Math. Commun. **31** (2026), 1–14
DOI: 10.1000/100

Mathematical Communications

Interpolatory necessary optimality conditions for reduced-order modeling of parametric linear time-invariant systems

Petar Mlinarić[1,*], Peter Benner[2], and Serkan Gugercin[3]

[1] Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička cesta 30, HR-10000 Zagreb, Croatia
[2] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, D-39106 Magdeburg, Germany
[3] Department of Mathematics and Division of Computational Modeling and Data Analytics, Academy of Data Science, Virginia Tech, 225 Stanger Street, Blacksburg, VA 24061, USA

Abstract. Interpolatory necessary optimality conditions for $\mathcal{H}_2$-optimal reduced-order modeling of non-parametric linear time-invariant (LTI) systems are known and well-investigated. In this paper, using the general framework of $\mathcal{L}_2$-optimal reduced-order modeling of parametric stationary problems, we derive interpolatory $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality conditions for parametric LTI systems with a general pole-residue form. Then we specialize this result to recover known conditions for systems with parameter-independent poles and develop new conditions for a certain class of systems with parameter-dependent poles.

AMS subject classifications: 30E10, 41A20, 46N10, 65D05, 65D15, 93A15, 93B40
Keywords: reduced-order modeling; parametric systems; optimization; linear systems
Received June 3, 2025; accepted October 30, 2025

1. Introduction

Consider a parametric linear time-invariant (LTI) system (full-order model (FOM))

$$\mathcal{E}(\mathsf{q})\dot{x}(t,\mathsf{q}) = \mathcal{A}(\mathsf{q})x(t,\mathsf{q}) + \mathcal{B}(\mathsf{q})u(t), \tag{1a}$$

$$y(t,\mathsf{q}) = \mathcal{C}(\mathsf{q})x(t,\mathsf{q}), \tag{1b}$$

where $\mathsf{q} \in \mathcal{Q} \subseteq \mathbb{C}^{n_\mathsf{q}}$ is the parameter vector; $u(t) \in \mathbb{C}^{n_\mathrm{i}}$ is the input; $x(t,\mathsf{q}) \in \mathbb{C}^n$ is the state; $y(t,\mathsf{q}) \in \mathbb{C}^{n_\mathrm{o}}$ is the output; and $\mathcal{E}(\mathsf{q}), \mathcal{A}(\mathsf{q}) \in \mathbb{C}^{n \times n}$, $\mathcal{B}(\mathsf{q}) \in \mathbb{C}^{n \times n_\mathrm{i}}$, and $\mathcal{C}(\mathsf{q}) \in \mathbb{C}^{n_\mathrm{o} \times n}$ are parametric matrices. Given the FOM in (1), the goal of parametric reduced-order modeling is to find a reduced parametric LTI system (reduced-order model (ROM))

$$\widehat{\mathcal{E}}(\mathsf{q})\dot{\widehat{x}}(t,\mathsf{q}) = \widehat{\mathcal{A}}(\mathsf{q})\widehat{x}(t,\mathsf{q}) + \widehat{\mathcal{B}}(\mathsf{q})u(t), \tag{2a}$$

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

where $\widehat{x}(t,\mathsf{q}) \in \mathbb{C}^r$ is the reduced state with $r \ll n$; $\widehat{y}(t,\mathsf{q}) \in \mathbb{C}^{n_\mathrm{o}}$ is the approximate output; and $\widehat{\mathcal{E}}(\mathsf{q}), \widehat{\mathcal{A}}(\mathsf{q}) \in \mathbb{C}^{r \times r}$, $\widehat{\mathcal{B}}(\mathsf{q}) \in \mathbb{C}^{r \times n_\mathrm{i}}$, and $\widehat{\mathcal{C}}(\mathsf{q}) \in \mathbb{C}^{n_\mathrm{o} \times r}$ are the reduced parametric matrices, such that $\widehat{y}$ approximates y for a wide range of inputs u and a set of parameters $\mathsf{q} \in \mathcal{Q}$. Parametric dynamical systems are ubiquitous in applications, ranging from inverse problems and uncertainty quantification to optimization, and model reduction of parametric systems has been a major research topic; we refer the reader to, e.g., [4, 3] for more details.

Both the FOM and the ROM can be fully described by their (parametric) transfer functions, given by

$$H(s,\mathsf{q}) = \mathcal{C}(\mathsf{q})(s\mathcal{E}(\mathsf{q}) - \mathcal{A}(\mathsf{q}))^{-1}\mathcal{B}(\mathsf{q}) \quad \text{and}$$
$$\widehat{H}(s,\mathsf{q}) = \widehat{\mathcal{C}}(\mathsf{q})\left(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q})\right)^{-1}\widehat{\mathcal{B}}(\mathsf{q}), \tag{3}$$

Email addresses: petar.mlinaric@math.hr (P. Mlinarić), benner@mpi-magdeburg.mpg.de (P. Benner), gugercin@vt.edu (S. Gugercin)
*Corresponding author.

https://www.mathos.unios.hr/mc

respectively. As in any approximation problem, one needs a metric to judge the quality of the approximation. For *non-parametric* LTI systems, i.e., when $\mathcal{E}, \mathcal{A}, \mathcal{B}, \mathcal{C}$ are constant matrices, the $\mathcal{H}_2$-norm has been one of the most commonly used metrics in (optimal) reduced-order modeling [1, 8, 12, 23]. For parametric LTI systems considered here, the $\mathcal{H}_2 \otimes \mathcal{L}_2$-norm introduced in [2] provides a natural extension. The goal of $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling is to find ROM that (locally) minimizes the $\mathcal{H}_2 \otimes \mathcal{L}_2$ error

$$\left\| H - \widehat{H} \right\|_{\mathcal{H}_2 \otimes \mathcal{L}_2} = \left(\int_{\mathcal{Q}} \left\| H(\cdot, \mathsf{q}) - \widehat{H}(\cdot, \mathsf{q}) \right\|_{\mathcal{H}_2}^2 \mathrm{d}\nu(\mathsf{q}) \right)^{1/2}, \tag{4}$$

where ν is a measure over $\mathcal{Q}$ and the $\mathcal{H}_2$ norm is given as

$$\|H(\cdot, \mathsf{q})\|_{\mathcal{H}_2} = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|H(\imath\omega, \mathsf{q})\|_{\mathrm{F}}^2 \, \mathrm{d}\omega \right)^{1/2},$$

assuming $H(\cdot, \mathsf{q})$ is analytic in the open right half-plane (otherwise, we say that the $\mathcal{H}_2$ norm is infinite). The $\mathcal{H}_2 \otimes \mathcal{L}_2$ error gives an upper bound for the output error

$$\|y - \widehat{y}\|_{\mathcal{L}_\infty \otimes \mathcal{L}_2} \leqslant \left\| H - \widehat{H} \right\|_{\mathcal{H}_2 \otimes \mathcal{L}_2} \|u\|_{\mathcal{L}_2},$$

where

$$\|y\|_{\mathcal{L}_\infty \otimes \mathcal{L}_2} = \left(\int_{\mathcal{Q}} \|y(\cdot, \mathsf{q})\|_{\mathcal{L}_\infty}^2 \, \mathrm{d}\nu(\mathsf{q}) \right)^{1/2}$$

is the $\mathcal{L}_\infty \otimes \mathcal{L}_2$ norm of the output, further justifying the use of the $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm in parametric reduced-order modeling.

The $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm with ν as the Lebesgue measure was introduced in Baur et al. [2]. There, for the special case where $\widehat{\mathcal{E}}$ and $\widehat{\mathcal{A}}$ are *parameter-independent*, the $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling problem was converted to a non-parametric $\mathcal{H}_2$-optimal reduced-order modeling problem and interpolatory optimality conditions could be established. For another simplified problem where the poles of $\widehat{H}$ do *not* vary with the parameter $\mathsf{q} \in \{|\mathsf{q}| = 1\} \subset \mathbb{C}$, Grimm [7] used an $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm, derived interpolatory conditions and proposed an optimization algorithm.

A common assumption in parametric reduced-order modeling methods is parameter-separability. For the ROM (2), this would mean that the reduced quantities can be written as

$$\widehat{\mathcal{E}}(\mathsf{q}) = \sum_{\ell=1}^{n_{\widehat{\mathcal{E}}}} \widehat{\varepsilon}_\ell(\mathsf{q}) \widehat{E}_\ell, \quad \widehat{\mathcal{A}}(\mathsf{q}) = \sum_{i=1}^{n_{\widehat{\mathcal{A}}}} \widehat{\alpha}_i(\mathsf{q}) \widehat{A}_i, \quad \widehat{\mathcal{B}}(\mathsf{q}) = \sum_{j=1}^{n_{\widehat{\mathcal{B}}}} \widehat{\beta}_j(\mathsf{q}) \widehat{B}_j, \quad \widehat{\mathcal{C}}(\mathsf{q}) = \sum_{k=1}^{n_{\widehat{\mathcal{C}}}} \widehat{\gamma}_k(\mathsf{q}) \widehat{C}_k, \tag{5}$$

for some functions $\widehat{\varepsilon}_\ell, \widehat{\alpha}_i, \widehat{\beta}_j, \widehat{\gamma}_k \colon \mathcal{Q} \to \mathbb{R}$, constant matrices $\widehat{E}_\ell$, $\widehat{A}_i$, $\widehat{B}_j$, $\widehat{C}_k$, and positive integers $n_{\widehat{\mathcal{E}}}, n_{\widehat{\mathcal{A}}}, n_{\widehat{\mathcal{B}}}, n_{\widehat{\mathcal{C}}}$. We call ROM of this form structured ROM (StROM). This form has also been considered in $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling methods. In particular, Petersson [19] considered the case of a discretized $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm, i.e., where ν is a sum of Dirac measures, proposing an optimization algorithm to find a locally $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal ROM. Additionally, Hund et al. [9] proposed an optimization algorithm for $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling using quadrature for the case of the Lebesgue measure. Both of these works used matrix equation-based, Wilson-type conditions [23] and not interpolation.

