$ exists and is given as

$$\mathscr{L}^{-1}(\mathcal{K}^{-1}(\cdot)) := \Phi(\cdot).$$

Then, $\Phi(t)$ is the *fundamental solution* associated to the linear operator $\mathcal{L}$. Indeed, $\Phi(t)$ is the solution of the functional differential equation

$$(\mathcal{L}\mathbf{x})(t) = \delta(t),$$

where $\delta(t)$ is the Dirac delta distribution and the initial conditions are all zero. Moreover, the inhomogeneous equation

$$(\mathcal{L}\mathbf{x})(t) = \mathbf{g}(t),$$

with $\mathbf{g}(\cdot)$ being a suitable function, has a solution in convolution form

$$\mathbf{x}(t) = \int_0^t \Phi(\sigma)\mathbf{g}(t-\sigma)d\sigma.$$

## C Tangential interpolation-based MOR for MIMO systems

Here, we discuss a construction of an interpolating ROM for MIMO polynomial systems. Similar to the SISO case, the leading generalized transfer functions for a MIMO polynomial system are given as follows:

$$\mathbf{F}_{\mathrm{L}}(s_1) = \mathbf{C}\mathcal{K}^{-1}(s_1)\mathbf{B}, \tag{51a}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(s_1, \ldots, s_{\xi+1}) = \mathbf{C}\mathcal{K}^{-1}(s_{\xi+1})\mathbf{H}_\xi \left(\mathcal{K}^{-1}(s_\xi)\mathbf{B} \otimes \cdots \otimes \mathcal{K}^{-1}(s_1)\mathbf{B}\right), \tag{51b}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(s_1, \ldots, s_{\eta+1}) = \mathbf{C}\mathcal{K}^{-1}(s_{\eta+1})\mathbf{N}_\eta \left(\mathbf{I}_m \otimes \mathcal{K}^{-1}(s_\eta)\mathbf{B} \otimes \cdots \otimes \mathcal{K}^{-1}(s_1)\mathbf{B}\right). \tag{51c}$$

**Lemma C.1** *Consider the original system as given in (3). Let $\sigma_i \in \mathbb{C}$, $i \in \{1, \ldots, \widetilde{r}\}$, be interpolation points such that $\mathcal{K}(s)$ is invertible for all $s \in \{\sigma_1, \ldots, \sigma_{\widetilde{r}}\}$, and $\mathbf{b}_i \in \mathbb{C}^m$ and $\mathbf{c}_i \in \mathbb{C}^q$ for $i \in \{1, \ldots, \widetilde{r}\}$ be right and left tangential directions corresponding to $\sigma_i$, respectively. Let $\mathbf{V}$ and $\mathbf{W}$ be defined as follows:*

$$\begin{aligned}
\mathcal{V}_{\mathrm{L}} &= \bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i\right),\\
\mathcal{V}_{\mathrm{N}} &= \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{N}_\eta\left(\mathbf{I}_m \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \cdots \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i\right)\right),\\
\mathcal{V}_{\mathrm{H}} &= \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{H}_\xi\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \cdots \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i\right)\right),\\
\mathcal{W}_{\mathrm{L}} &= \bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-\top}(\sigma_i)\mathbf{C}^\top\mathbf{c}_i\right),\\
\mathcal{W}_{\mathrm{N}} &= \bigcup_{\eta=1}^{d-1}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\left(\mathbf{N}_\eta\right)_{(2)}\left(\mathbf{I}_m \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \cdots \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \mathcal{K}^{-\top}(\sigma_i)\mathbf{C}^\top\mathbf{c}_i\right)\right),\\
\mathcal{W}_{\mathrm{H}} &= \bigcup_{\xi=2}^{d}\bigcup_{i=1}^{\widetilde{r}} \operatorname{range}\left(\mathcal{K}^{-1}(\sigma_i)\left(\mathbf{H}_\xi\right)_{(2)}\left(\mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \cdots \otimes \mathcal{K}^{-1}(\sigma_i)\mathbf{B}\mathbf{b}_i \otimes \mathcal{K}^{-\top}(\mu_i)\mathbf{C}^\top\mathbf{c}_i\right)\right),\\
\operatorname{range}(\mathbf{V}) &= \mathcal{V}_{\mathrm{L}} + \mathcal{V}_{\mathrm{N}} + \mathcal{V}_{\mathrm{H}},\\
\operatorname{range}(\mathbf{W}) &= \mathcal{W}_{\mathrm{L}} + \mathcal{W}_{\mathrm{N}} + \mathcal{W}_{\mathrm{H}}.
\end{aligned}$$

