Structure-Preserving Interpolation for Model Reduction of Parametric Bilinear Systems ⋆

Peter Benner [a,b], Serkan Gugercin [c], Steffen W. R. Werner [a]

[a] *Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany*

[b] *Otto von Guericke University, Faculty of Mathematics, Universitätsplatz 2, 39106 Magdeburg, Germany*

[c] *Department of Mathematics and Computational Modeling and Data Analytics Division, Academy of Integrated Science, Virginia Tech, Blacksburg, VA 24061, USA*

Abstract

In this paper, we present a projection-based interpolation framework for structure-preserving model order reduction of parametric bilinear dynamical systems. We introduce a general setting, covering a broad variety of different structures for parametric bilinear systems, and then provide conditions on projection spaces for the interpolation of structured subsystem transfer functions such that the system structure and parameter dependencies are preserved in the reduced-order model. Two benchmark examples with different parameter dependencies are used to demonstrate the theoretical analysis.

Key words: model order reduction; parametric bilinear systems; moment matching; structure-preserving approximation; structured parametric interpolation.

1 Introduction

Design and control processes usually involve simulating systems of differential equations describing the underlying dynamics. An important class of such systems are parametric bilinear time-invariant systems. These systems naturally appear in the modeling of thermal and mechanical systems (see Hu & Wang (2002), Mohler (1973)), plasma devices (see Ou (2010)), electrical circuits (see Al-Baiyat et al. (1993)), or medical applications (see Saputra et al. (2019)). They are also an important tool in the analysis of linear stochastic systems like in Benner & Damm (2011) and linear parameter-varying systems as in Benner & Breiten (2011). Bilinear systems also appear from in the linearization process of more general nonlinear systems using the Carleman linearization method; see Carleman (1932). In most cases, these bilinear systems have special structures resulting from the underlying physical model and the dynamics are parameter dependent. For example, in case of parametric bilinear mechanical systems, they have the form

$$\begin{aligned} M(\mu)\ddot{x}(t;\mu) + D(\mu)\dot{x}(t;\mu) + K(\mu)x(t;\mu) &= B_{\mathrm{u}}(\mu)u(t)\\ + \sum_{j=1}^{m} N_{\mathrm{p},j}(\mu)x(t;\mu)u_j(t) &+ \sum_{j=1}^{m} N_{\mathrm{v},j}(\mu)\dot{x}(t;\mu)u_j(t),\\ y(t;\mu) = C_{\mathrm{p}}(\mu)x(t;\mu) + C_{\mathrm{v}}(\mu)\dot{x}(t;\mu),& \end{aligned} \tag{1}$$

where $M(\mu)$, $D(\mu)$, $K(\mu)$, $N_{\mathrm{p},j}(\mu)$, $N_{\mathrm{v},j}(\mu) \in \mathbb{R}^{n\times n}$, for $j = 1,\ldots,m$; $B_{\mathrm{u}}(\mu) \in \mathbb{R}^{n\times m}$ and $C_{\mathrm{p}}(\mu), C_{\mathrm{v}}(\mu) \in \mathbb{R}^{p\times n}$ are constant matrices; and $\mu \in \mathbb{M} \subset \mathbb{R}^d$ represents the time-invariant parameters affecting the dynamics. In (1), $u(t) = \begin{bmatrix} u_1(t), \ u_2(t), \ \ldots, \ u_m(t)\end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^m$ denotes the inputs (forcing), $y(t;\mu) \in \mathbb{R}^p$ the outputs (measurements), and $x(t;\mu) \in \mathbb{R}^{n\times n}$ the internal variables. The parameter μ may represent variations in, e.g., material properties or system geometry.

Due to an increasing demand for accuracy in the modeling stage, systems as in (1) become larger and larger, e.g., $n > 10^6$, imposing overwhelming demands on computational resources like time and memory. The situation is even more prominent in the parametric problems we consider here due to the need to evaluate/simulate (1) for many samples of μ. The aim of parametric model order reduction is to construct a cheap-to-evaluate approximation of the input-to-output behavior of the original

⋆ This paper was not presented at any IFAC meeting. Corresponding author S. W. R. Werner.

Email addresses: benner@mpi-magdeburg.mpg.de (Peter Benner), gugercin@vt.edu (Serkan Gugercin), werner@mpi-magdeburg.mpg.de (Steffen W. R. Werner).

system by reducing the state-space dimension, i.e., the number of equations n, in such a way that the reduced model provides a high-fidelity approximation to the original one for the parameter range of interest. Additionally, the reduced-order model should have the same internal structure as well as the parameter dependencies as the original to retain the underlying physical structure. For example, for the system (1), the structure-preserving parametric reduced-order model will have the form

$$
\begin{aligned}
\widehat{M}(\mu)\ddot{\hat{x}}(t;\mu) + \widehat{D}(\mu)\dot{\hat{x}}(t;\mu) + \widehat{K}(\mu)\hat{x}(t;\mu) &= \widehat{B}_{\mathrm{u}}(\mu)u(t) \\
+ \sum_{j=1}^{m} \widehat{N}_{\mathrm{p},j}(\mu)\hat{x}(t;\mu)u_j(t) &+ \sum_{j=1}^{m} \widehat{N}_{\mathrm{v},j}(\mu)\dot{\hat{x}}(t;\mu)u_j(t), \\
\hat{y}(t;\mu) = \widehat{C}_{\mathrm{p}}(\mu)\hat{x}(t;\mu) &+ \widehat{C}_{\mathrm{v}}(\mu)\dot{\hat{x}}(t;\mu),
\end{aligned}
\tag{2}
$$

with $\widehat{M}(\mu), \widehat{D}(\mu), \widehat{K}(\mu), \widehat{N}_{\mathrm{p},j}(\mu), \widehat{N}_{\mathrm{v},j}(\mu) \in \mathbb{R}^{r\times r}$, for $j = 1,\ldots,m$, $\widehat{B}_{\mathrm{u}}(\mu) \in \mathbb{R}^{r\times m}$, $\widehat{C}_{\mathrm{p}}(\mu), \widehat{C}_{\mathrm{v}}(\mu) \in \mathbb{R}^{p\times r}$, and $r \ll n$. Note that the reduced-order model (2) has the same structure as (1) and can be interpreted as a physically meaningful reduced-order mechanical system. The structure preservation can also be very beneficial in terms of computational speed and accuracy; see, e.g., Benner et al. (2021).

Model reduction for linear and general nonlinear (parametric) systems has been studied heavily, especially over the last three decades, using a variety of approaches; see, e.g., Benner et al. (2017), Quarteroni & Rozza (2014), Scarciotti & Astolfi (2017), Schilders et al. (2008). In recent years, the class of bilinear control systems received additional focus as an important link between linear and nonlinear systems, since they only involve the multiplication of states and inputs as nonlinearities. In this paper, we will concentrate on structure-preserving model reduction for parametric bilinear systems. For parametric *unstructured (classical)* bilinear systems, i.e., for systems of the form

$$
\begin{aligned}
E(\mu)\dot{x}(t;\mu) &= A(\mu)x(t;\mu) + B(\mu)u(t) \\
&\quad + \sum_{j=1}^{m} N_j(\mu)x(t;\mu)u_j(t), \\
y(t;\mu) &= C(\mu)x(t;\mu),
\end{aligned}
\tag{3}
$$

the interpolatory parametric model reduction framework was developed in Rodriguez et al. (2018) by synthesizing the interpolation theory for parametric linear dynamical systems from, e.g., Antoulas et al. (2020), Baur et al. (2011), with the subsystem interpolation approaches for bilinear systems; see Antoulas et al. (2020), Bai & Skoogh (2006), Breiten & Damm (2010), Condon & Ivanov (2007). There are other approaches to model reduction of unstructured bilinear systems, e.g., bilinear balanced truncation from Al-Baiyat et al. (1993), Benner & Damm (2011), Hsu et al. (1983), Volterra series interpolation as in Benner & Breiten (2012), Flagg & Gugercin (2015), Zhang & Lam (2002), or the bilinear Loewner framework from Antoulas et al. (2016), Gosea et al. (2019). Those approaches do neither provide extensions for the parametric bilinear system case nor have extensions for structured systems and, therefore, will not be further discussed in this paper. Recently in Benner et al. (2021), the structured interpolation framework of Beattie & Gugercin (2009) for linear dynamical systems has been extended to the case of *non-parametric* structured bilinear systems. In this paper, we will extend this interpolation theory to the case of *structured parametric* bilinear systems by developing the subspace conditions to be enforced in the projection-based model reduction framework.