The $\mathcal{H}_2 \otimes \mathcal{L}_2$-norm was also used by Brunsch [6] to derive error bounds within a reduced-order modeling framework for parametric LTI systems with symmetric positive definite $\mathcal{E}(\mathsf{q})$ and $-\mathcal{A}(\mathsf{q})$. The method is based on sparse-grid interpolation in the parameter domain. It satisfies (Hermite) interpolation conditions and preserves stability, but has no proven optimality properties. See also [10, 5, 20] for some data-driven approaches.

In our recent work on $\mathcal{L}_2$-optimal reduced-order modeling [15], we covered both LTI systems and parametric stationary problems. We developed interpolatory necessary optimality conditions in [16] for certain types of structured ROMs (StROMs), including *non-parametric* LTI systems and parametric stationary problems. We also showed that the interpolatory conditions of [7] can be derived from our generalized $\mathcal{L}_2$-optimality conditions. However, as stated before, [7] assumes the poles are fixed.

Therefore, unlike for non-parametric LTI problems for which interpolatory optimality conditions for $\mathcal{H}_2$ model reduction have been well-established [12, 8, 1], there is a significant gap in the development of interpolatory optimality conditions for $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal parametric ROM construction, except for the special cases mentioned above. Our goal in this paper is to close this gap and to develop interpolatory optimality conditions for a more general setting of parametric LTI systems. Additionally, we show that our analysis contains the earlier conditions from [2] as a special case.

We provide background in Section 2. While Section 3 covers the general parametric diagonal StROMs (D-StROMs) case, Sections 4 and 5 focus on simplified cases, leading to optimality conditions that can be directly linked to the bitangential Hermite interpolation framework. We conclude with Section 7.

2. Background

Here we recall one of the main results of [14], specifically, the necessary $\mathcal{L}_2$-optimality conditions for D-StROMs (based on prior work in [15, 16]), which will form the foundation of our analysis. For generality, we switch the notation from $\mathcal{Q}$ and q to $\mathcal{P}$ and p, and in the later sections we will have $\mathsf{p} = (s, \mathsf{q})$.

Given a parameter-to-output mapping

$$H \colon \mathcal{P} \to \mathbb{C}^{n_{\mathrm{o}} \times n_{\mathrm{i}}},$$

the goal in [14] is to construct StROM

$$\widehat{\mathcal{K}}(\mathsf{p})\widehat{X}(\mathsf{p}) = \widehat{\mathcal{F}}(\mathsf{p}), \tag{6a}$$

$$\widehat{H}(\mathsf{p}) = \widehat{\mathcal{G}}(\mathsf{p})\widehat{X}(\mathsf{p}), \tag{6b}$$

with a parameter-separable form

$$\widehat{\mathcal{K}}(\mathsf{p}) = \sum_{i=1}^{n_{\widehat{\mathcal{K}}}} \widehat{\kappa}_i(\mathsf{p})\widehat{K}_i, \quad \widehat{\mathcal{F}}(\mathsf{p}) = \sum_{j=1}^{n_{\widehat{\mathcal{F}}}} \widehat{\zeta}_j(\mathsf{p})\widehat{F}_j, \quad \widehat{\mathcal{G}}(\mathsf{p}) = \sum_{k=1}^{n_{\widehat{\mathcal{G}}}} \widehat{\eta}_k(\mathsf{p})\widehat{G}_k, \tag{7}$$

where $\widehat{X}(\mathsf{p}) \in \mathbb{C}^{r \times n_{\mathrm{i}}}$ is the reduced state, $\widehat{H}(\mathsf{p}) \in \mathbb{C}^{n_{\mathrm{o}} \times n_{\mathrm{i}}}$ is the approximate output, $\widehat{\mathcal{K}}(\mathsf{p}) \in \mathbb{C}^{r \times r}$, $\widehat{\mathcal{F}}(\mathsf{p}) \in \mathbb{C}^{r \times n_{\mathrm{i}}}$, $\widehat{\mathcal{G}}(\mathsf{p}) \in \mathbb{C}^{n_{\mathrm{o}} \times r}$, $\widehat{\kappa}_i, \widehat{\zeta}_j, \widehat{\eta}_k \colon \mathcal{P} \to \mathbb{C}$, $\widehat{K}_i \in \mathbb{C}^{r \times r}$, $\widehat{F}_j \in \mathbb{C}^{r \times n_{\mathrm{i}}}$, and $\widehat{G}_k \in \mathbb{C}^{n_{\mathrm{o}} \times r}$. The goal is to construct $\widehat{\mathcal{K}}(\mathsf{p})$, $\widehat{\mathcal{F}}(\mathsf{p})$, and $\widehat{\mathcal{G}}(\mathsf{p})$ such that $\widehat{H}(\mathsf{p}) = \widehat{\mathcal{G}}(\mathsf{p})\widehat{\mathcal{K}}(\mathsf{p})^{-1}\widehat{\mathcal{F}}(\mathsf{p})$ is an optimal $\mathcal{L}_2$-approximation to the original mapping $H(\mathsf{p})$, i.e.,

$$\left\| H - \widehat{H} \right\|_{\mathcal{L}_2} = \left(\int_{\mathcal{P}} \left\| H(\mathsf{p}) - \widehat{H}(\mathsf{p}) \right\|_{\mathrm{F}}^2 \mathrm{d}\mu(\mathsf{p}) \right)^{1/2} \tag{8}$$

is minimized, where μ is a measure over $\mathcal{P}$. The $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm (4) is a special case of the $\mathcal{L}_2$-norm (8) for appropriately defined p and μ, a fact we exploit in Sections 3 to 5. We will use the notation $(\widehat{K}_i, \widehat{F}_j, \widehat{G}_k)$ to denote the StROM specified by (6) and (7).

We assume $(\widehat{K}_i, \widehat{F}_j, \widehat{G}_k)$ is D-StROM, i.e., all $\widehat{K}_i$'s are diagonal, and in return so is $\widehat{\mathcal{K}}(\mathsf{p})$ in (6). Then $\widehat{H}$ has a "pole-residue" form

$$\widehat{H}(\mathsf{p}) = \widehat{\mathcal{G}}(\mathsf{p})\widehat{\mathcal{K}}(\mathsf{p})^{-1}\widehat{\mathcal{F}}(\mathsf{p}) = \sum_{\ell=1}^{r} \frac{g_\ell(\mathsf{p}) f_\ell(\mathsf{p})^*}{k_\ell(\mathsf{p})}, \tag{9}$$

where $k_\ell(\mathsf{p})$ is the ℓth diagonal entry of $\widehat{\mathcal{K}}(\mathsf{p})$, $f_\ell(\mathsf{p}) = \widehat{\mathcal{F}}(\mathsf{p})^* e_\ell$, and $g_\ell(\mathsf{p}) = \widehat{\mathcal{G}}(\mathsf{p}) e_\ell$, with e_ℓ denoting the ℓth canonical basis vector of appropriate size. With this pole-residue form in hand, we have the optimality conditions for D-StROMs (Corollary 2.4 in [14]).

Theorem 1. *Suppose that $\mathcal{P} \subseteq \mathbb{C}^{n_{\mathsf{p}}}$; μ is a measure over $\mathcal{P}$; the function H is in $\mathcal{L}_2(\mathcal{P}, \mu; \mathbb{C}^{n_{\mathrm{o}} \times n_{\mathrm{i}}})$; functions $\widehat{\kappa}_i, \widehat{\zeta}_j, \widehat{\eta}_k \colon \mathcal{P} \to \mathbb{C}$ are measurable and satisfy*

$$\int_{\mathcal{P}} \left(\frac{\sum_{j=1}^{n_{\widehat{\mathcal{F}}}} \left| \widehat{\zeta}_j(\mathsf{p}) \right| \sum_{k=1}^{n_{\widehat{\mathcal{G}}}} \left| \widehat{\eta}_k(\mathsf{p}) \right|}{\sum_{i=1}^{n_{\widehat{\mathcal{K}}}} \left| \widehat{\kappa}_i(\mathsf{p}) \right|} \right)^2 \mathrm{d}\mu(\mathsf{p}) < \infty, \tag{10}$$

$\widehat{K}_i \in \mathbb{C}^{r \times r}$, $\widehat{F}_j \in \mathbb{C}^{r \times n_{\mathrm{i}}}$, $\widehat{G}_k \in \mathbb{C}^{n_{\mathrm{o}} \times r}$; and

$$\operatorname*{ess\,sup}_{\mathsf{p} \in \mathcal{P}} \left\| \widehat{\kappa}_i(\mathsf{p}) \widehat{\mathcal{K}}(\mathsf{p})^{-1} \right\|_{\mathrm{F}} < \infty, \quad i = 1, 2, \dots, n_{\widehat{\mathcal{K}}}, \tag{11}$$

where $\widehat{\mathcal{K}}$ is as in (7). Furthermore, let $(\widehat{K}_i, \widehat{F}_j, \widehat{G}_k)$ be an $\mathcal{L}_2$-optimal D-StROM of H with $\widehat{H}$ as in (9). Then

$$\int_{\mathcal{P}} \frac{\overline{\widehat{\eta}_k(\mathsf{p})} H(\mathsf{p}) f_\ell(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}} \,\mathrm{d}\mu(\mathsf{p}) = \int_{\mathcal{P}} \frac{\overline{\widehat{\eta}_k(\mathsf{p})} \widehat{H}(\mathsf{p}) f_\ell(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}} \,\mathrm{d}\mu(\mathsf{p}), \tag{12a}$$

$$\int_{\mathcal{P}} \frac{\overline{\widehat{\zeta}_j(\mathsf{p})} g_\ell(\mathsf{p})^* H(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}} \,\mathrm{d}\mu(\mathsf{p}) = \int_{\mathcal{P}} \frac{\overline{\widehat{\zeta}_j(\mathsf{p})} g_\ell(\mathsf{p})^* \widehat{H}(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}} \,\mathrm{d}\mu(\mathsf{p}), \tag{12b}$$

$$\int_{\mathcal{P}} \frac{\overline{\widehat{\kappa}_i(\mathsf{p})} g_\ell(\mathsf{p})^* H(\mathsf{p}) f_\ell(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}^2} \,\mathrm{d}\mu(\mathsf{p}) = \int_{\mathcal{P}} \frac{\overline{\widehat{\kappa}_i(\mathsf{p})} g_\ell(\mathsf{p})^* \widehat{H}(\mathsf{p}) f_\ell(\mathsf{p})}{\overline{k_\ell(\mathsf{p})}^2} \,\mathrm{d}\mu(\mathsf{p}), \tag{12c}$$

for $i = 1, 2, \dots, n_{\widehat{\mathcal{K}}}$, $j = 1, 2, \dots, n_{\widehat{\mathcal{F}}}$, $k = 1, 2, \dots, n_{\widehat{\mathcal{G}}}$, and $\ell = 1, 2, \dots, r$.