*If a ROM is computed as shown in (7) using the projection matrices $\mathbf{V}$ and $\mathbf{W}$, where we assume $\mathbf{V}$ and $\mathbf{W}$ to be of full rank, then the following interpolation conditions are fulfilled:*

$$\mathbf{F}_{\mathrm{L}}(\sigma_i)\mathbf{b}_i = \widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i)\mathbf{b}_i, \tag{53a}$$

$$\mathbf{c}_i^\top\mathbf{F}_{\mathrm{L}}(\sigma_i) = \mathbf{c}_i^\top\widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i), \tag{53b}$$

$$\frac{d}{ds_1}\mathbf{c}_i^\top\mathbf{F}_{\mathrm{L}}(\sigma_i)\mathbf{b}_i = \frac{d}{ds_1}\mathbf{c}_i^\top\widehat{\mathbf{F}}_{\mathrm{L}}(\sigma_i)\mathbf{b}_i, \tag{53c}$$

$$\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m \otimes \mathbf{b}_i^{\textcircled{\eta}}\right) = \widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m \otimes \mathbf{b}_i^{\textcircled{\eta}}\right), \tag{53d}$$

$$\mathbf{c}_i^\top\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m^{\textcircled{2}} \otimes \mathbf{b}_i^{\textcircled{\eta-1}}\right) = \mathbf{c}_i^\top\widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m^{\textcircled{2}} \otimes \mathbf{b}_i^{\textcircled{\eta-1}}\right) \tag{53e}$$

$$\frac{\partial}{\partial s_j}\mathbf{c}_i^\top\mathbf{F}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m \otimes \mathbf{b}_i^{\textcircled{\eta}}\right) = \frac{\partial}{\partial s_j}\mathbf{c}_i^\top\widehat{\mathbf{F}}_{\mathrm{N}}^{(\eta)}(\sigma_i, \ldots, \sigma_i)\left(\mathbf{I}_m \otimes \mathbf{b}_i^{\textcircled{\eta}}\right), \tag{53f}$$

$$\mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i, \ldots, \sigma_i)\mathbf{b}_i^{\textcircled{\xi}} = \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i, \ldots, \sigma_i)\mathbf{b}_i^{\textcircled{\xi}}, \tag{53g}$$

$$\mathbf{c}_i^\top \mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i)\left(\mathbf{I}_m \otimes b_i^{\textcircled{\scriptsize{$\xi-1$}}}\right) = \mathbf{c}_i^\top \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i)\left(\mathbf{I}_m \otimes \mathbf{b}_i^{\textcircled{\scriptsize{$\xi-1$}}}\right), \tag{53h}$$

$$\frac{\partial}{\partial s_j}\mathbf{c}_i^\top \mathbf{F}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i)\mathbf{b}_i^{\textcircled{\scriptsize{$\xi$}}} = \frac{\partial}{\partial s_j}\mathbf{c}_i^\top \widehat{\mathbf{F}}_{\mathrm{H}}^{(\xi)}(\sigma_i,\ldots,\sigma_i)\mathbf{b}_i^{\textcircled{\scriptsize{$\xi$}}} \tag{53i}$$

*where* $i \in \{1,\ldots,\widetilde{r}\}$, $\xi \in \{2,\ldots,d\}$, $\eta \in \{1,\ldots,d\}$ *and* $\frac{\partial}{\partial s_j}$ *denotes the partial derivative with respect to* $s_j$ *of a given function.*

***Proof*** The proof of these interpolation conditions follows exactly the one of Theorem 3.1.

**Acknowledgements** We would like to express our gratitude to Dr. Steffen W. R. Werner for providing the data and code from [16, 17].

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of Interest** The authors declared that they have no conflict of interest.