In Section 2, we introduce basic mathematical concepts and notation. We prove the structure-preserving interpolation framework for parametric bilinear systems in Section 3. The established theory is then extended in Section 4 to the interpolation of parameter sensitivities. Section 5 illustrates the analysis in two numerical benchmark examples, followed by conclusions in Section 6.

2 Mathematical preliminaries

For a complex-valued matrix $X \in \mathbb{C}^{n\times m}$, $X^{\mathsf{H}} := \overline{X}^{\mathsf{T}}$ will denote its conjugate transpose. Given two matrices $A \in \mathbb{C}^{n\times m}$ and $B \in \mathbb{C}^{p\times q}$, $(A \otimes B)$ will denote the Kronecker product, i.e.,

$$
A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix} \in \mathbb{C}^{np\times mq},
$$

where a_{ij} is the (i,j)-th element of A.

Under some mild assumptions, the output of the bilinear system (3) can be rewritten in terms of a Volterra series, i.e.,

$$
\begin{aligned}
y(t;\mu) = \sum_{k=1}^{\infty} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} g_k(t_1,\ldots,t_k,\mu) \\
\times \left(u(t - \sum_{i=1}^{k} t_i) \otimes \cdots \otimes u(t-t_1) \right) \mathrm{d}t_k \cdots \mathrm{d}t_1,
\end{aligned}
$$

where g_k denotes the k-th regular Volterra kernel; see, e.g., Rugh (1981). Using the multivariate Laplace transformation from Rugh (1981), the regular Volterra kernels yield the frequency representation (4), as the k-th regular subsystem transfer function of (3), where $N(\mu) = \begin{bmatrix} N_1(\mu), & \ldots, & N_m(\mu) \end{bmatrix}$. The model reduction theory in Rodriguez et al. (2018) is based on the interpolation of (4), i.e., unstructured (classical) parametric subsystems.

In this paper, we consider a much more general setting of multivariate transfer functions. The interpolation

$$G_k(s_1,\ldots,s_k,\mu) = C(\mu)(s_kE(\mu)-A(\mu))^{-1}\left(\prod_{j=1}^{k-1}(I_{m^{j-1}}\otimes N(\mu))(I_{m^j}\otimes(s_{k-j}E(\mu)-A(\mu))^{-1})\right)(I_{m^{k-1}}\otimes B(\mu)),\quad k\geq 1 \tag{4}$$

$$G_k(s_1,\ldots,s_k,\mu) = \mathcal{C}(s_k,\mu)\mathcal{K}(s_k,\mu)^{-1}\left(\prod_{j=1}^{k-1}(I_{m^{j-1}}\otimes\mathcal{N}(s_{k-j},\mu))(I_{m^j}\otimes\mathcal{K}(s_{k-j},\mu)^{-1})\right)(I_{m^{k-1}}\otimes\mathcal{B}(s_1,\mu)),\quad k\geq 1 \tag{5}$$

$$\widehat{G}_k(s_1,\ldots,s_k,\mu) = \widehat{\mathcal{C}}(s_k,\mu)\widehat{\mathcal{K}}(s_k,\mu)^{-1}\left(\prod_{j=1}^{k-1}(I_{m^{j-1}}\otimes\widehat{\mathcal{N}}(s_{k-j},\mu))(I_{m^j}\otimes\widehat{\mathcal{K}}(s_{k-j},\mu)^{-1})\right)(I_{m^{k-1}}\otimes\widehat{\mathcal{B}}(s_1,\mu)),\quad k\geq 1 \tag{6}$$

of structured transfer functions for linear systems was developed in Beattie & Gugercin (2009) and then extended to the parametric setting in Antoulas et al. (2010). As the structured transfer functions were recently extended to non-parametric bilinear systems in Benner et al. (2021), we consider here *structured parametric multivariate transfer functions* of the form (5) with frequency points $s_1,\ \ldots,\ s_k \in \mathbb{C}$, parameters $\mu \in \mathbb{M} \subset \mathbb{R}^d$, $\mathcal{N}(s,\mu) = \left[\mathcal{N}_1(s,\mu),\ \ldots,\ \mathcal{N}_m(s,\mu)\right]$, and matrix functions

$$\begin{aligned} \mathcal{C}&\colon \mathbb{C}\times\mathbb{M}\to\mathbb{C}^{p\times n}, & \mathcal{K}&\colon \mathbb{C}\times\mathbb{M}\to\mathbb{C}^{n\times n},\\ \mathcal{B}&\colon \mathbb{C}\times\mathbb{M}\to\mathbb{C}^{n\times m}, & \mathcal{N}_j&\colon \mathbb{C}\times\mathbb{M}\to\mathbb{C}^{n\times n}, \end{aligned}$$

for $j = 1,\ldots,m$. For the parametric bilinear mechanical system (1), these matrix functions are realized by

$$\begin{aligned} \mathcal{K}(s,\mu) &= s^2M(\mu) + sD(\mu) + K(\mu),\\ \mathcal{N}_j(s,\mu) &= N_{\mathrm{p},j}(\mu) + sN_{\mathrm{v},j}(\mu) \text{ for } j=1,\ldots,m,\\ \mathcal{B}(s,\mu) &= B_{\mathrm{u}}(\mu),\ \text{and } \mathcal{C}(s,\mu) = C_{\mathrm{p}}(\mu) + sC_{\mathrm{v}}(\mu). \end{aligned}$$

The reduced-order models are then computed by projection: Given model reduction bases $V, W \in \mathbb{C}^{n\times r}$, the reduced-order model $\widehat{G}$ is described by the reduced-order matrix functions

$$\begin{aligned} \widehat{\mathcal{C}}(s,\mu) &= \mathcal{C}(s,\mu)V, & \widehat{\mathcal{K}}(s,\mu) &= W^{\mathsf{H}}\mathcal{K}(s,\mu)V,\\ \widehat{\mathcal{B}}(s,\mu) &= W^{\mathsf{H}}\mathcal{B}(s,\mu), & \widehat{\mathcal{N}}_j(s,\mu) &= W^{\mathsf{H}}\mathcal{N}_j(s,\mu)V, \end{aligned} \tag{7}$$

for $j = 1,\ldots,m$. Model reduction by projection in the sense of (7) is structure-preserving by nature. In general, every matrix-valued function can be affinely decomposed with respect to its arguments, here frequency and parameter, and we can write

$$\mathcal{K}(s,\mu) = \sum_{j=1}^{n_{\mathcal{K}}} h_{\mathcal{K},j}(s,\mu)\mathcal{K}_j, \tag{8}$$

where $h_{\mathcal{K},j}\colon \mathbb{C}\times\mathbb{M}\to\mathbb{C}$ are scalar functions depending on frequency and parameter, and $\mathcal{K}_j \in \mathbb{C}^{n\times n}$ are constant matrices, for $j = 1,\ldots,n_{\mathcal{K}}$. In the worst-case scenario, we have $n_{\mathcal{K}} = n^2$ and the $\mathcal{K}_j$'s are elementary matrices. However, we are interested in cases where $n_{\mathcal{K}} \ll n$, which is true in most applications. In the numerical examples we present in Section 5, $n_{\mathcal{K}}$ is at most 3. The choice of the scalar functions $h_{\mathcal{K},j}$ encodes the internal structure of the system. Using the affine decomposition, the reduced-order matrix function is given by

$$\widehat{\mathcal{K}}(s,\mu) = \sum_{j=1}^{n_{\mathcal{K}}} h_{\mathcal{K},j}(s,\mu)W^{\mathsf{H}}\mathcal{K}_jV = \sum_{j=1}^{n_{\mathcal{K}}} h_{\mathcal{K},j}(s,\mu)\widehat{\mathcal{K}}_j.$$

This works analogously for the other matrix functions in (7), which gives a computable realization of the reduced-order model. Since the functions $h_{\mathcal{K},j}$ stay unchanged, the internal structure and parameter dependency of the original matrix functions (and thus of the original system) are retained. The reduced-order model is then given by replacing the original system matrices in the affine decomposition (8) by their reduced-order counterparts.

In the following, we will use an abbreviation for the notion of partial derivatives, namely we denote

$$\partial_{s_1^{j_1}\cdots s_k^{j_k}} f(z_1,\ldots,z_k) := \frac{\partial^{j_1+\ldots+j_k} f}{\partial s_1^{j_1}\cdots\partial s_k^{j_k}}(t_1,\ldots,t_k),$$

for the differentiation of an analytic function $f\colon \mathbb{C}^k \to \mathbb{C}^\ell$ with respect to the variables $s_1,\ldots,s_k$ and evaluated at $z_1,\ \ldots,\ z_k$. Also, we denote the vertical concatenation of the bilinear terms by $\widetilde{\mathcal{N}}(s,\mu) = \begin{bmatrix}\mathcal{N}_1(s,\mu)\\ \vdots\\ \mathcal{N}_m(s,\mu)\end{bmatrix}$.