Theorem 1 establishes the interpolatory optimality conditions (12) for $\mathcal{L}_2$-optimal approximation. We showed in [14] that various structured reduced-order modeling problems appear as a special case of Theorem 1 and derived interpolatory optimality conditions for important classes of *non-parametric* structured LTI systems. In this paper, we extend this analysis to parametric LTI systems.

3. $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal parametric interpolation

Here we use Theorem 1 to derive interpolatory conditions for $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order approximation of the FOM (1) using D-StROMs. But first we need to establish what assumptions (10) and (11) appearing in $\mathcal{L}_2$-optimal approximation the structure (7) correspond to in the case of $\mathcal{H}_2 \otimes \mathcal{L}_2$ approximation with the structures of StROMs in (5).

Lemma 1. *Let $\mathsf{p} = (s, \mathsf{q})$, $\mathcal{P} = \imath\mathbb{R} \times \mathcal{Q}$, and $\mu = \frac{1}{2\pi} \lambda_{\imath\mathbb{R}} \times \nu$ where $\lambda_{\imath\mathbb{R}}$ is the Lebesgue measure over $\imath\mathbb{R}$ and ν is a measure over $\mathcal{Q} \subseteq \mathbb{C}^{n_{\mathsf{q}}}$. Furthermore, for StROM in (5), let the functions $\widehat{\varepsilon}_\ell, \widehat{\alpha}_i, \widehat{\beta}_j, \widehat{\gamma}_k \colon \mathcal{Q} \to \mathbb{C}$ be measurable. Then the condition*

$$\int_{\mathcal{Q}} \frac{\left(\sum_{j=1}^{n_{\widehat{\mathcal{B}}}} \left| \widehat{\beta}_j(\mathsf{q}) \right| \sum_{k=1}^{n_{\widehat{\mathcal{C}}}} \left| \widehat{\gamma}_k(\mathsf{q}) \right| \right)^2}{\sum_{\ell=1}^{n_{\widehat{\mathcal{E}}}} \left| \widehat{\varepsilon}_\ell(\mathsf{q}) \right| \sum_{i=1}^{n_{\widehat{\mathcal{A}}}} \left| \widehat{\alpha}_i(\mathsf{q}) \right|} \,\mathrm{d}\nu(\mathsf{q}) < \infty \tag{13}$$

is equivalent to (10), and the conditions

$$\operatorname*{ess\,sup}_{\mathsf{q} \in \mathcal{Q}} \left| \widehat{\varepsilon}_\ell(\mathsf{q}) \right| \left\| s \left(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q}) \right)^{-1} \right\|_{\mathcal{L}_\infty} < \infty, \tag{14a}$$

$$\operatorname*{ess\,sup}_{\mathsf{q} \in \mathcal{Q}} \left| \widehat{\alpha}_i(\mathsf{q}) \right| \left\| \left(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q}) \right)^{-1} \right\|_{\mathcal{L}_\infty} < \infty, \tag{14b}$$

for $\ell = 1, \dots, n_{\widehat{\mathcal{E}}}$ and $i = 1, \dots, n_{\widehat{\mathcal{A}}}$, are equivalent to (11).

Proof. First note that with the choices of $\mathsf{p} = (s, \mathsf{q})$, $\mathcal{P} = \imath\mathbb{R} \times \mathcal{Q}$, and $\mu = \frac{1}{2\pi} \lambda_{\imath\mathbb{R}} \times \nu$, the $\mathcal{L}_2$-norm in (8) recovers the $\mathcal{H}_2 \otimes \mathcal{L}_2$norm in (4). Now note that the integral in (10), for the StROM $(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q}), \widehat{\mathcal{B}}(\mathsf{q}), \widehat{\mathcal{C}}(\mathsf{q}))$ as in (5), takes the form

$$\int_{\mathcal{Q}} \int_{-\infty}^{\infty} \left(\frac{\sum_{j=1}^{n_{\widehat{\mathcal{B}}}} \left| \widehat{\beta}_j(\mathsf{q}) \right| \sum_{k=1}^{n_{\widehat{\mathcal{C}}}} \left| \widehat{\gamma}_k(\mathsf{q}) \right|}{|\omega| \sum_{\ell=1}^{n_{\widehat{\mathcal{E}}}} \left| \widehat{\varepsilon}_\ell(\mathsf{q}) \right| + \sum_{i=1}^{n_{\widehat{\mathcal{A}}}} \left| \widehat{\alpha}_i(\mathsf{q}) \right|} \right)^2 \mathrm{d}\omega \,\mathrm{d}\nu(\mathsf{q}).$$

Using that $\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{(a|x|+b)^2} = \frac{2}{ab}$ for positive a and b, the above integral becomes equal to the one in (13), up to scaling by 2.

Next, the conditions in (11) become

$$\operatorname*{ess\,sup}_{\mathsf{q}\in\mathcal{Q}} \operatorname*{ess\,sup}_{\omega\in\mathbb{R}} \left\| \imath\omega\widehat{\varepsilon}_\ell(\mathsf{q}) \left(\imath\omega\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q})\right)^{-1} \right\|_{\mathrm{F}} < \infty,$$
$$\operatorname*{ess\,sup}_{\mathsf{q}\in\mathcal{Q}} \operatorname*{ess\,sup}_{\omega\in\mathbb{R}} \left\| \widehat{\alpha}_i(\mathsf{q}) \left(\imath\omega\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q})\right)^{-1} \right\|_{\mathrm{F}} < \infty,$$

which simplify to

$$\operatorname*{ess\,sup}_{\mathsf{q}\in\mathcal{Q}} |\widehat{\varepsilon}_\ell(\mathsf{q})| \operatorname*{ess\,sup}_{s\in\imath\mathbb{R}} \left\| s\left(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q})\right)^{-1} \right\|_{\mathrm{F}} < \infty,$$
$$\operatorname*{ess\,sup}_{\mathsf{q}\in\mathcal{Q}} |\widehat{\alpha}_i(\mathsf{q})| \operatorname*{ess\,sup}_{s\in\imath\mathbb{R}} \left\| \left(s\widehat{\mathcal{E}}(\mathsf{q}) - \widehat{\mathcal{A}}(\mathsf{q})\right)^{-1} \right\|_{\mathrm{F}} < \infty.$$

Since $\|\cdot\|_{\mathrm{F}}$ and $\|\cdot\|_2$ are equivalent norms, the above conditions are equivalent to (14). □

In [9] it was assumed that $\mathcal{Q} \subset \mathbb{R}^{n_{\mathsf{q}}}$ is compact, ν is a finite Borel measure over $\mathcal{Q}$, $\widehat{\varepsilon}_\ell, \widehat{\alpha}_i, \widehat{\beta}_j, \widehat{\gamma}_k \colon \mathcal{Q} \to \mathbb{R}$ are continuous, $\widehat{\mathcal{E}}(\mathsf{q})$ is invertible, and $\widehat{\mathcal{E}}(\mathsf{q})^{-1}\widehat{\mathcal{A}}(\mathsf{q})$ has all eigenvalues in the open left half-plane for all $\mathsf{q} \in \mathcal{Q}$. Therefore, we see that the assumptions of the earlier work [9] on $\mathcal{H}_2 \otimes \mathcal{L}_2$ approximation are indeed a special case of the ones we derived in Lemma 1.

Now that we have established the assumptions of Theorem 1 for parametric LTI systems, we are ready to derive the corresponding interpolatory $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality conditions based on conditions (12). We take that $\widehat{\mathcal{E}}(\mathsf{q}) = I$ and all $\widehat{A}_i$ are diagonal. (The result can be extended to parametric diagonal $\widehat{\mathcal{E}}$; only the expressions become more involved.)