## References

1. Ahmad, M.I., Benner, P., Jaimoukha, I.: Krylov subspace projection methods for model reduction of quadratic-bilinear systems. IET Control Theory Appl. **10**(16), 2010–2018 (2016). https://doi.org/10.1049/iet-cta.2016.0415
2. Antoulas, A.C., Beattie, C.A., Gugercin, S.: Interpolatory model reduction of large-scale dynamical systems. In: Mohammadpour, J., Grigoriadis, K.M. (eds.) Efficient Modeling and Control of Large-Scale Systems, pp. 3–58. Springer US (2010). https://doi.org/10.1007/978-1-4419-5757-3_1
3. Antoulas, A.C., Gosea, I.V., Ionita, A.C.: Model reduction of bilinear systems in the Loewner framework. SIAM J. Sci. Comput. **38**(5), B889–B916 (2016)
4. Barrault, M., Maday, Y., Nguyen, N.C., Patera, A.T.: An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. C. R. Math. Acad. Sci. Paris **339**(9), 667–672 (2004)
5. Baur, U., Benner, P., Feng, L.: Model order reduction for linear and nonlinear systems: a system-theoretic perspective. Arch. Comput. Methods Eng. **21**(4), 331–358 (2014). https://doi.org/10.1007/s11831-014-9111-2
6. Beattie, C.A., Gugercin, S.: Interpolatory projection methods for structure-preserving model reduction. Systems Control Lett. **58**(3), 225–232 (2009). https://doi.org/10.1016/j.sysconle.2008.10.016
7. Beddig, R.S., Benner, P., Dorschky, I., Reis, T., Schwerdtner, P., Voigt, M., Werner, S.W.R.: Structure-preserving model reduction for dissipative mechanical systems. Calm, Smooth and Smart: Novel

Approaches for Influencing Vibrations by Means of Deliberately Introduced Dissipation, pp. 209–230. Springer (2023)
8. Bellman, R.E., Cooke, K.L.: Differential-Difference Equations. Rand Corporation (1963)
9. Benner, P., Breiten, T.: Two-sided projection method for nonlinear model reduction. SIAM J. Sci. Comput. **37**(2), B239–B260 (2015). https://doi.org/10.1137/14097255X
10. Benner, P., Goyal, P.: Balanced truncation model order reduction for quadratic-bilinear systems. arXiv:1705.00160 (2017). https://arxiv.org/abs/1705.00160
11. Benner, P., Goyal, P.: Interpolation-based model order reduction for polynomial parametric systems. arXiv:1904.11891 (2019). https://arxiv.org/abs/1904.11891
12. Benner, P., Goyal, P.: Interpolation-based model order reduction for polynomial systems. SIAM J. Sci. Comp. **43**(1), A84–A108 (2021). https://doi.org/10.1137/19M1259171
13. Benner, P., Goyal, P., Gugercin, S.: $\mathcal{H}_2$-quasi-optimal model order reduction for quadratic-bilinear control systems. SIAM J. Matrix Anal. Appl. **39**(2), 983–1032 (2018). https://doi.org/10.1137/16M1098280
14. Benner, P., Goyal, P., Pontes Duff, I.: Identification of dominant subspaces for linear structured parametric systems and model reduction. arXiv:1910.13945 (2019). https://arxiv.org/abs/1910.13945
15. Benner, P., Gugercin, S., Werner, S.W.: Structured interpolation for multivariate transfer functions of quadratic-bilinear systems. Adv. Comput. Math. **50**(2), 18 (2024)
16. Benner, P., Gugercin, S., Werner, S.W.R.: Structure-preserving interpolation for model reduction of parametric bilinear systems. Automatica J. IFAC **132**, 109799 (2021). https://doi.org/10.1016/j.automatica.2021.109799
17. Benner, P., Gugercin, S., Werner, S.W.R.: Structure-preserving interpolation of bilinear control systems. Adv. Comput. Math. **47**(3), 43 (2021). https://doi.org/10.1007/s10444-021-09863-w
18. Benner, P., Schilders, W., Grivet-Talocia, S., Quarteroni, A., Rozza, G., Miguel Silveira, L.: Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms. De Gruyter (2020)
19. Breiten, T.: Structure-preserving model reduction for integro-differential equations. SIAM J. Cont. Optim. **54**(6), 2992–3015 (2016). https://doi.org/10.1137/15M1032296
20. Breiten, T., Damm, T.: Krylov subspace methods for model order reduction of bilinear control systems. Systems Control Lett. **59**(10), 443–450 (2010)
21. Carlberg, K., Farhat, C., Cortial, J., Amsallem, D.: The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows. J. Comput. Phys. **242**, 623–647 (2013)
22. Castañé-Selga, R., Lohmann, B., Eid, R.: Stability preservation in projection-based model order reduction of large-scale systems. Eur. J. Control. **18**(2), 122–132 (2012)
23. Chahlaoui, Y., Lemonnier, D., Vandendorpe, A., Van Dooren, P.: Second-order balanced truncation. Linear Algebra Appl. **415**(2–3), 373–384 (2006). https://doi.org/10.1016/j.laa.2004.03.032
24. Chaturantabut, S., Sorensen, D.C.: Nonlinear model reduction via discrete empirical interpolation. SIAM J. Sci. Comput. **32**(5), 2737–2764 (2010)
25. Chellappa, S., Feng, L., Benner, P.: An adaptive sampling approach for the reduced basis method. In: Beattie, C., Benner, P., Embree, M., Gugercin, S., Lefteriu, S. (eds.) Realization and Model Reduction of Dynamical Systems - A Festschrift in Honor of the 70th Birthday of Thanos Antoulas. Springer (2022). https://arxiv.org/abs/1910.00298
26. Eid, R., Salimbahrami, B., Lohmann, B., Rudnyi, E.B., Korvink, J.G.: Parametric order reduction of proportionally damped second-order systems. Sensors and Materials **19**(3), 149–164 (2007)
27. Gosea, I.V., Antoulas, A.C.: Data-driven model order reduction of quadratic-bilinear systems. Numer. Lin. Alg. Appl. **25**(6), e2200 (2018)
28. Gosea, I.V., Pontes Duff, I., Benner, P., Antoulas, A.C.: Model order reduction of bilinear time-delay systems. In: 18th European Control Conference (ECC), pp. 2289–2294 (2019). https://doi.org/10.23919/ECC.2019.8796085
29. Gubisch, M., Volkwein, S.: Proper orthogonal decomposition for linear–quadratic optimal control. In: Benner, P., Cohen, A., Ohlberger, M., Willcox, K. (eds.) Model Reduction and Approximation: Theory and Algorithms, pp. 3–64. SIAM (2017). https://doi.org/10.1137/1.9781611974829.ch1
30. Hale, J.K., Lunel, S.M.V.: Introduction to Functional Differential Equations, vol. 99. Springer Science & Business Media (2013)
31. Ionita, A.C., Antoulas, A.C.: Data-driven parametrized model reduction in the Loewner framework. SIAM J. Sci. Comput. **36**(3), A984–A1007 (2014). https://doi.org/10.1137/130914619