3 Structured interpolation

Interpolatory model reduction has been one of the most commonly used and effective approaches to model reduction and shown to provide locally optimal reduced models for linear, bilinear, quadratic-bilinear dynamical systems; we refer the reader to Antoulas et al. (2020), Baur et al. (2014), Scarciotti & Astolfi (2017) and references therein for details on interpolatory model reduction for linear and nonlinear systems. In this setting, one chooses V and W in (7) such that the reduced-order transfer functions interpolate the transfer functions of the original system at selected points. In the setting of parametric structured multivariate transfer functions G_k in (5), we want to construct V and W such that the reduced transfer functions $\widehat{G}_k$ in (6) satisfy

$$G_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) = \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) \quad \text{and} \tag{9}$$
$$\nabla G_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) = \nabla \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}), \tag{10}$$

for given frequency interpolation points $\sigma_1, \ldots, \sigma_k \in \mathbb{C}$, the parameter interpolation point $\hat{\mu} \in \mathbb{M}$, and where ∇G_k denotes the Jacobian matrix

$$\nabla G_k = \left[\partial_{s_1} G_k, \; \ldots, \; \partial_{s_k} G_k, \; \partial_{\mu_1} G_k, \; \ldots, \; \partial_{\mu_d} G_k\right].$$

We emphasize that for multi-input/multi-output (MIMO) systems we consider here, transfer functions G_k are matrix valued. Therefore, conditions in (9) and (10) enforce matrix interpolation. This is not usually needed. For MIMO linear dynamical systems, for example, one enforces tangential interpolation, meaning matrix-interpolation along selected directions; see, e.g., Antoulas et al. (2020). However, for brevity and to keep the notation concise, we will focus on matrix interpolation.

Even though we have only listed two sets of interpolation conditions in (9) and (10), Theorems 1 and 2 below will show how to construct V and W to enforce interpolation for more general cases, including higher-order partial derivatives. The recent work in Benner et al. (2021) showed how to enforce (9) and (10) for *non-parametric* structured bilinear systems. Our theory below will extend these results to the parametric case. Note that the first condition (9) does not involve any differentiation with respect to the parameter $\hat{\mu}$ and can be viewed as interpolation for a fixed parameter $\mu = \hat{\mu}$. Therefore, we might expect that the subspace constructions from Benner et al. (2021) for the non-parametric problem might yield the desired subspaces. This is indeed what we discuss first in Theorems 1 and 2. However, the second condition (10) involves matching sensitivity with respect to the parameter as well, which will be discussed in Section 4.

Theorem 1 (Structured matrix interpolation) *Let G be a parametric bilinear system, with its structured subsystem transfer functions G_k in* (5)*, and $\widehat{G}$ be the reduced-order parametric bilinear system, constructed as in* (7) *with its subsystem transfer functions $\widehat{G}_k$ in* (6)*. Let the matrix functions $\mathcal{C}(s,\mu)$, $\mathcal{K}(s,\mu)^{-1}$, $\mathcal{N}(s,\mu)$, $\mathcal{B}(s,\mu)$, and $\widehat{\mathcal{K}}(s,\mu)^{-1}$ be defined for given sets of frequency interpolation points $\sigma_1, \ldots, \sigma_k \in \mathbb{C}$ and $\varsigma_1, \ldots, \varsigma_\theta \in \mathbb{C}$, and the parameter interpolation point $\hat{\mu} \in \mathbb{M}$.*

(a) If V is constructed such that

$$\operatorname{span}(V) \supseteq \operatorname{span}([V_1, \ldots, V_k]),$$

where

$$\begin{aligned} V_1 &= \mathcal{K}(\sigma_1, \hat{\mu})^{-1}\mathcal{B}(\sigma_1, \hat{\mu}) \text{ and} \\ V_j &= \mathcal{K}(\sigma_j, \hat{\mu})^{-1}\mathcal{N}(\sigma_{j-1}, \hat{\mu})(I_m \otimes V_{j-1}), \end{aligned} \tag{11}$$

for $2 \leq j \leq k$, then the following interpolation conditions hold true:

$$G_j(\sigma_1, \ldots, \sigma_j, \hat{\mu}) = \widehat{G}_j(\sigma_1, \ldots, \sigma_j, \hat{\mu}), \tag{12}$$

for $j = 1, \ldots, k$.

(b) If W is constructed such that

$$\operatorname{span}(W) \supseteq \operatorname{span}([W_1, \ldots, W_\theta]),$$

where

$$\begin{aligned} W_1 &= \mathcal{K}(\varsigma_\theta, \hat{\mu})^{-\mathsf{H}}\mathcal{C}(\varsigma_\theta, \hat{\mu})^{\mathsf{H}} \text{ and} \\ W_i &= \mathcal{K}(\varsigma_{\theta-i+1}, \hat{\mu})^{-\mathsf{H}}\widetilde{\mathcal{N}}(\varsigma_{\theta-i+1}, \hat{\mu})^{\mathsf{H}}(I_m \otimes W_{i-1}), \end{aligned}$$

for $2 \leq i \leq \theta$, then the following interpolation conditions hold true:

$$G_i(\varsigma_{\theta-i+1}, \ldots, \varsigma_\theta, \hat{\mu}) = \widehat{G}_i(\varsigma_{\theta-i+1}, \ldots, \varsigma_\theta, \hat{\mu}), \tag{13}$$

for $i = 1, \ldots, \theta$.

(c) Let V be constructed as in Part (a) and W as in Part (b). Then, in addition to (12) *and* (13)*, the interpolation conditions*

$$\begin{aligned} &G_{q+\eta}(\sigma_1, \ldots, \sigma_q, \varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu}) \\ &= \widehat{G}_{q+\eta}(\sigma_1, \ldots, \sigma_q, \varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu}), \end{aligned} \tag{14}$$

hold for $1 \leq q \leq k$ and $1 \leq \eta \leq \theta$.

PROOF. Given the fixed parameter $\hat{\mu} \in \mathbb{M}$, the matrix functions $\mathcal{C}(s,\hat{\mu})$, $\mathcal{K}(s,\hat{\mu})$, $\mathcal{N}(s,\hat{\mu})$ and $\mathcal{B}(s,\hat{\mu})$ can be viewed as the realization of a non-parametric bilinear system. Then, the interpolation conditions (12)–(14) can be considered as subsystem interpolation of a non-parametric bilinear system as these conditions do not involve any variation/sensitivity with respect to μ. Therefore, the subspace conditions in (Benner et al. 2021, Theorem 8), for interpolating a non-parametric structured bilinear system, apply here as well, which are precisely the subspace conditions listed in Parts (a)–(c). However, to make the paper self-contained and the proof of Theorem 4 in Section 4 easier to follow, we will still prove Part (a) for $k = 2$. By induction over k, the rest of the result in (a) follows directly using the same arguments. Using (6), the second reduced-order transfer function is given by

$$\begin{aligned} \widehat{G}_2(\sigma_1, \sigma_2, \hat{\mu}) = \widehat{\mathcal{C}}(\sigma_2, \hat{\mu})\widehat{\mathcal{K}}(\sigma_2, \hat{\mu})^{-1}\widehat{\mathcal{N}}(\sigma_1, \hat{\mu}) \\ \times (I_m \otimes \widehat{\mathcal{K}}(\sigma_1, \hat{\mu})^{-1})(I_m \otimes \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})). \end{aligned}$$

We observe that with (7) it holds

$$
\begin{aligned}
&(I_m \otimes V)(I_m \otimes \widehat{\mathcal{K}}(\sigma_1, \hat{\mu})^{-1})(I_m \otimes \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})) \\
&= (I_m \otimes \underbrace{V\widehat{\mathcal{K}}(\sigma_1, \hat{\mu})^{-1} W^{\mathsf{H}} \mathcal{K}(\sigma_1, \hat{\mu})}_{P_{\mathrm{V}_1}} \underbrace{\mathcal{K}(\sigma_1, \hat{\mu})^{-1} \mathcal{B}(\sigma_1, \hat{\mu})}_{V_1}),
\end{aligned}
$$

where P_{V_1} is a projector onto $\operatorname{span}(V)$ and V_1 is as defined in (11). By construction, we have $\operatorname{span}(V_1) \subseteq \operatorname{span}(V)$; thus $P_{\mathrm{V}_1} V_1 = V_1$ and, therefore

$$
\begin{aligned}
&(I_m \otimes V)(I_m \otimes \widehat{\mathcal{K}}(\sigma_1, \hat{\mu})^{-1})(I_m \otimes \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})) \\
&= (I_m \otimes \mathcal{K}(\sigma_1, \hat{\mu})^{-1})(I_m \otimes \mathcal{B}(\sigma_1, \hat{\mu})).
\end{aligned}
$$