Theorem 2. *Given the full-order parametric transfer function H of the finite $\mathcal{H}_2 \otimes \mathcal{L}_2$ norm, let $\widehat{H}$ in (3) be an $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal D-StROM for H with $\widehat{\mathcal{E}}(\mathsf{q}) = I$ and all $\widehat{A}_i$ diagonal in (5). Let $\lambda_\ell(\mathsf{q})$ denote the ℓth diagonal entry of $\widehat{\mathcal{A}}(\mathsf{q})$. Moreover, define $c_\ell(\mathsf{q}) = \widehat{\mathcal{C}}(\mathsf{q})e_\ell$ and $b_\ell(\mathsf{q}) = \widehat{\mathcal{B}}(\mathsf{q})^* e_\ell$, where $\widehat{\mathcal{B}}(\mathsf{q})$ and $\widehat{\mathcal{C}}(\mathsf{q})$ are as defined in (5). Then*

$$\widehat{H}(s, \mathsf{q}) = \sum_{\ell=1}^{r} \frac{c_\ell(\mathsf{q}) b_\ell(\mathsf{q})^*}{s - \lambda_\ell(\mathsf{q})} \tag{15}$$

and

$$\int_{\mathcal{Q}} \widehat{\gamma}_k(\mathsf{q}) H\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) b_\ell(\mathsf{q}) \,\mathrm{d}\nu(\mathsf{q}) = \int_{\mathcal{Q}} \widehat{\gamma}_k(\mathsf{q}) \widehat{H}\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) b_\ell(\mathsf{q}) \,\mathrm{d}\nu(\mathsf{q}), \tag{16a}$$

$$\int_{\mathcal{Q}} \widehat{\beta}_j(\mathsf{q}) c_\ell(\mathsf{q})^* H\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\nu(\mathsf{q}) = \int_{\mathcal{Q}} \widehat{\beta}_j(\mathsf{q}) c_\ell(\mathsf{q})^* \widehat{H}\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\nu(\mathsf{q}), \tag{16b}$$

$$\int_{\mathcal{Q}} \widehat{\alpha}_i(\mathsf{q}) c_\ell(\mathsf{q})^* \frac{\partial H}{\partial s}\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) b_\ell(\mathsf{q}) \,\mathrm{d}\nu(\mathsf{q}) = \int_{\mathcal{Q}} \widehat{\alpha}_i(\mathsf{q}) c_\ell(\mathsf{q})^* \frac{\partial \widehat{H}}{\partial s}\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) b_\ell(\mathsf{q}) \,\mathrm{d}\nu(\mathsf{q}), \tag{16c}$$

for $\ell = 1, 2, \ldots, r$, $k = 1, 2, \ldots, n_{\widehat{\mathcal{C}}}$, $j = 1, 2, \ldots, n_{\widehat{\mathcal{B}}}$, $i = 1, 2, \ldots n_{\widehat{\mathcal{A}}}$, where $\widehat{\alpha}_i$, $\widehat{\beta}_j$, and $\widehat{\gamma}_k$ are as defined in (5).

Proof. The pole-residue form (15) follows from the general diagonal pole-residue form (9), with $k_\ell(s, \mathsf{q}) = s - \lambda_\ell(\mathsf{q})$, $f_\ell(s, \mathsf{q}) = b_\ell(\mathsf{q})$, and $g_\ell(s, \mathsf{q}) = c_\ell(\mathsf{q})$. Then, with this structure, optimality conditions (16) follow from diagonal conditions after applying the Cauchy integral formula. For instance, the left-hand side of the right tangential Lagrange condition (12a) becomes

$$\begin{aligned}
\int_{\mathcal{P}} \frac{\overline{\widehat{\gamma}_k(\mathsf{p})} H(\mathsf{p}) b_\ell(\mathsf{p})}{\overline{a_\ell(\mathsf{p})}} \,\mathrm{d}\mu(\mathsf{p}) &= \int_{\mathcal{Q}} \int_{-\infty}^{\infty} \frac{\widehat{\gamma}_k(\mathsf{q}) H(\imath\omega, \mathsf{q}) b_\ell(\mathsf{q})}{\overline{\imath\omega - \lambda_\ell(\mathsf{q})}} \,\mathrm{d}\omega \,\mathrm{d}\nu(\mathsf{q}) \\
&= \int_{\mathcal{Q}} \int_{-\infty}^{\infty} \frac{\widehat{\gamma}_k(\mathsf{q}) H(\imath\omega, \mathsf{q}) b_\ell(\mathsf{q})}{-\imath\omega - \overline{\lambda_\ell(\mathsf{q})}} \,\mathrm{d}\omega \,\mathrm{d}\nu(\mathsf{q}) \\
&= \frac{1}{\imath} \int_{\mathcal{Q}} \oint_{\imath\mathbb{R}} \frac{\widehat{\gamma}_k(\mathsf{q}) H(s, \mathsf{q}) b_\ell(\mathsf{q})}{-s - \overline{\lambda_\ell(\mathsf{q})}} \,\mathrm{d}s \,\mathrm{d}\nu(\mathsf{q}) \\
&= \frac{2\pi}{\imath} \int_{\mathcal{Q}} \widehat{\gamma}_k(\mathsf{q}) H\left(-\overline{\lambda_\ell(\mathsf{q})}, \mathsf{q}\right) b_\ell(\mathsf{q}) \,\mathrm{d}\nu(\mathsf{q}),
\end{aligned}$$

which yields (16a). The remaining two conditions (16b)–(16c) follow similarly from (12b) and (12c). □

Recall that $\mathcal{H}_2$-optimal approximation of *non-parametric* LTI systems requires bitangential Hermite interpolation of the FOM transfer function H at the mirror images of the reduced-order poles [8, 1]. We showed in our earlier papers [15, 16, 14] that bitangential Hermite interpolation as necessary conditions for optimality extends to many other $\mathcal{H}_2/\mathcal{L}_2$ approximation settings as well. Even though the $\mathcal{H}_2 \otimes \mathcal{L}_2$ optimality conditions (16) derived here have an integral form, they still have a similar bitangential Hermite interpolation structure as before. To arrive at this more familiar form of bitangential Hermite interpolations, we need to have explicit expressions for the functions $\widehat{\alpha}_i, \widehat{\beta}_j, \widehat{\gamma}_k, \lambda_\ell, b_\ell, c_\ell$. In the next two sections we focus on such cases.

4. Parameters in inputs and outputs

In [2], the authors considered parametric LTI systems with parameters only in $\widehat{\mathcal{B}}$ and $\widehat{\mathcal{C}}$; specifically, the ROM of the form

$$\widehat{\mathcal{E}}(\mathsf{q}) = I,\ \ \widehat{\mathcal{A}}(\mathsf{q}) = \widehat{A},\ \ \widehat{\mathcal{B}}(\mathsf{q}) = \widehat{B}_1 + \mathsf{q}_1 \widehat{B}_2,\ \ \widehat{\mathcal{C}}(\mathsf{q}) = \widehat{C}_1 + \mathsf{q}_2 \widehat{C}_2, \tag{17}$$

with $\widehat{A} = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$ and $\mathcal{Q} = [0, 1]^2$. Therefore,

$$\begin{aligned}
\mathsf{q} = (\mathsf{q}_1, \mathsf{q}_2),\ n_{\widehat{\mathcal{A}}} = 1,\ \widehat{\alpha}_1(\mathsf{q}) = 1, \\
n_{\widehat{\mathcal{B}}} = 2,\ \widehat{\beta}_1(\mathsf{q}) = 1,\ \widehat{\beta}_2(\mathsf{q}) = \mathsf{q}_1, \\
n_{\widehat{\mathcal{C}}} = 2,\ \widehat{\gamma}_1(\mathsf{q}) = 1,\ \widehat{\gamma}_2(\mathsf{q}) = \mathsf{q}_2, \\
\lambda_\ell(\mathsf{q}) = \lambda_\ell,\ b_\ell(\mathsf{q}) = b_{\ell,1} + \mathsf{q}_1 b_{\ell,2},\ c_\ell(\mathsf{q}) = c_{\ell,1} + \mathsf{q}_2 c_{\ell,2},
\end{aligned}$$

where $b_{\ell,i} = \widehat{B}_i^* e_\ell$ and $c_{\ell,i} = \widehat{C}_i e_\ell$ for $i = 1, 2$. In [2], only single-input single-output (SISO) systems were considered. Here we consider multiple-input multiple-output (MIMO) systems. (Further extensions are possible, see e.g., [11].)

Note that the reduced transfer function is bilinear in terms of the parameters q_1 and q_2:

$$\widehat{H}(s, \mathsf{q}) = \left(\widehat{C}_1 + \mathsf{q}_2 \widehat{C}_2\right) \left(sI - \widehat{A}\right)^{-1} \left(\widehat{B}_1 + \mathsf{q}_1 \widehat{B}_2\right) = \widehat{H}_{11}(s) + \mathsf{q}_1 \widehat{H}_{12}(s) + \mathsf{q}_2 \widehat{H}_{21}(s) + \mathsf{q}_1 \mathsf{q}_2 \widehat{H}_{22}(s) \tag{18}$$

where

$$\widehat{H}_{ij}(s) = \widehat{C}_i \left(sI - \widehat{A}\right)^{-1} \widehat{B}_j, \quad i, j \in \{1, 2\}.$$

We assume the same form for the full transfer function

$$H(s, \mathsf{q}) = H_{11}(s) + \mathsf{q}_1 H_{12}(s) + \mathsf{q}_2 H_{21}(s) + \mathsf{q}_1 \mathsf{q}_2 H_{22}(s), \tag{19}$$

where $H_{ij} \in \mathcal{H}_2$ for $i, j \in \{1, 2\}$. However, contrary to [2], we do not need to assume that H_{ij} has a finite-dimensional state space, i.e., they can contain non-rational terms. We only require the ROM to have a finite-dimensional state space. Thus, the theory we develop applies not only to MIMO systems but also to irrational transfer functions.

Following [2], we define the auxiliary transfer functions

$$\mathcal{H}(s) = \begin{bmatrix} H_{11}(s) & H_{12}(s) \\ H_{21}(s) & H_{22}(s) \end{bmatrix}, \tag{20a}$$

$$\widehat{\mathcal{H}}(s) = \begin{bmatrix} \widehat{H}_{11}(s) & \widehat{H}_{12}(s) \\ \widehat{H}_{21}(s) & \widehat{H}_{22}(s) \end{bmatrix} = \begin{bmatrix} \widehat{C}_1 \\ \widehat{C}_2 \end{bmatrix} \left(sI - \widehat{A}\right)^{-1} \begin{bmatrix} \widehat{B}_1 & \widehat{B}_2 \end{bmatrix}. \tag{20b}$$

Note that H and $\widehat{H}$ can be obtained from $\mathcal{H}$ and $\widehat{\mathcal{H}}$ via

$$H(s, \mathsf{q}) = \begin{bmatrix} I_{n_\mathrm{o}} & \mathsf{q}_2 I_{n_\mathrm{o}} \end{bmatrix} \mathcal{H}(s) \begin{bmatrix} I_{n_\mathrm{i}} \\ \mathsf{q}_1 I_{n_\mathrm{i}} \end{bmatrix}, \tag{21a}$$

$$\widehat{H}(s, \mathsf{q}) = \begin{bmatrix} I_{n_\mathrm{o}} & \mathsf{q}_2 I_{n_\mathrm{o}} \end{bmatrix} \widehat{\mathcal{H}}(s) \begin{bmatrix} I_{n_\mathrm{i}} \\ \mathsf{q}_1 I_{n_\mathrm{i}} \end{bmatrix}. \tag{21b}$$

We obtain the following result, where we take ν to be the Lebesgue measure over $\mathcal{Q}$ (as in [2]).