32. Jarlebring, E., Damm, T., Michiels, W.: Model reduction of time-delay systems using position balancing and delay Lyapunov equations. Math. Control Signals Systems **25**(2), 147–166 (2013)
33. Kolda, T.G.: Multilinear operators for higher-order decompositions. Tech. Rep. SAND2006-2081, Sandia National Laboratories (2006)
34. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. **51**(3), 455–500 (2009). https://doi.org/10.1137/07070111X
35. Quarteroni, A., Manzoni, A., Negri, F.: Reduced Basis Methods for Partial Differential Equations, La Matematica per il 3 + 2, vol. 92. Springer International Publishing (2016). ISBN: 978-3-319-15430-5
36. Reis, T., Stykel., T.: Balanced truncation model reduction of second-order systems. Math. Comput. Model. Dyn. Syst. **14**(5), 391–406 (2008)
37. Rodriguez, A.C., Gugercin, S., Borggaard, J.: Interpolatory model reduction of parameterized bilinear dynamical systems. Adv. Comput. Math. **44**(6), 1887–1916 (2018)
38. Rugh, W.J.: Nonlinear System Theory. The Johns Hopkins University Press, Baltimore, MD (1981)
39. Schulze, P., Unger, B., Beattie, C., Gugercin, S.: Data-driven structured realization. Linear Algebra Appl. **537**, 250–286 (2018)
40. Werner, S.W.R.: Structure-preserving model reduction for mechanical systems. Dissertation, Otto-von-Guericke-Universität, Magdeburg, Germany (2021). https://doi.org/10.25673/38617