Then, $\widehat{G}_2$ can be written as

$$
\begin{aligned}
\widehat{G}_2(\sigma_1, \sigma_2, \hat{\mu}) &= \mathcal{C}(\sigma_2, \hat{\mu}) V \widehat{\mathcal{K}}(\sigma_2, \hat{\mu})^{-1} W^{\mathsf{H}} \mathcal{N}(\sigma_1, \hat{\mu}) \\
&\quad \times (I_m \otimes V_1).
\end{aligned}
$$

Also, it holds that

$$
\begin{aligned}
&V\widehat{\mathcal{K}}(\sigma_2, \hat{\mu})^{-1} W^{\mathsf{H}} \mathcal{N}(\sigma_1, \hat{\mu})(I_m \otimes V_1) \\
&= \underbrace{V\widehat{\mathcal{K}}(\sigma_2, \hat{\mu})^{-1} W^{\mathsf{H}} \mathcal{K}(\sigma_2, \hat{\mu})}_{P_{\mathrm{V}_2}} \\
&\quad \times \underbrace{\mathcal{K}(\sigma_2, \hat{\mu})^{-1} \mathcal{N}(\sigma_1, \hat{\mu})(I_m \otimes V_1)}_{V_2} \\
&= \mathcal{K}(\sigma_2, \hat{\mu})^{-1} \mathcal{N}(\sigma_1, \hat{\mu})(I_m \otimes V_1),
\end{aligned}
$$

using the fact that P_{V_2} is another projector onto $\operatorname{span}(V)$ and that $\operatorname{span}(V_2) \subseteq \operatorname{span}(V)$. Inserting this last equality into the second reduced-order transfer function yields

$$
\widehat{G}_2(\sigma_1, \sigma_2, \hat{\mu}) = G_2(\sigma_1, \sigma_2, \hat{\mu}).
$$

Constructing further projectors onto $\operatorname{span}(V)$ for higher-order transfer functions gives the result in (a). The result in Part (b) follows exactly the same way by using the Hermitian transposed matrix functions and constructing now projectors onto $\operatorname{span}(W)$. Part (c) is then resulting from the application of both types of projectors onto $\operatorname{span}(V)$ and $\operatorname{span}(W)$.

In practice, one would construct the final basis matrices V and W via a rank-revealing orthogonalization of the concatenation, e.g., $[V_1, \ldots, V_k]$. This could be done, for example, via a rank-revealing QR decomposition or SVD afterwards, or by a repeated re-orthogonalization process in every step after each computation of the next V_i. This yields basis matrices with orthonormal columns and reveals rank deficiency in the constructed matrices, leading to smaller subspace dimension and thus a smaller reduced order.

In Theorem 1, only function values are matched, i.e., the zeroth derivative. The following theorem extends these results to matching higher-order derivatives in the frequency arguments, i.e., to enforcing Hermite interpolation conditions.

Theorem 2 (Hermite matrix interpolation) *Let G be a parametric bilinear system, with its structured subsystem transfer functions G_k in (5) and $\widehat{G}$ be the reduced-order parametric bilinear system, constructed as in (7) with its subsystem transfer functions $\widehat{G}_k$ in (6). Let the matrix functions $\mathcal{C}(s, \mu)$, $\mathcal{K}(s, \mu)^{-1}$, $\mathcal{N}(s, \mu)$, $\mathcal{B}(s, \mu)$, and $\widehat{\mathcal{K}}(s, \mu)^{-1}$ be analytic for given sets of frequency interpolation points $\sigma_1, \ldots, \sigma_k \in \mathbb{C}$ and $\varsigma_1, \ldots, \varsigma_\theta \in \mathbb{C}$, and the parameter interpolation point $\hat{\mu} \in \mathbb{M}$.*

(a) If V is constructed such that

$$
\operatorname{span}(V) \supseteq \operatorname{span}([V_{1,0}, \ldots, V_{k,\ell_k}]),
$$

where

$$
\begin{aligned}
V_{1,j_1} &= \partial_{s^{j_1}} (\mathcal{K}^{-1}\mathcal{B})(\sigma_1, \hat{\mu}) \text{ and} \\
V_{q,j_q} &= \partial_{s^{j_q}} \mathcal{K}^{-1}(\sigma_q, \hat{\mu}) \\
&\quad \times \left(\prod_{j=1}^{q-2} \partial_{s^{\ell_{q-j}}} \big((I_{m^{j-1}} \otimes \mathcal{N}) \right. \\
&\quad \left. \times (I_{m^j} \otimes \mathcal{K}) \big)(\sigma_{q-j}, \hat{\mu}) \right) \\
&\quad \times \partial_{s^{\ell_1}} \big((I_{m^{q-2}} \otimes \mathcal{N})(I_{m^{q-1}} \otimes \mathcal{K}) \\
&\quad \times (I_{m^{q-1}} \otimes \mathcal{B}) \big)(\sigma_1, \hat{\mu}) \,,
\end{aligned}
$$

for $2 \leq q \leq k$ and $0 \leq j_1 \leq \ell_1$; $0 \leq j_q \leq \ell_q$, then the following interpolation conditions hold true:

$$
\begin{aligned}
&\partial_{s_1^{\ell_1} \cdots s_{q-1}^{\ell_{q-1}} s_q^{j_q}} G_q(\sigma_1, \ldots, \sigma_q, \hat{\mu}) \\
&= \partial_{s_1^{\ell_1} \cdots s_{q-1}^{\ell_{q-1}} s_q^{j_q}} \widehat{G}_q(\sigma_1, \ldots, \sigma_q, \hat{\mu}),
\end{aligned}
\tag{15}
$$

for $q = 1, \ldots, k$ and $j_q = 0, \ldots, \ell_q$.

(b) If W is constructed such that

$$
\operatorname{span}(W) \supseteq \operatorname{span}([W_{1,0}, \ldots, W_{\theta,\nu_\theta}]),
$$

where

$$
\begin{aligned}
W_{1,i_\theta} &= \partial_{s^{i_\theta}} (\mathcal{K}^{-\mathsf{H}} \mathcal{C}^{\mathsf{H}})(\varsigma_\theta, \hat{\mu}) \text{ and} \\
W_{\eta, i_{\theta-\eta+1}} &= \partial_{s^{i_{\theta-\eta+1}}} (\mathcal{K}^{-\mathsf{H}} \widetilde{\mathcal{N}}^{\mathsf{H}})(\varsigma_{\theta-\eta+1}, \hat{\mu}) \\
&\quad \times \left(\prod_{i=\theta-\eta+2}^{\theta-1} \partial_{s^{\nu_i}} (I_{m^{i-1}} \otimes \mathcal{K}^{-\mathsf{H}} \widetilde{\mathcal{N}}^{\mathsf{H}})(\varsigma_i, \hat{\mu}) \right) \\
&\quad \times \left(I_{m^{\theta-1}} \otimes \partial_{s^{\nu_\theta}} (\mathcal{K}^{-\mathsf{H}} \mathcal{C}^{\mathsf{H}})(\varsigma_\theta, \hat{\mu}) \right),
\end{aligned}
$$

for $2 \leq \eta \leq \theta$ and $0 \leq i_\theta \leq \nu_\theta$; $0 \leq i_{\theta-\eta+1} \leq \nu_{\theta-\eta+1}$, then the following interpolation conditions hold true:

$$\partial_{s_1^{i_{\theta-\eta+1}} s_2^{\nu_{\theta-\eta+2}} \cdots s_\theta^{\nu_\theta}} G_\eta(\varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu}) = \partial_{s_1^{i_{\theta-\eta+1}} s_2^{\nu_{\theta-\eta+2}} \cdots s_\theta^{\nu_\theta}} \widehat{G}_\eta(\varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu}), \tag{16}$$

for $\eta = 1, \ldots, \theta$ and $i_{\theta-\eta+1} = 0, \ldots, \nu_{\theta-\eta+1}$.