Theorem 3. *Let $H, \widehat{H}$ be as in (19) and (17) and $\mathcal{H}, \widehat{\mathcal{H}}$ as in (20a) and (20b). Furthermore, let $\widehat{H}$ be an $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal ROM for H. Define*

$$\mathfrak{b}_\ell = \begin{bmatrix} I_{n_\mathrm{i}} & \frac{1}{2} I_{n_\mathrm{i}} \\ \frac{1}{2} I_{n_\mathrm{i}} & \frac{1}{3} I_{n_\mathrm{i}} \end{bmatrix} \begin{bmatrix} b_{\ell,1} \\ b_{\ell,2} \end{bmatrix}, \quad \mathfrak{c}_\ell = \begin{bmatrix} I_{n_\mathrm{o}} & \frac{1}{2} I_{n_\mathrm{o}} \\ \frac{1}{2} I_{n_\mathrm{o}} & \frac{1}{3} I_{n_\mathrm{o}} \end{bmatrix} \begin{bmatrix} c_{\ell,1} \\ c_{\ell,2} \end{bmatrix}. \tag{22}$$

Then for $\ell = 1, 2, \ldots, r$, we have

$$\mathcal{H}\left(-\overline{\lambda_\ell}\right) \mathfrak{b}_\ell = \widehat{\mathcal{H}}\left(-\overline{\lambda_\ell}\right) \mathfrak{b}_\ell, \tag{23a}$$

$$\mathfrak{c}_\ell^* \mathcal{H}\left(-\overline{\lambda_\ell}\right) = \mathfrak{c}_\ell^* \widehat{\mathcal{H}}\left(-\overline{\lambda_\ell}\right), \tag{23b}$$

$$\mathfrak{c}_\ell^* \mathcal{H}'\left(-\overline{\lambda_\ell}\right) \mathfrak{b}_\ell = \mathfrak{c}_\ell^* \widehat{\mathcal{H}}'\left(-\overline{\lambda_\ell}\right) \mathfrak{b}_\ell. \tag{23c}$$

Proof. Based on the conditions in (16), we need to compute the integrals

$$\int_{\mathcal{Q}} H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) b_\ell(\mathsf{q}) \, \mathsf{d}\mathsf{q}, \quad \int_{\mathcal{Q}} \mathsf{q}_2 H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) b_\ell(\mathsf{q}) \, \mathsf{d}\mathsf{q}, \quad \int_{\mathcal{Q}} c_\ell(\mathsf{q})^* H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) \mathsf{d}\mathsf{q},$$
$$\int_{\mathcal{Q}} \mathsf{q}_1 c_\ell(\mathsf{q})^* H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) \mathsf{d}\mathsf{q}, \quad \int_{\mathcal{Q}} c_\ell(\mathsf{q})^* \frac{\partial H}{\partial s}\left(-\overline{\lambda_\ell}, \mathsf{q}\right) b_\ell(\mathsf{q}) \, \mathsf{d}\mathsf{q},$$

and similarly with $\widehat{H}$. Starting with the first, we find that

$$\begin{aligned}
\int_{\mathcal{Q}} H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) b_\ell(\mathsf{q}) \, \mathsf{d}\mathsf{q} &= \int_{\mathcal{Q}} \left(H_{11}\left(-\overline{\lambda_\ell}\right) + \mathsf{q}_1 H_{12}\left(-\overline{\lambda_\ell}\right) + \mathsf{q}_2 H_{21}\left(-\overline{\lambda_\ell}\right) + \mathsf{q}_1 \mathsf{q}_2 H_{22}\left(-\overline{\lambda_\ell}\right)\right) \left(b_{\ell,1} + \mathsf{q}_1 b_{\ell,2}\right) \mathsf{d}\mathsf{q} \\
&= H_{11}\left(-\overline{\lambda_\ell}\right) b_{\ell,1} + \frac{1}{2} H_{12}\left(-\overline{\lambda_\ell}\right) b_{\ell,1} + \frac{1}{2} H_{21}\left(-\overline{\lambda_\ell}\right) b_{\ell,1} + \frac{1}{4} H_{22}\left(-\overline{\lambda_\ell}\right) b_{\ell,1} \\
&\quad + \frac{1}{2} H_{11}\left(-\overline{\lambda_\ell}\right) b_{\ell,2} + \frac{1}{3} H_{12}\left(-\overline{\lambda_\ell}\right) b_{\ell,2} + \frac{1}{4} H_{21}\left(-\overline{\lambda_\ell}\right) b_{\ell,2} + \frac{1}{6} H_{22}\left(-\overline{\lambda_\ell}\right) b_{\ell,2} \\
&= \begin{bmatrix} I_{n_\mathrm{o}} & \frac{1}{2} I_{n_\mathrm{o}} \end{bmatrix} \mathcal{H}\left(-\overline{\lambda_\ell}\right) \begin{bmatrix} I_{n_\mathrm{i}} & \frac{1}{2} I_{n_\mathrm{i}} \\ \frac{1}{2} I_{n_\mathrm{i}} & \frac{1}{3} I_{n_\mathrm{i}} \end{bmatrix} \begin{bmatrix} b_{\ell,1} \\ b_{\ell,2} \end{bmatrix},
\end{aligned}$$

where we used (19) and (21a). The second integral becomes

$$\int_{\mathcal{Q}} \mathsf{q}_2 H\left(-\overline{\lambda_\ell}, \mathsf{q}\right) b_\ell(\mathsf{q}) \, \mathsf{d}\mathsf{q} = \begin{bmatrix} \frac{1}{2} I_{n_\mathrm{o}} & \frac{1}{3} I_{n_\mathrm{o}} \end{bmatrix} \mathcal{H}\left(-\overline{\lambda_\ell}\right) \begin{bmatrix} I_{n_\mathrm{i}} & \frac{1}{2} I_{n_\mathrm{i}} \\ \frac{1}{2} I_{n_\mathrm{i}} & \frac{1}{3} I_{n_\mathrm{i}} \end{bmatrix} \begin{bmatrix} b_{\ell,1} \\ b_{\ell,2} \end{bmatrix}.$$

Stacking these two vertically (and using (22)) gives us the left-hand side in the right Lagrange tangential condition (23a). The other conditions follow similarly. □

Therefore, for this special case of (17), we obtain a more familiar bitangential Hermite interpolation. More specifically, $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling of H with $\widehat{H}$ in (17) is equivalent to a *weighted* $\mathcal{H}_2$-optimal reduced-order modeling for $\mathcal{H}$ with $\widehat{\mathcal{H}}$. Thus, our general framework in Theorem 2 not only recovers the results from [2] but also extends them to MIMO systems and eliminates the need for the FOM to have a rational transfer function.

Interpolatory conditions of Theorem 3 are in terms of the parametric function $\widehat{\mathcal{H}}$, not the original parametric transfer function H. The next result gives explicit interpolatory conditions in terms of H for SISO systems.

Corollary 1. *Let the assumptions in Theorem 3 hold. Furthermore, let $n_{\mathrm{i}} = n_{\mathrm{o}} = 1$ such that $\mathfrak{b}_\ell, \mathfrak{c}_\ell \in \mathbb{C}^2$ and denote their components as*

$$\mathfrak{b}_\ell = \begin{bmatrix} \mathfrak{b}_{\ell,1} \\ \mathfrak{b}_{\ell,2} \end{bmatrix} \quad \textit{and} \quad \mathfrak{c}_\ell = \begin{bmatrix} \mathfrak{c}_{\ell,1} \\ \mathfrak{c}_{\ell,2} \end{bmatrix}.$$

If $\mathfrak{b}_{\ell,1} \neq 0$ and $\mathfrak{c}_{\ell,1} \neq 0$, then

$$\begin{aligned}
H\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \mathsf{q}_2\right) &= \widehat{H}\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \mathsf{q}_2\right), \\
\partial_{\mathsf{q}_2} H\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \mathsf{q}_2\right) &= \partial_{\mathsf{q}_2} \widehat{H}\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \mathsf{q}_2\right), \\
H\left(-\overline{\lambda_\ell}, \mathsf{q}_1, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right) &= \widehat{H}\left(-\overline{\lambda_\ell}, \mathsf{q}_1, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right), \\
\partial_{\mathsf{q}_1} H\left(-\overline{\lambda_\ell}, \mathsf{q}_1, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right) &= \partial_{\mathsf{q}_1} \widehat{H}\left(-\overline{\lambda_\ell}, \mathsf{q}_1, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right), \\
\partial_s H\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right) &= \partial_s \widehat{H}\left(-\overline{\lambda_\ell}, \frac{\mathfrak{b}_{\ell,2}}{\mathfrak{b}_{\ell,1}}, \overline{\frac{\mathfrak{c}_{\ell,2}}{\mathfrak{c}_{\ell,1}}}\right),
\end{aligned}$$

for all $\mathsf{q}_1, \mathsf{q}_2 \in \mathbb{C}$ and $\ell = 1, 2, \ldots, r$.

Proof. The proof follows directly from (23) using (21) and

$$\partial_{\mathsf{q}_2} H(s, \mathsf{q}) = \begin{bmatrix} 0 & 1 \end{bmatrix} \mathcal{H}(s) \begin{bmatrix} 1 \\ \mathsf{q}_1 \end{bmatrix}$$

and similar expressions for $\partial_{\mathsf{q}_2} \widehat{H}$, $\partial_{\mathsf{q}_1} H$, and $\partial_{\mathsf{q}_1} \widehat{H}$. □

This states that $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality for the full-order and reduced-order structure in (19) and (18) requires that, in the parameter space, interpolation be enforced over lines rather than at only a finite number of points (as generically done for parametric systems). Furthermore, note that partial derivatives with respect to the parameters $\mathsf{q}_1, \mathsf{q}_2$ also appear in the optimality conditions, thus having a similar role as the Laplace variable s in the necessary optimality conditions.