(c) *Let V be constructed as in (a) and W as in (b). Then, in addition to (15) and (16), the interpolation conditions (17) hold for $j_q = 0, \ldots, \ell_q$; $i_{\theta-\eta+1} = 0, \ldots, \nu_{\theta-\eta+1}$; $1 \leq q \leq k$ and $1 \leq \eta \leq \theta$.*

PROOF. As in Theorem 1, all the interpolation conditions are for a fixed parameter $\hat{\mu} \in \mathbb{M}$, i.e., they can be proven using a similar construction of projectors onto suitable subspaces as in Theorem 1. Therefore, the subspace conditions in (Benner et al. 2021, Theorem 9) can be applied here, which are precisely the subspace conditions listed in Theorem 2.

Remark 3 *We note that the major computational cost stems from solving (sparse) linear systems of equations to construct the basis matrices V and W, as common to the interpolatory model reduction framework in general.*

4 Matching parameter sensitivities

So far, the interpolation conditions enforced did not show variability with respect to the parameter μ. Even in the Hermite conditions matched in Theorem 2, the matched derivatives (sensitivities) are with respect to the frequency points. This enabled us to directly employ the conditions and analysis from Benner et al. (2021). However, for parametric systems it is important to match the parameter sensitivity with respect to the parameter variation as well. This is what we establish in the next result, extending the similar results from linear dynamics in Baur et al. (2011) and unstructured bilinear dynamics in Rodriguez et al. (2018) to the new parametric structured framework. An important conclusion is that the parameter sensitivity is matched implicitly, i.e., without ever explicitly computing it. This is achieved by using the same set of frequency interpolation points for V and W.

Theorem 4 (Two-sided matrix interpolation) *Let G be a parametric bilinear system, with its structured subsystem transfer functions G_k in (5) and $\widehat{G}$ be the reduced-order parametric bilinear system, constructed as in (7) with its subsystem transfer functions $\widehat{G}_k$ in (6). Let the matrix functions $\mathcal{C}(s,\mu)$, $\mathcal{K}(s,\mu)^{-1}$, $\mathcal{N}(s,\mu)$, $\mathcal{B}(s,\mu)$, and $\widehat{\mathcal{K}}(s,\mu)^{-1}$ be analytic for a given set of frequency interpolation points $\sigma_1, \ldots, \sigma_k \in \mathbb{C}$ and the parameter interpolation point $\hat{\mu} \in \mathbb{M}$.*

(a) *Let V be constructed as in Theorem 1 Part (a) and W be constructed as in Theorem 1 Part (b) with $\varsigma_i = \sigma_i$ for $i = 1, 2, \ldots, k$. Then, in addition to (12)–(14) it holds*

$$\nabla G_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) = \nabla \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}). \tag{18}$$

(b) *Let V be constructed as in Theorem 2 Part (a) and W be constructed as in Theorem 2 Part (b) with $\varsigma_i = \sigma_i$ for $i = 1, 2, \ldots, k$. Then, in addition to (15)–(17), it holds*

$$\nabla \left(\partial_{s_1^{\ell_1} \cdots s_k^{\ell_k}} G_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) \right) = \nabla \left(\partial_{s_1^{\ell_1} \cdots s_k^{\ell_k}} \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) \right). \tag{19}$$

PROOF. For brevity, we only prove (18). The proof of (19) follows analogously. As in the proof of Theorem 1, we will construct appropriate projectors onto the projection spaces span(V) or span(W). In contrast to Theorem 2, we now also interpolate the derivative with respect to the parameters. Using the product rule, the partial derivative of $\widehat{G}_k$ with respect to a single parameter entry μ_i, for $1 \leq i \leq d$, is given by

$$\begin{aligned} &\partial_{\mu_i} \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) \\ &= \sum_{\alpha \in \mathbb{A}} \left(\partial_{\mu_i^{\alpha_1}} \widehat{\mathcal{C}}(\sigma_k, \hat{\mu}) \right) \left(\partial_{\mu_i^{\alpha_2}} \widehat{\mathcal{K}}^{-1}(\sigma_k, \hat{\mu}) \right) \\ &\quad \times \left(\prod_{j=1}^{k-1} (I_{m^{j-1}} \otimes \partial_{\mu_i^{\alpha_{2j+1}}} \widehat{\mathcal{N}}(\sigma_{k-j}, \hat{\mu})) \right. \\ &\quad \left. \times (I_{m^j} \otimes \partial_{\mu_i^{\alpha_{2j+2}}} \widehat{\mathcal{K}}^{-1}(\sigma_{k-j}, \hat{\mu})) \right) \\ &\quad \times (I_{m^{k-1}} \otimes \partial_{\mu_i^{\alpha_{2k+1}}} \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})), \end{aligned} \tag{20}$$

where $\mathbb{A}$ denotes the set of all columns of the identity matrix of size $2k+1$. In other words, (20) is a sum of $2k+1$ terms where each term corresponds to the vector α taking a value from this set of columns. Therefore, in each term only a single matrix function is differentiated. We will show that every single term in the sum (20) matches the same term in the full-order model, thus, summed together, proving the desired interpolation property (18). Consider, e.g., the second term in (20), i.e., the term in which α is the second column of the identity matrix: $\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{2k+1} \end{bmatrix}^\mathsf{T} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \end{bmatrix}^\mathsf{T}$. Denote the corresponding term by $\widehat{\mathcal{H}}_2$. Then,

$$\begin{aligned} \widehat{\mathcal{H}}_2 &:= \widehat{\mathcal{C}}(\sigma_k, \hat{\mu}) \left(\partial_{\mu_i} \widehat{\mathcal{K}}^{-1}(\sigma_k, \hat{\mu}) \right) \\ &\quad \times \left(\prod_{j=1}^{k-1} (I_{m^{j-1}} \otimes \widehat{\mathcal{N}}(\sigma_{k-j}, \hat{\mu})) \right. \end{aligned}$$

$$
\begin{aligned}
&\partial_{s_1^{\ell_1} \cdots s_{q-1}^{\ell_{q-1}} s_q^{j_q} s_{q+1}^{i_{\theta-\eta+1}} s_{q+2}^{\nu_{\theta-\eta+2}} \cdots s_{q+\eta}^{\nu_\theta}} G_{q+\eta}(\sigma_1, \ldots, \sigma_q, \varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu}) \\
&= \partial_{s_1^{\ell_1} \cdots s_{q-1}^{\ell_{q-1}} s_q^{j_q} s_{q+1}^{i_{\theta-\eta+1}} s_{q+2}^{\nu_{\theta-\eta+2}} \cdots s_{q+\eta}^{\nu_\theta}} \widehat{G}_{q+\eta}(\sigma_1, \ldots, \sigma_q, \varsigma_{\theta-\eta+1}, \ldots, \varsigma_\theta, \hat{\mu})
\end{aligned} \tag{17}
$$

$$
\times \left(I_{m^j} \otimes \widehat{\mathcal{K}}(\sigma_{k-j}, \hat{\mu})^{-1}\right)\Bigg) (I_{m^{k-1}} \otimes \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})).
$$

The derivative of the inverse appearing in $\widehat{\mathcal{H}}_2$ is given by

$$
\partial_{\mu_i} \widehat{\mathcal{K}}^{-1}(\sigma_k, \hat{\mu}) = -\widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-1} \Big(\partial_{\mu_i} \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})\Big) \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-1}.
$$

Therefore, $\widehat{\mathcal{H}}_2$ can be rewritten as

$$
\begin{aligned}
\widehat{\mathcal{H}}_2 &= -\widehat{\mathcal{C}}(\sigma_k, \hat{\mu}) \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-1} \Big(\partial_{\mu_i} \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})\Big) \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-1} \\
&\quad \times \Bigg(\prod_{j=1}^{k-1} (I_{m^{j-1}} \otimes \widehat{\mathcal{N}}(\sigma_{k-j}, \hat{\mu})) \\
&\quad \times (I_{m^j} \otimes \widehat{\mathcal{K}}(\sigma_{k-j}, \hat{\mu})^{-1}) \Bigg) (I_{m^{k-1}} \otimes \widehat{\mathcal{B}}(\sigma_1, \hat{\mu})) \\
&=: -\widehat{W}_1^{\mathsf{H}} \Big(\partial_{\mu_i} \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})\Big) \widehat{V}_k.
\end{aligned}
$$

Noting that the model reduction matrix V is constructed as in Theorem 1, we obtain

$$
V \widehat{V}_k = \underbrace{V \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-1} W^{\mathsf{H}} \mathcal{K}(\sigma_k, \hat{\mu})}_{P_{\mathrm{V}_k}} V_k = V_k,
$$

where P_{V_k} is a projector onto $\operatorname{span}(V)$. Similarly, we have

$$
\begin{aligned}
W \widehat{W}_1 &= \underbrace{W \widehat{\mathcal{K}}(\sigma_k, \hat{\mu})^{-\mathsf{H}} V \mathcal{K}(\sigma_k, \hat{\mu})^{\mathsf{H}}}_{P_{\mathrm{W}_1}} \underbrace{\mathcal{K}(\sigma_k, \hat{\mu})^{-\mathsf{H}} \mathcal{C}(\sigma_k, \hat{\mu})^{\mathsf{H}}}_{W_1} \\
&= W_1,
\end{aligned}
$$

with P_{W_1} a projector onto $\operatorname{span}(W)$. Using those two identities, we obtain

$$
\begin{aligned}
\widehat{\mathcal{H}}_2 &= -\widehat{W}_1^{\mathsf{H}} W^{\mathsf{H}} \Big(\partial_{\mu_i} \mathcal{K}(\sigma_k, \hat{\mu})\Big) V \widehat{V}_k \\
&= -W_1^{\mathsf{H}} \Big(\partial_{\mu_i} \mathcal{K}(\sigma_k, \hat{\mu})\Big) V_k,
\end{aligned}
$$

i.e., $\widehat{\mathcal{H}}_2$ is identical to the term using the original matrix functions. Since the same technique can be used for all other α values corresponding the other columns in the set $\mathbb{A}$, we obtain, for all $1 \leq i \leq d$,

$$
\partial_{\mu_i} \widehat{G}_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}) = \partial_{\mu_i} G_k(\sigma_1, \ldots, \sigma_k, \hat{\mu}). \tag{21}
$$

Interpolation of the partial derivatives with respect to the frequency parameters follows by using the fixed parameter $\hat{\mu}$ in (Benner et al. 2021, Corollary 2). Together with (21), this proves (18).

Remark 5 *Theorem 4 shows how to match the parameter sensitivity implicitly without ever computing this quantity. Matching the parameter sensitivities is important, especially in the setting of optimization and design. These results can be extended to match the parameter Hessian as well; compare to Rodriguez et al. (2018). However, we skip those details for brevity.*

Remark 6 *All the results in Theorems 1 to 4 are formulated for a single parameter interpolation point $\hat{\mu} \in \mathbb{M}$. However, the results directly extend to interpolation at multiple parameter sampling points $\hat{\mu}^{(1)}, \ldots, \hat{\mu}^{(q)} \in \mathbb{M}$ by constructing the projection spaces for every parameter sample and then concatenating the resulting spaces into a single global projection space. As example, consider the task of interpolating*

$$
\begin{aligned}
&G_1(\sigma_1, \hat{\mu}^{(1)}), \quad G_2(\sigma_1, \sigma_2, \hat{\mu}^{(1)}), \\
&G_1(\sigma_3, \hat{\mu}^{(2)}), \quad G_2(\sigma_3, \sigma_4, \hat{\mu}^{(2)}),
\end{aligned} \tag{22}
$$

with the four frequency points $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ and the two parameter points $\hat{\mu}^{(1)}, \hat{\mu}^{(2)}$. Using Theorem 1 Part (a), we can construct basis matrices $V^{(1)}$, $V^{(2)}$ for the interpolation in either $\hat{\mu}^{(1)}$ or $\hat{\mu}^{(2)}$, respectively. The construction of a reduced-order model that satisfies all interpolation conditions (22) *is then given by constructing V such that*

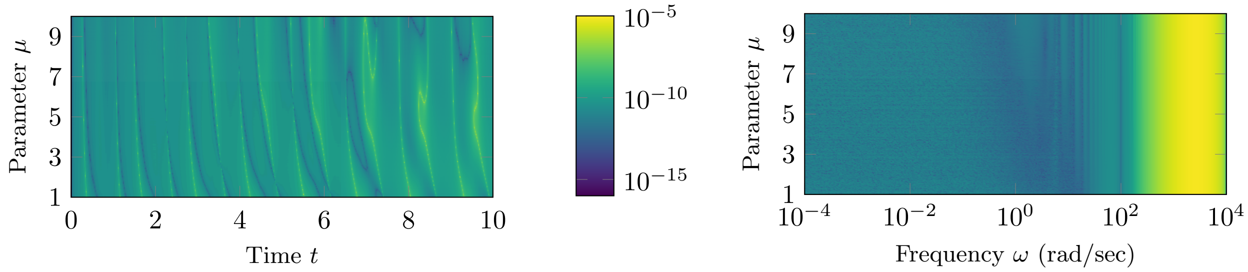
$$
\operatorname{span}(V) \supseteq \operatorname{span}([V^{(1)},\ V^{(2)}]).
$$

Remark 7 *The results simplify drastically for single-input/single-output (SISO) systems. In that case, the multivariate transfer functions corresponding to bilinear systems* (5) *can be written without Kronecker products* (23) *and the construction of the corresponding projection spaces simplifies such that no Kronecker products are involved anymore.*

5 Numerical examples

We illustrate the analysis with two benchmark examples. The experiments reported here have been executed on a machine with 2 Intel(R) Xeon(R) Silver 4110 CPU processors running at 2.10GHz and equipped with 192 GB total main memory. The computer is run on CentOS Linux release 7.5.1804 (Core) with MATLAB 9.7.0.1190202 (R2019b).

$$G_k(s_1, \dots, s_k, \mu) = \mathcal{C}(s_k, \mu)\mathcal{K}(s_k, \mu)^{-1} \left(\prod_{j=1}^{k-1} \mathcal{N}(s_{k-j}, \mu)\mathcal{K}(s_{k-j}, \mu)^{-1}) \right) \mathcal{B}(s_1, \mu) \tag{23}$$



(a) Relative time domain error.

(b) Relative error of the first transfer function.

Fig. 1. Relative errors for the time-delay system.

5.1 Parametric bilinear time-delay system

In the first example from Gosea et al. (2019), we consider a time-delayed heated rod modeled by a one-dimensional heat equation

$$\partial_t v(\zeta, t) = \partial_\zeta^2 v(\zeta, t) + a_1(\zeta)v(\zeta, t) + a_2(\zeta)v(\zeta, t-1) + u(t),$$

with homogeneous Dirichlet boundary conditions. We parameterize the diffusivity using the coefficients

$$a_1 = -\mu \sin(\zeta) \text{ and } a_2 = \mu \sin(\zeta), \quad \text{for} \quad \mu \in [1, 10].$$

The non-parametric example in Gosea et al. (2019) is recovered for $\mu = 2$. After a spatial discretization, we obtain a parametric bilinear system of the form

$$\begin{aligned} E\dot{x}(t) &= (A_0 - \mu A_{\mathrm{d}})x(t) + \mu A_{\mathrm{d}} x(t-1) \\ &\quad + Nx(t)u(t) + Bu(t), \\ y(t) &= Cx(t), \end{aligned}$$

with $m = p = 1$ and $n = 5\,000$. In our structured parametric setting, this model corresponds to the matrix functions

$$\begin{aligned} &\mathcal{K}(s, \mu) = sE - (A_0 - \mu A_{\mathrm{d}}) - \mu e^{-s} A_{\mathrm{d}}, \\ &\mathcal{B}(s, \mu) = B, \ \mathcal{N}(s, \mu) = N, \text{ and } \mathcal{C}(s, \mu) = C. \end{aligned}$$

The reduced-order model is constructed via Theorem 4 Part (a) with the frequency sampling points $\{\pm 10^{-4}\mathrm{i}, \pm 10^4\mathrm{i}\}$ and the parameter sampling points $\{1, 5.5, 10\}$ for the first two transfer functions. By construction, the reduced-order model has the same parametric time-delay structure as the original one, where the reduced matrices are given by

$$\begin{aligned} \widehat{E} &= W^{\mathsf{H}} E V, & \widehat{A}_0 &= W^{\mathsf{H}} A_0 V, & \widehat{A}_{\mathrm{d}} &= W^{\mathsf{H}} A_{\mathrm{d}} V, \\ \widehat{N} &= W^{\mathsf{H}} N V, & \widehat{B} &= W^{\mathsf{H}} B, & \widehat{C} &= CV, \end{aligned}$$

using the orthogonal truncation matrices V and W. The reduced-order system has the state-space dimension $r = 24$.