5. Parameter in dynamics

In the previous section, we considered a special case where the parametric dependence was only in $\widehat{\mathcal{B}}$ and $\widehat{\mathcal{C}}$. Now, we consider the case where $\mathcal{Q} = [a, b] \subset \mathbb{R}$, $a < b$, and the FOM and ROM have the form

$$\mathcal{E}(\mathsf{q}) = I, \quad \mathcal{A}(\mathsf{q}) = A_1 + \mathsf{q} A_2, \quad \mathcal{B}(\mathsf{q}) = B, \quad \mathcal{C}(\mathsf{q}) = C \tag{24}$$

and

$$\widehat{\mathcal{E}}(\mathsf{q}) = I, \quad \widehat{\mathcal{A}}(\mathsf{q}) = \widehat{A}_1 + \mathsf{q} \widehat{A}_2, \quad \widehat{\mathcal{B}}(\mathsf{q}) = \widehat{B}, \quad \widehat{\mathcal{C}}(\mathsf{q}) = \widehat{C}, \tag{25}$$

respectively, with

$$A_k = \operatorname{diag}(\nu_{k,1}, \nu_{k,2}, \ldots, \nu_{k,n}) \quad \text{and} \quad \widehat{A}_k = \operatorname{diag}(\lambda_{k,1}, \lambda_{k,2}, \ldots, \lambda_{k,r}),$$

for $k = 1, 2$. In other words, we assume the parametric dependencies appear only in the dynamics matrices $\mathcal{A}$ and $\widehat{\mathcal{A}}$ and they are both composed of only two terms which are simultaneously diagonalizable. With these parametric forms, the full-order and reduced-order transfer functions have the pole-residue forms

$$H(s, \mathsf{q}) = \sum_{i=1}^{n} \frac{\Phi_i}{s - \nu_i(\mathsf{q})}, \quad \widehat{H}(s, \mathsf{q}) = \sum_{i=1}^{r} \frac{c_i b_i^*}{s - \lambda_i(\mathsf{q})}, \tag{26}$$

where

$$\nu_i(\mathsf{q}) = \nu_{1,i} + \mathsf{q}\nu_{2,i} \quad \text{and} \quad \lambda_i(\mathsf{q}) = \lambda_{1,i} + \mathsf{q}\lambda_{2,i}. \tag{27}$$

Let ν be the Lebesgue measure over $[a, b]$. Furthermore, for any $\sigma_a, \sigma_b \in \mathbb{C}_-$, define $f_{\sigma_a,\sigma_b} \colon \mathbb{C}_+^2 \to \mathbb{C}$ as

$$f_{\sigma_a,\sigma_b}(s_a, s_b) = \frac{b - a}{(s_b - \sigma_b) - (s_a - \sigma_a)} \ln\left(\frac{s_b - \sigma_b}{s_a - \sigma_a}\right). \tag{28}$$

Additionally, define the functions $G, \widehat{G} \colon \mathbb{C}_+^2 \to \mathbb{C}^{n_\mathrm{o} \times n_\mathrm{i}}$ by

$$G(s_a, s_b) = \sum_{i=1}^{n} f_{\nu_i(a),\nu_i(b)}(s_a, s_b)\Phi_i \quad \text{and} \tag{29a}$$

$$\widehat{G}(s_a, s_b) = \sum_{i=1}^{r} f_{\lambda_i(a),\lambda_i(b)}(s_a, s_b) c_i b_i^*. \tag{29b}$$

Note that G and $\widehat{G}$ depend on the pole-residue forms (26) of H and $\widehat{H}$, respectively. Thus, one can consider G as the full-order *modified* function and $\widehat{G}$ the reduced-order one. Based on this setup, we are ready to state the interpolatory optimality conditions in this setting.

Theorem 4. *Let H and $\widehat{H}$ be as given in (26) and let G and $\widehat{G}$ be as defined in (29). If $\widehat{H}$ is an $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal D-StROM for H, then*

$$G\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i = \widehat{G}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i, \tag{30a}$$

$$c_i^* G\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) = c_i^* \widehat{G}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right), \tag{30b}$$

$$c_i^* \frac{\partial G}{\partial s_a}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i = c_i^* \frac{\partial \widehat{G}}{\partial s_a}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i, \tag{30c}$$

$$c_i^* \frac{\partial G}{\partial s_b}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i = c_i^* \frac{\partial \widehat{G}}{\partial s_b}\left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\right) b_i, \tag{30d}$$

for $i = 1, 2, \ldots, r$.

Proof. Due to the special form of the ROM in (25) and its transfer function $\widehat{H}$ in (26), the quantities $b_\ell, c_\ell, \widehat{\gamma}_k, \widehat{\beta}_j$ in (15) and (16) in Theorem 2 are parameter-independent. Thus, to analyze the first Lagrange conditions (16a) and (16b) for the special case (24) and (25), it is enough to focus on the integrals

$$\int_{\mathcal{Q}} H\left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\nu(\mathsf{q}) \quad \text{and} \quad \int_{\mathcal{Q}} \widehat{H}\left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\nu(\mathsf{q}).$$

We start with the first integral involving H. Using the pole-residue form of H from (26) and the expression for $\lambda_i(\mathsf{q})$ from (27), we obtain

$$\begin{aligned}
\int_{\mathcal{Q}} H\left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\nu(\mathsf{q}) = \int_a^b H\left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\right) \mathrm{d}\mathsf{q} &= \int_a^b \sum_{j=1}^{n} \frac{\Phi_j}{-\overline{\lambda_i(\mathsf{q})} - \nu_{1,j} - \mathsf{q}\nu_{2,j}} \mathrm{d}\mathsf{q} \\
&= \sum_{j=1}^{n} \int_a^b \frac{\Phi_j}{-\overline{\lambda_{1,i}} - \mathsf{q}\overline{\lambda_{2,i}} - \nu_{1,j} - \mathsf{q}\nu_{2,j}} \mathrm{d}\mathsf{q} \\
&= \sum_{j=1}^{n} \int_a^b \frac{\Phi_j}{-\overline{\lambda_{1,i}} - \nu_{1,j} + \mathsf{q}\left(-\overline{\lambda_{2,i}} - \nu_{2,j}\right)} \mathrm{d}\mathsf{q}.
\end{aligned}$$

Then integrating the last equality gives

$$\begin{aligned}\int_{\mathcal{Q}} H\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\nu(\mathsf{q}) &= \sum_{j=1}^{n} \frac{\Phi_j}{-\overline{\lambda_{2,i}} - \nu_{2,j}} \ln\left(\frac{-\overline{\lambda_{1,i}} - \nu_{1,j} + b\Big(-\overline{\lambda_{2,i}} - \nu_{2,j}\Big)}{-\overline{\lambda_{1,i}} - \nu_{1,j} + a\Big(-\overline{\lambda_{2,i}} - \nu_{2,j}\Big)}\right) \\ &= \sum_{j=1}^{n} \frac{\Phi_j(b-a)}{\Big(-\overline{\lambda_i(b)} - \nu_j(b)\Big) - \Big(-\overline{\lambda_i(a)} - \nu_j(a)\Big)} \ln\left(\frac{-\overline{\lambda_i(b)} - \nu_j(b)}{-\overline{\lambda_i(a)} - \nu_j(a)}\right) \\ &= G\Big(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\Big).\end{aligned}$$

Similarly, one can show that

$$\int_{\mathcal{Q}} \widehat{H}\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\nu(\mathsf{q}) = \widehat{G}\Big(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\Big).$$

Therefore, the first two optimality conditions (16a) and (16b) in Theorem 2 lead to the interpolatory conditions (30a) and (30b).

To derive the remaining two conditions (30c) and (30d) from (16c), we now consider the integrals

$$\int_{\mathcal{Q}} \mathsf{q}^k \frac{\partial H}{\partial s}\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\nu(\mathsf{q}), \quad k = 0, 1. \tag{31}$$

We will need the expressions for the partial derivatives of G. It directly follows from (29a) that

$$\frac{\partial G}{\partial s_a}(s_a, s_b) = \sum_{i=1}^{n} \frac{\partial f_{\nu_i(a),\nu_i(b)}}{\partial s_a}(s_a, s_b)\Phi_i. \tag{32}$$

Similar expressions hold for $\frac{\partial G}{\partial s_b}(s_a, s_b)$, $\frac{\partial \widehat{G}}{\partial s_a}(s_a, s_b)$, and $\frac{\partial \widehat{G}}{\partial s_b}(s_a, s_b)$ as well. Thus, to compute these partial derivatives, we simply focus on f_{σ_a,σ_b} and obtain, via direct differentiation of (28), that

$$\frac{\partial f_{\sigma_a,\sigma_b}}{\partial s_a}(s_a, s_b) = \frac{b-a}{((s_b-\sigma_b)-(s_a-\sigma_a))^2} \ln\left(\frac{s_b-\sigma_b}{s_a-\sigma_a}\right) - \frac{b-a}{(s_b-\sigma_b)-(s_a-\sigma_a)} \cdot \frac{1}{s_a-\sigma_a}, \quad \text{and}$$

$$\frac{\partial f_{\sigma_a,\sigma_b}}{\partial s_b}(s_a, s_b) = -\frac{b-a}{((s_b-\sigma_b)-(s_a-\sigma_a))^2} \ln\left(\frac{s_b-\sigma_b}{s_a-\sigma_a}\right) + \frac{b-a}{(s_b-\sigma_b)-(s_a-\sigma_a)} \cdot \frac{1}{s_b-\sigma_b}.$$