Figure 1a shows the relative time response error in the output, given by

$$\mathrm{err}_{1,\mathrm{t}}(t, \mu) := \frac{|y(t; \mu) - \hat{y}(t; \mu)|}{|y(t; \mu)|},$$

for $t \in [0, 10]$ and $\mu \in [1, 10]$, using the same test input signal as in Gosea et al. (2019), namely, $u(t) = 0.05\,(\cos(10t) + \cos(5t))$. The maximum error in the time and parameter domain is

$$\max_{\mu \in [1,10]} \left(\max_{t \in [0,10]} \mathrm{err}_{1,\mathrm{t}}(t, \mu) \right) \approx 9.993 \cdot 10^{-6},$$

illustrating a high-fidelity parametric reduced model over the full parameter domain. Figure 1b depicts the relative error in the first transfer function over the parameter range, computed as

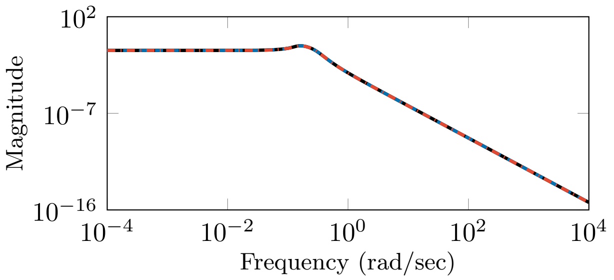
$$\mathrm{err}_{1,\mathrm{f}}(\omega_1, \mu) := \frac{|G_1(\omega_1 \mathrm{i}, \mu) - \widehat{G}_1(\omega_1 \mathrm{i}, \mu)|}{|G_1(\omega_1 \mathrm{i}, \mu)|},$$

where $\omega_1 \in [10^{-4}, 10^4]$ and $\mu \in [1, 10]$. As for the time domain error, we computed the maximum error to obtain

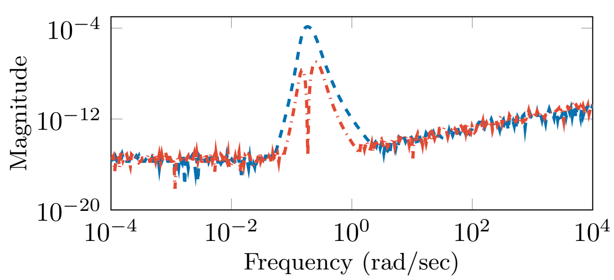
$$\max_{\mu \in [1,10]} \left(\max_{\omega_1 \in [10^{-4}, 10^4]} \mathrm{err}_{1,\mathrm{f}}(\omega_1, \mu) \right) \approx 7.002 \cdot 10^{-6},$$

showing the accuracy of the parametric reduced model in the frequency domain as well. We computed the maximum relative error in the second transfer function $G_2(s_1, s_2, \mu)$ as well to obtain

$$\max_{\mu \in [1,10]} \left(\max_{\omega_1, \omega_2 \in [10^{-4}, 10^{+4}]} \mathrm{err}_{1,\mathrm{f}}(\omega_1, \omega_2, \mu) \right) \approx 6.657 \cdot 10^{-4},$$



(a) Sigma plots.



(b) Relative errors.

Fig. 2. First transfer functions for the damped mass-spring system.

where

$$\mathrm{err}_{1,\mathrm{f}}(\omega_1, \omega_2, \mu) := \frac{|G_2(\omega_1 \mathrm{i}, \omega_2 \mathrm{i}, \mu) - \widehat{G}_2(\omega_1 \mathrm{i}, \omega_2 \mathrm{i}, \mu)|}{|G_2(\omega_1 \mathrm{i}, \omega_2 \mathrm{i}, \mu)|}.$$

All these results show that the structure-preserving parametric reduced-order model is an accurate approximation of the original system over the full parameter domain.

5.2 *Parametric bilinear mechanical system*

As second example, we consider a parametrized version of the multi-input/multi-output damped mass-spring system from Benner et al. (2021), a special case of the model (1), given by

$$\begin{aligned} M\ddot{x}(t;\mu) + D\dot{x}(t;\mu) + Kx(t;\mu) &= B_{\mathrm{u}}u(t) \\ &\quad + \mu_1 N_{\mathrm{p},1}x(t)u_1(t) + \mu_2 N_{\mathrm{p},2}x(t)u_2(t), \\ y(t;\mu) &= C_{\mathrm{p}}x(t;\mu)\dot{x}(t;\mu), \end{aligned}$$

where $\mu = (\mu_1, \mu_2)$ is the parameter entering through the bilinear terms and all the other matrices are exactly as in Benner et al. (2021), except for C_{p}, which we set as $C_{\mathrm{p}} = [e_2, e_{n-3}]^{\mathsf{T}}$, where e_j denotes the j-th column of the n-dimensional identity matrix. We have then $n = 1\,000$ masses, $m = 2$ inputs and $p = 2$ outputs. The parameter set is $\mathbb{M} = [0,1] \times [0,1]$. Note that for $\mu = (0,0)$, the system becomes linear as the bilinear terms are multiplied with 0. In our setting, this parametric bilinear model corresponds to

$$\begin{aligned} &\mathcal{K}(s,\mu) = s^2 M + sD + K, \ \mathcal{B}(s,\mu) = B_{\mathrm{u}}, \\ &\mathcal{N}(s,\mu) = \begin{bmatrix} \mu_1 N_{\mathrm{p},1}, \ \mu_2 N_{\mathrm{p},2} \end{bmatrix}, \text{ and } \mathcal{C}(s,\mu) = C_{\mathrm{p}}. \end{aligned}$$

Two reduced-order models are constructed via Theorem 1 to illustrate, on the one hand, the qualitative behavior of the structure-preserving reduced-order models, and, on the other hand, the effect of interpolation point selection. For the first reduced-order model, the interpolation points $\{\pm 10^{-4}\mathrm{i}, \pm 10^{4}\mathrm{i}\}$ are used in the frequency domain for the first two transfer function levels and combined with $\{(0,1), (1,0)\}$ in the parameter domain. To preserve the structural properties, such as positive definiteness of the mass, damping and stiffness matrices, we use a one-sided projection, i.e., we choose $W = V$. Since the first transfer function (the linear term) is independent of the parameter, some of the vectors in the construction of V are redundant and removed, yielding a structured parametric reduced-order model with $r = 40$. The reduced-order system matrices are then given by

$$\begin{aligned} &\widehat{M} = V^{\mathsf{H}}MV, && \widehat{D} = V^{\mathsf{H}}DV, && \widehat{K} = V^{\mathsf{H}}KV, \\ &\widehat{N}_{\mathrm{p},1} = V^{\mathsf{H}}N_{\mathrm{p},1}V, && \widehat{N}_{\mathrm{p},2} = V^{\mathsf{H}}N_{\mathrm{p},2}V, \\ &\widehat{B}_{\mathrm{u}} = V^{\mathsf{H}}B_{\mathrm{u}}, && \widehat{C}_{\mathrm{p}} = C_{\mathrm{p}}V, \end{aligned}$$

with the orthogonal truncation matrix V.

Next, we investigate the first transfer function of this reduced-order model in Figure 2. For most frequencies, the relative error, computed by

$$\mathrm{err}_{2,\mathrm{f}}(\omega_1) := \frac{\|G_1(\omega_1 \mathrm{i}) - \widehat{G}_1(\omega_1 \mathrm{i})\|_2}{\|G(\omega \mathrm{i})\|_2},$$

over the frequency range $\omega_1 \in [10^{-4}, 10^{+4}]$, is at machine precision except for a bump in the middle, where the transfer function behavior changes. To reduce the error in this region, we construct a second reduced-order model by adding an additional frequency interpolation point where the first reduced-order model attains its maximum error, around the frequency $\pm 1.85\mathrm{i}$. The second reduced-order model has the order $r = 60$.

Figure 2 illustrates the expected error behavior. The interpolation is numerically exact in the additional interpolation point and, additionally, has a significantly reduced error in the surrounding region. For a more detailed comparison, we have computed the maximum errors in frequency and time domain, which are provided

Table 1
Maximum errors for the damped mass-spring system.