Using the pole-residue form of H from (26) in the first integral in (31) gives

$$\begin{aligned}\int_{\mathcal{Q}} \frac{\partial H}{\partial s}\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\nu(\mathsf{q}) &= \int_a^b \frac{\partial H}{\partial s}\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\mathsf{q} = \int_a^b \sum_{j=1}^{n} \frac{-\Phi_j}{\Big(-\overline{\lambda_i(\mathsf{q})} - \nu_{1,j} - \mathsf{q}\nu_{2,j}\Big)^2} \,\mathrm{d}\mathsf{q} \\ &= \sum_{j=1}^{n} \int_a^b \frac{-\Phi_j}{\Big(-\overline{\lambda_{1,i}} - \nu_{1,j} + \mathsf{q}\Big(-\overline{\lambda_{2,i}} - \nu_{2,j}\Big)\Big)^2} \,\mathrm{d}\mathsf{q} \\ &= \sum_{j=1}^{n} \frac{\Phi_j}{-\overline{\lambda_{2,i}} - \nu_{2,j}} \left(\frac{1}{-\overline{\lambda_{1,i}} - \nu_{1,j} + b\Big(-\overline{\lambda_{2,i}} - \nu_{2,j}\Big)} - \frac{1}{-\overline{\lambda_{1,i}} - \nu_{1,j} + a\Big(-\overline{\lambda_{2,i}} - \nu_{2,j}\Big)}\right).\end{aligned}$$

After various algebraic manipulations to replace $\lambda_{k,i}$ and $\nu_{k,j}$ by $\lambda_i(\cdot)$ and $\nu_j(\cdot)$ and using (32), we obtain

$$\begin{aligned}\int_{\mathcal{Q}} \frac{\partial H}{\partial s}\Big(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q}\Big) \,\mathrm{d}\nu(\mathsf{q}) &= \sum_{j=1}^{n} \frac{\Phi_j(b-a)}{\Big(-\overline{\lambda_i(b)} - \nu_j(b)\Big) - \Big(-\overline{\lambda_i(a)} - \nu_j(a)\Big)} \\ &\quad \times \left(\frac{1}{-\overline{\lambda_i(b)} - \nu_j(b)} - \frac{1}{-\overline{\lambda_i(a)} - \nu_j(a)}\right) \\ &= \frac{\partial G}{\partial s_a}\Big(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\Big) + \frac{\partial G}{\partial s_b}\Big(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)}\Big).\end{aligned}$$

Following the same derivations, one obtains similar expressions involving $\widehat{H}$ and $\widehat{G}$, which shows that

$$\begin{aligned} &c_i^* \left(\frac{\partial G}{\partial s_a} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) + \frac{\partial G}{\partial s_b} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) \right) b_i \\ &= c_i^* \left(\frac{\partial \widehat{G}}{\partial s_a} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) + \frac{\partial \widehat{G}}{\partial s_b} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) \right) b_i. \end{aligned} \tag{33}$$

Focusing on the second integral in (31), we find

$$\begin{aligned} \int_{\mathcal{Q}} \mathsf{q} \frac{\partial H}{\partial s} \left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q} \right) \mathrm{d}\nu(\mathsf{q}) &= \int_a^b \mathsf{q} \frac{\partial H}{\partial s} \left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q} \right) \mathrm{d}\mathsf{q} = \sum_{j=1}^n \int_a^b \frac{-\mathsf{q}\Phi_j}{\left(-\overline{\lambda_{1,i}} - \nu_{1,j} + \mathsf{q}\left(-\overline{\lambda_{2,i}} - \nu_{2,j} \right) \right)^2} \mathrm{d}\mathsf{q} \\ &= \sum_{j=1}^n \frac{-\Phi_j}{\left(-\overline{\lambda_{2,i}} - \nu_{2,j} \right)^2} \left(\frac{-\overline{\lambda_{1,i}} - \nu_{1,j}}{-\overline{\lambda_i(b)} - \nu_j(b)} - \frac{-\overline{\lambda_{1,i}} - \nu_{1,j}}{-\overline{\lambda_i(a)} - \nu_j(a)} + \ln\left(\frac{-\overline{\lambda_i(b)} - \nu_j(b)}{-\overline{\lambda_i(a)} - \nu_j(a)} \right) \right). \end{aligned}$$

After more tedious algebraic manipulations, we obtain

$$\begin{aligned} \int_{\mathcal{Q}} \mathsf{q} \frac{\partial H}{\partial s} \left(-\overline{\lambda_i(\mathsf{q})}, \mathsf{q} \right) \mathrm{d}\nu(\mathsf{q}) &= \sum_{j=1}^n \frac{\Phi_j(b-a)}{\left(-\overline{\lambda_i(b)} - \nu_j(b) \right) - \left(-\overline{\lambda_i(a)} - \nu_j(a) \right)} \left(\frac{b}{-\overline{\lambda_i(b)} - \nu_j(b)} - \frac{a}{-\overline{\lambda_i(a)} - \nu_j(a)} \right) \\ &\quad + \sum_{j=1}^n \frac{-\Phi_j(b-a)^2}{\left(\left(-\overline{\lambda_i(b)} - \nu_j(b) \right) - \left(-\overline{\lambda_i(a)} - \nu_j(a) \right) \right)^2} \ln\left(\frac{-\overline{\lambda_i(b)} - \nu_j(b)}{-\overline{\lambda_i(a)} - \nu_j(a)} \right) \\ &= a \frac{\partial G}{\partial s_a} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) + b \frac{\partial G}{\partial s_b} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right). \end{aligned}$$

As before, with similar expressions involving $\widehat{H}$ and $\widehat{G}$, we obtain

$$\begin{aligned} &c_i^* \left(a \frac{\partial G}{\partial s_a} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) + b \frac{\partial G}{\partial s_b} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) \right) b_i \\ &= c_i^* \left(a \frac{\partial \widehat{G}}{\partial s_a} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) + b \frac{\partial \widehat{G}}{\partial s_b} \left(-\overline{\lambda_i(a)}, -\overline{\lambda_i(b)} \right) \right) b_i. \end{aligned} \tag{34}$$

Then (33) and (34) give the last two optimality conditions (30c) and (30d), thus concluding the proof. □

Theorem 4 proves that for this class of parametric LTI systems, bitangential Hermite interpolation, once again, forms the foundation of the $\mathcal{L}_2$-optimal approximation. The interpolation is based on a modified, two-variable transfer function G, and has to be enforced at the reflected boundary values of the poles. This is the first such result for parametric LTI systems where the system poles vary with the parameters. Therefore, we have extended the classical bitangential Hermite interpolation conditions from non-parametric $\mathcal{H}_2$-optimal approximation to parametric $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal approximation.

6. Numerical experiments

In this section, we numerically demonstrate Theorem 4 on two numerical examples. In both cases, we start with FOM of the form in (24) and numerically find a locally $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal ROM of the form in (25) via the BFGS method implemented in SciPy [22]. The gradients are computed based on the expressions in [9] and using SciPy's numerical quadrature. The $\mathcal{H}_2$ norms are computed using pyMOR [17]. Then Theorem 4 is verified by computing the relative errors in (30), i.e., computing the absolute difference between the left- and the right-hand sides and dividing by the absolute value of the left-hand side (since we focus on SISO systems in the numerical examples, there is only one Lagrange interpolation condition). In other words, in both cases we show that optimal D-StROMs satisfy the developed interpolatory optimal conditions.

The code to reproduce the results is available in [13].

6.1. Synthetic parametric model

We first consider a variant of the synthetic parametric model from MOR Wiki [21]. In particular, we consider FOM of order 6 (instead of 100), i.e., $\mathcal{E}(\mathsf{q}) = I$,

$$\mathcal{A}(\mathsf{q}) = \begin{bmatrix} -10\mathsf{q} & 10 & & & & \\ -10 & -10\mathsf{q} & & & & \\ & & -30\mathsf{q} & 30 & & \\ & & -30 & -30\mathsf{q} & & \\ & & & & 50\mathsf{q} & 50 \\ & & & & -50 & 50\mathsf{q} \end{bmatrix}, \quad \mathcal{B}(\mathsf{q}) = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \quad \mathcal{C}(\mathsf{q}) = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix},$$

over the parameter space $\mathcal{P} = [\frac{1}{50}, 1]$. To obtain a locally $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal D-StROM of order 4, we enforce the complex diagonal structure with $\widehat{\mathcal{E}}(\mathsf{q}) = I$,

$$\widehat{\mathcal{A}}(\mathsf{q}) = \begin{bmatrix} x_1 + \mathsf{q}x_2 & x_3 + \mathsf{q}x_4 & & \\ -x_3 - \mathsf{q}x_4 & x_1 + \mathsf{q}x_2 & & \\ & & x_5 + \mathsf{q}x_6 & x_7 + \mathsf{q}x_8 \\ & & -x_7 - \mathsf{q}x_8 & x_5 + \mathsf{q}x_6 \end{bmatrix}, \quad \widehat{\mathcal{B}}(\mathsf{q}) = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \quad \widehat{\mathcal{C}}(\mathsf{q}) = \begin{bmatrix} x_9 & x_{10} & x_{11} & x_{12} \end{bmatrix}.$$

Note that for this example, H and $\widehat{H}$ are of the form in (26), i.e., A_1 and A_2 in $\mathcal{A}(\mathsf{q}) = A_1 + \mathsf{q}A_2$ are simultaneously diagonalizable, as well as $\widehat{A}_1$ and $\widehat{A}_2$ in $\widehat{\mathcal{A}}(\mathsf{q}) = \widehat{A}_1 + \mathsf{q}\widehat{A}_2$. Therefore, the assumptions of Theorem 4 are satisfied.

We initialize the BFGS-based minimization with the FOM truncated to the first 4 states. Upon the convergence of BFGS, we obtain

$$\begin{aligned} \widehat{\mathcal{A}}(\mathsf{q}) &= \operatorname{diag}\left(\widehat{\mathcal{A}}_1(\mathsf{q}), \widehat{\mathcal{A}}_2(\mathsf{q})\right), \\ \widehat{\mathcal{A}}_1(\mathsf{q}) &= \begin{bmatrix} -7.0213 \times 10^{-3} - 11.014\mathsf{q} & 9.9975 + 0.24074\mathsf{q} \\ -9.9975 - 0.24074\mathsf{q} & -7.0213 \times 10^{-3} - 11.014\mathsf{q} \end{bmatrix}, \\ \widehat{\mathcal{A}}_2(\mathsf{q}) &= \begin{bmatrix} -1.6795 - 39.184\mathsf{q} & 29.261 + 0.95464\mathsf{q} \\ -29.261 - 0.95464\mathsf{q} & -1.6795 - 39.184\mathsf{q} \end{bmatrix}, \\ \widehat{\mathcal{C}}(\mathsf{q}) &= \begin{bmatrix} 1.1211 & -0.019113 & 1.7966 & 0.65666 \end{bmatrix}. \end{aligned}$$

Then, we check whether the converged ROM satisfies the newly developed optimality conditions. The relative errors in Lagrange interpolation of G in Theorem 4 for the two complex conjugate pairs are 8.496×10^{-9} and 2.105×10^{-8}. For interpolation of $\partial G / \partial s_a$, we have 1.6114×10^{-8} and 1.2166×10^{-7}. Finally, for interpolation of $\partial G / \partial s_b$, we have 4.9016×10^{-8} and 2.0029×10^{-7}. Thus, the resulting optimal ROM satisfies the interpolatory optimality conditions (to the accuracy of the gradient-based stopping criteria of the minimization algorithm).

6.2. Parametric Penzl's FOM

Next, we consider a parametric variant of Penzl's FOM from the SLICOT benchmark collection [18]. Specifically, we consider FOM of order 12 with $\mathcal{E}(\mathsf{q}) = I$,

$$\mathcal{A}(\mathsf{q}) = \operatorname{diag}(\mathcal{A}_1(\mathsf{q}), \mathcal{A}_2), \quad \mathcal{A}_1(\mathsf{q}) = \begin{bmatrix} -1 & \mathsf{q} \\ -\mathsf{q} & -1 \end{bmatrix}, \quad \mathcal{A}_2 = \operatorname{diag}(-1, -2, \ldots, -10),$$
$$\mathcal{B}(\mathsf{q})^{\mathrm{T}} = \mathcal{C}(\mathsf{q}) = \begin{bmatrix} 5 & 5 & 1 & 1 & \cdots & 1 \end{bmatrix},$$

and $\mathcal{P} = [1, 100]$.

Setting the initial ROM based on the FOM truncated to the first 3 states and running minimization to find a locally $\mathcal{H}_2 \otimes \mathcal{L}_2$ optimal D-StROM of the form $\widehat{\mathcal{E}}(\mathsf{q}) = I$,

$$\widehat{\mathcal{A}}(\mathsf{q}) = \begin{bmatrix} x_1 + \mathsf{q}x_2 & x_3 + \mathsf{q}x_4 & \\ -x_3 - \mathsf{q}x_4 & x_1 + \mathsf{q}x_2 & \\ & & x_5 + \mathsf{q}x_6 \end{bmatrix}, \quad \widehat{\mathcal{B}}(\mathsf{q}) = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \widehat{\mathcal{C}}(\mathsf{q}) = \begin{bmatrix} x_7 & x_8 & x_9 \end{bmatrix},$$

we obtain

$$\begin{aligned}
\widehat{\mathcal{A}}(\mathsf{q}) &= \operatorname{diag}\left(\widehat{\mathcal{A}}_1(\mathsf{q}), \widehat{\mathcal{A}}_2(\mathsf{q})\right),\\
\widehat{\mathcal{A}}_1(\mathsf{q}) &= \begin{bmatrix} -1.0030 + 7.2387 \times 10^{-6}\mathsf{q} & 2.2567 \times 10^{-3} + 1.0000\mathsf{q} \\ -2.2567 \times 10^{-3} - 1.0000\mathsf{q} & -1.0030 + 7.2387 \times 10^{-6}\mathsf{q} \end{bmatrix},\\
\widehat{\mathcal{A}}_2(\mathsf{q}) &= -3.5530 + 2.4940 \times 10^{-4}\mathsf{q},\\
\widehat{\mathcal{C}}(\mathsf{q}) &= \begin{bmatrix} 25.063 & -0.053279 & 8.7695 \end{bmatrix}.
\end{aligned}$$

Note that, just as in the previous example, the assumptions of Theorem 4 are satisfied.

The relative errors in the Lagrange and the two Hermite conditions for the first complex conjugate pair of poles are

$$1.0660 \times 10^{-10},\ 1.9085 \times 10^{-9},\ \text{and } 1.5356 \times 10^{-9}.$$

For the real pole they are

$$4.4460 \times 10^{-10},\ 4.4054 \times 10^{-10},\ \text{and } 1.6685 \times 10^{-9}.$$

Therefore, similarly to the previous example, we find good agreement with the theory.

7. Conclusions

We derived interpolatory necessary optimality conditions for $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal reduced-order modeling of parametric LTI systems with a general diagonal structure. Then we give conditions for special cases where only inputs and outputs or only the dynamics are parameterized. Future work includes the derivation of an iterative, IRKA-like algorithm to compute $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal StROMs.

References

[1] A. C. Antoulas, C. A. Beattie, and S. Gugercin, *Interpolatory methods for model reduction*, *Computational Science and Engineering 21*, SIAM, Philadelphia, PA, 2020.

[2] U. Baur, C. A. Beattie, P. Benner, and S. Gugercin, Interpolatory projection methods for parameterized model reduction, *SIAM J. Sci. Comput.* **33** (2011), no. 5, 2489–2518.

[3] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. M. Silveira (eds.), *Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms*, De Gruyter, Berlin, Boston, 2020.

[4] P. Benner, S. Gugercin, and K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems, *SIAM Rev.* **57** (2015), no. 4, 483–531.

[5] T. Bradde, S. Grivet-Talocia, A. Zanco, and G. C. Calafiore, Data-driven extraction of uniformly stable and passive parameterized macromodels, *IEEE Access* **10** (2022), 15786–15804.

[6] S. Brunsch, *Parametric Model Order Reduction Using Sparse Grids*, Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2017. Available at `https://hdl.handle.net/20.500.11811/7247`.

[7] A. R. Grimm, *Parametric Dynamical Systems: Transient Analysis and Data Driven Modeling*, Ph.D. thesis, Virginia Polytechnic Institute and State University, 2018. Available at `http://hdl.handle.net/10919/83840`.

[8] S. Gugercin, A. C. Antoulas, and C. Beattie, $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems, *SIAM J. Matrix Anal. Appl.* **30** (2008), no. 2, 609–638.

[9] M. Hund, T. Mitchell, P. Mlinarić, and J. Saak, Optimization-based parametric model order reduction via $\mathcal{H}_2 \otimes \mathcal{L}_2$ first-order necessary conditions, *SIAM J. Sci. Comput.* **44** (2022), no. 3, A1554–A1578.

[10] A. C. Ionita and A. C. Antoulas, Data-driven parametrized model reduction in the Loewner framework, *SIAM J. Sci. Comput.* **36** (2014), no. 3, A984–A1007.

[11] V. Kleyman, H. Gernandt, K. Worthmann, H. S. Abbas, R. Brinkmann, and M. A. Müller, Modeling and parameter identification for real-time temperature controlled retinal laser therapies, *at - Automatisierungstechnik* **68** (2020), no. 11, 953–966.

[12] L. Meier and D. Luenberger, Approximation of linear constant systems, *IEEE Trans. Autom. Control* **12** (1967), no. 5, 585–588.

[13] P. Mlinarić, Numerical examples for parametric $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimality conditions, Zenodo, 2024.

[14] P. Mlinarić, P. Benner, and S. Gugercin, Interpolatory $\mathcal{H}_2$-optimality conditions for structured linear time-invariant systems, *SIAM J. Numer. Anal.* **63** (2025), no. 2, 949–975.

[15] P. Mlinarić and S. Gugercin, $\mathcal{L}_2$-optimal reduced-order modeling using parameter-separable forms, *SIAM J. Sci. Comput.* **45** (2023), no. 2, A554–A578.

[16] P. Mlinarić and S. Gugercin, A unifying framework for interpolatory $\mathcal{L}_2$-optimal reduced-order modeling, *SIAM J. Numer. Anal.* **61** (2023), no. 5, 2133–2156.

[17] P. Mlinarić, S. Rave, and J. Saak, Parametric model order reduction using pyMOR, in *Model Reduction of Complex Dynamical Systems* (P. Benner, T. Breiten, H. Faßbender, M. Hinze, T. Stykel, and R. Zimmermann, eds.), *International Series of Numerical Mathematics* **171**, Birkhäuser, Cham, August 2021, pp. 357–367.

[18] Niconet e.V., `http://www.slicot.org`, *SLICOT - Subroutine Library in Systems and Control Theory*.

[19] D. Petersson, *A Nonlinear Optimization Approach to $\mathcal{H}_2$-Optimal Modeling and Control*, Ph.D. thesis, Linköping University, 2013. Available at `http://liu.diva-portal.org/smash/get/diva2:647068/FULLTEXT01.pdf`.

[20] A. C. Rodriguez, L. Balicki, and S. Gugercin, The p-AAA algorithm for data-driven modeling of parametric dynamical systems, *SIAM J. Sci. Comput.* **45** (2023), no. 3, A1332–A1358.

[21] The MORwiki Community, Synthetic parametric model, hosted at MORwiki – Model Order Reduction Wiki, 2005. Available at `http://modelreduction.org/index.php/Synthetic_parametric_model`.

[22] P. Virtanen, R. Gommers, T. E. Oliphant, and others, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17** (2020), 261–272.

[23] D. A. Wilson, Optimum solution of model-reduction problem, *Proc. Inst. Electr. Eng.* **117** (1970), no. 6, 1161–1165.