	StrInt ($r = 40$)	StrInt ($r = 60$)
$\max_{\mu} \mathrm{err}_{2,\mathrm{t}}(\mu)$	2.7665e-3	8.3947e-6
$\max_{\omega_1} \mathrm{err}_{2,\mathrm{f}}(\omega_1)$	1.3407e-4	9.9974e-8
$\max_{\mu,\omega_1,\omega_2} \mathrm{err}_{2,\mathrm{f}}(\omega_1, \omega_1, \mu)$	1.4242e-3	9.1448e-6

in Table 1, with

$$\mathrm{err}_{2,\mathrm{t}}(\mu) := \max_{j\in\{1,2\}} \left(\max_{t\in[0,100]} \frac{|y_j(.;\mu) - \hat{y}_j(.;\mu)|}{|y_j(.;\mu)|} \right),$$

for the time simulations using the input signal $u(t) = \left[\sin(200t) + 200,\ -\cos(200t) - 200\right]^{\mathsf{T}}$, and

$$\mathrm{err}_{2,\mathrm{f}}(\omega_1, \omega_1, \mu) := \frac{\|G_2(\omega_1\mathrm{i}, \omega_2\mathrm{i}, \mu) - \widehat{G}_2(\omega_1\mathrm{i}, \omega_2\mathrm{i}, \mu)\|_2}{\|G_2(\omega_1\mathrm{i}, \omega_2\mathrm{i}, \mu)\|_2},$$

for the second transfer functions with $\omega_1, \omega_2 \in [10^{-4}, 10^{+4}]$ and $\mu \in [0,1]^2$. In both frequency and time domain, the errors of the larger reduced-order model with the additional interpolation point are significantly smaller than for the smaller reduced system. This suggests that a greedy procedure based on an error estimator, and selecting the next interpolation point based on minimizing the error with respect to the estimator as suggested for linear parametric time-invariant systems in Feng & Benner (2019) will be a promising future research direction.

6 Conclusions

We have presented a structure-preserving interpolation framework for model order reduction of parametric bilinear systems. We have established the subspace conditions to enforce interpolation both in the frequency and parameter domains. Two numerical examples illustrate that the approach is well suited for efficient structure-preserving model order reduction of parametric bilinear systems. The presented approach covers arbitrary parameter dependencies of the system as well as more system structures than shown in the examples.

An important open question is the appropriate choice of interpolation points in the frequency as well as the parameter domains to minimize the approximation error in some appropriate measure. In the parametric linear system case, this problem can be solved using error estimators in a greedy interpolation point selection. But for parametric bilinear systems, such error estimators are not yet developed. Preservation of stability in the structured parametric bilinear reduced-order model is another important avenue to investigate. While there are some special structures where stability can be preserved, this is a topic of future research for the general bilinear structure we considered in this paper.

Acknowledgements

Benner and Werner were supported by the German Research Foundation (DFG) Research Training Group 2297 "MathCoRe", Magdeburg. Gugercin was supported in parts by National Science Foundation under Grant No. DMS-1720257 and DMS-1819110. Part of this material is based upon work supported by the National Science Foundation under Grant No. DMS-1439786 and by the Simons Foundation Grant No. 507536 while Gugercin and Benner were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the "Model and dimension reduction in uncertain and dynamic systems" program.

References

Al-Baiyat, S., Farag, A. S. & Bettayeb, M. (1993), 'Transient approximation of a bilinear two-area interconnected power system', *Electric Power Systems Research* **26**(1), 11–19.

Antoulas, A. C., Beattie, C. A. & Gugercin, S. (2010), Interpolatory model reduction of large-scale dynamical systems, *in* J. Mohammadpour & K. M. Grigoriadis, eds, 'Efficient Modeling and Control of Large-Scale Systems', Springer US, pp. 3–58.

Antoulas, A. C., Beattie, C. A. & Gugercin, S. (2020), *Interpolatory Methods for Model Reduction*, Computational Science & Engineering, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Antoulas, A. C., Gosea, I. V. & Ionita, A. C. (2016), 'Model reduction of bilinear systems in the Loewner framework', *SIAM J. Sci. Comput.* **38**(5), B889–B916.

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

Baur, U., Beattie, C. A., Benner, P. & Gugercin, S. (2011), 'Interpolatory projection methods for parameterized model reduction', *SIAM J. Sci. Comput.* **33**(5), 2489–2518.

Baur, U., Benner, P. & Feng, L. (2014), 'Model order reduction for linear and nonlinear systems: A system-theoretic perspective', *Arch. Comput. Methods Eng.* **21**(4), 331–358.

Beattie, C. A. & Gugercin, S. (2009), 'Interpolatory projection methods for structure-preserving model reduction', *Syst. Control Lett.* **58**(3), 225–232.

Benner, P. & Breiten, T. (2011), 'On H_2-model reduction of linear parameter-varying systems', *Proc. Appl. Math. Mech.* **11**(1), 805–806.

Benner, P. & Breiten, T. (2012), ‘Interpolation-based $\mathcal{H}_2$-model reduction of bilinear control systems’, *SIAM J. Matrix Anal. Appl.* **33**(3), 859–885.

Benner, P., Cohen, A., Ohlberger, M. & Willcox, K., eds (2017), *Model Reduction and Approximation: Theory and Algorithms*, Computational Science & Engineering, SIAM Publications, Philadelphia, PA.

Benner, P. & Damm, T. (2011), ‘Lyapunov equations, energy functionals, and model order reduction of bilinear and stochastic systems’, *SIAM J. Control Optim.* **49**(2), 686–711.

Benner, P., Gugercin, S. & Werner, S. W. R. (2021), ‘Structure-preserving interpolation of bilinear control systems’, *Adv. Comput. Math.* **47**(3), 43.

Breiten, T. & Damm, T. (2010), ‘Krylov subspace methods for model order reduction of bilinear control systems’, *Syst. Control Lett.* **59**(8), 443–450.

Carleman, T. (1932), ‘Application de la théorie des équations intégrales linéaires aux systèmes d’équations différentielles non linéaires’, *Acta Math.* **59**, 63–87.

Condon, M. & Ivanov, R. (2007), ‘Krylov subspaces from bilinear representations of nonlinear systems’, *Compel-Int. J. Comp. Math. Electr. Electron. Eng.* **26**(2), 399–406.

Feng, L. & Benner, P. (2019), ‘A new error estimator for reduced-order modeling of linear parametric systems’, *IEEE Trans. Microw. Theory Techn.* **67**(12), 4848–4859.

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

Gosea, I. V., Pontes Duff, I., Benner, P. & Antoulas, A. C. (2019), Model order reduction of bilinear time-delay systems, *in* ‘18th European Control Conference (ECC)’, pp. 2289–2294.

Hsu, C. S., Desai, U. B. & Crawley, C. A. (1983), Realization algorithms and approximation methods of bilinear systems, *in* ‘The 22nd IEEE Conference on Decision and Control, San Antonio, USA’, pp. 783–788.

Hu, H. Y. & Wang, Z. H. (2002), *Dynamics of Controlled Mechanical Systems with Delayed Feedback*, Springer-Verlag, Berlin, Heidelberg.

Mohler, R. R. (1973), *Bilinear Control Processes: With Applications to Engineering, Ecology and Medicine*, Vol. 106 of *Mathematics in Science and Engineering*, Academic Press, New York, London.

Ou, Y. (2010), Optimal Control of a Class of Nonlinear Parabolic PDE Systems Arising in Fusion Plasma Current Profile Dynamics, PhD thesis, Lehigh University, Bethlehem, Pennsylvania, USA.

Quarteroni, A. & Rozza, G. (2014), *Reduced Order Methods for Modeling and Computational Reduction*, Vol. 9 of *MS&A – Modeling, Simulation and Applications*, Springer International Publishing, Cham.

Rodriguez, A. C., Gugercin, S. & Boggaard, J. (2018), ‘Interpolatory model reduction of parameterized bilinear dynamical systems’, *Adv. Comput. Math.* **44**(6), 1887–1916.

Rugh, W. J. (1981), *Nonlinear System Theory: The Volterra/Wiener Approach*, The Johns Hopkins University Press, Baltimore.

Saputra, J., Saragih, R. & Handayani, D. (2019), ‘Robust H_∞ controller for bilinear system to minimize HIV concentration in blood plasma’, *J. Phys.: Conf. Ser.* **1245**, 012055.

Scarciotti, G. & Astolfi, A. (2017), ‘Nonlinear model reduction by moment matching’, *Foundations and Trends® in Systems and Control* **4**(3–4), 224–409.

Schilders, W. H. A., van der Vorst, H. A. & Rommes, J. (2008), *Model Order Reduction: Theory, Research Aspects and Applications*, Vol. 13 of *Mathematics in Industry (MATHINDUSTRY)*, Springer-Verlag, Berlin, Heidelberg.

Zhang, L. & Lam, J. (2002), ‘On H_2 model reduction of bilinear systems’, *Automatica J. IFAC* **38**(2), 205–216.