Parametric model reduction via rational interpolation along parameters

Ion Victor Gosea* Serkan Gugercin† Benjamin Unger‡

*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany.
Email: gosea@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3580-4116
†Department of Mathematics and Computational Modeling and Data Analytics Division, Academy of Integrated Science, Virginia Tech, Blacksburg, VA 24061, USA.
Email: gugercin@vt.edu, ORCID: 0000-0003-4564-5999
‡SC Simulation Technology, University of Stuttgart, Germany.
Email: benjamin.unger@simtech.uni-stuttgart.de, ORCID: 0000-0003-4272-1079

We present a novel projection-based model reduction framework for parametric linear time-invariant systems that allows interpolating the transfer function at a given frequency point along parameter-dependent curves as opposed to the standard approach where transfer function interpolation is achieved for a discrete set of parameter and frequency samples. We accomplish this goal by using parameter-dependent projection spaces. Our main result shows that for holomorphic system matrices, the corresponding interpolatory projection spaces are also holomorphic. The coefficients of the power series representation of the projection spaces can be computed iteratively using standard methods. We illustrate the analysis on three numerical examples.

Keywords: model reduction, parametric systems, rational interpolation, Sylvester equations, holomorphic functions.

1 Introduction

For a parameter vector $p \in \mathbb{P} \subseteq \mathbb{R}^{\nu}$, consider the parametric dynamical system in the state-space form

$$\Sigma(p)\colon \begin{cases} E(p)\dot{x}(t;p) = A(p)x(t;p) + B(p)u(t), \\ y(t;p) = C(p)x(t;p), \\ x(0;p) = 0, \end{cases} \tag{1}$$

with matrix functions $E, A\colon \mathbb{P} \to \mathbb{R}^{N\times N}$, $B\colon \mathbb{P} \to \mathbb{R}^{N\times m}$, and $C\colon \mathbb{P} \to \mathbb{R}^{\ell\times N}$. We assume that $E(p)$ is nonsingular for every $p \in \mathbb{P}$. In (1), we refer to x, u, and y as the *states*, *inputs*, and *outputs*, respectively. The parametric dynamical systems of the form (1) arise in many applications ranging from inverse problems to optimal control to uncertainty quantification and the parameter vector p enters the model in various ways, representing, for example, material properties, system geometry, and operating conditions; see, e.g., [7, 13, 20] and the references therein. Our standing assumption is that N is large and hence simulating (1) for a given input u and a given parameter p is expensive. Therefore, as required in many prominent applications, the need to repeat these simulations/computations for many parameter values and input selections leads to a big computational burden. This is what *parametric model reduction* (PMOR) aims to resolve. The goal of PMOR is to replace the *full-order model* (FOM) (1) by a *reduced-order model* (ROM) of the form

$$\widehat{\Sigma}(p)\colon \begin{cases} \widehat{E}(p)\dot{\widehat{x}}(t;p) = \widehat{A}(p)\widehat{x}(t;p) + \widehat{B}(p)u(t), \\ \widehat{y}(t;p) = \widehat{C}(p)\widehat{x}(t;p), \\ \widehat{x}(0;p) = 0, \end{cases} \tag{2}$$

with $\widehat{E}, \widehat{A}\colon \mathbb{P} \to \mathbb{R}^{n\times n}$, $\widehat{B}\colon \mathbb{P} \to \mathbb{R}^{n\times m}$, $\widehat{C}\colon \mathbb{P} \to \mathbb{R}^{n\times \ell}$, and $n \ll N$ such that the output $\widehat{y}(t;p)$ of the ROM approximates the output $y(t;p)$ of the FOM with high fidelity over a wide range of parameters and input selection. More precisely, we want the approximation error $\|y - \widehat{y}\|$ to be small for any $u \in \mathcal{L}_2(0,\infty,\mathbb{R}^m)$ and any parameter $p \in \mathbb{P}$.

1.1 Projection-based PMOR

There are plethora of methods to construct the ROM $\widehat{\Sigma}(p)$, we refer the reader to [2, 6, 7, 13, 20] for details. Common to most of these approaches is that they can be realized via a Petrov-Galerkin framework: Construct two MOR bases $V, W \in \mathbb{R}^{N\times n}$ such that $x(t,p) \approx V\widehat{x}(t,p)$. Then, substitute this approximation into (1) and enforce a Petrov-Galerkin condition on the residual to obtain the reduced-order matrices as

$$\begin{aligned} \widehat{E}(p) &:= W^\top E(p)V, \quad \widehat{A}(p) := W^\top A(p)V, \\ \widehat{B}(p) &:= W^\top B(p), \qquad \widehat{C}(p) := C(p)V. \end{aligned} \tag{3}$$

The task of model reduction is thus essentially equivalent to determining n-dimensional subspaces $\mathcal{V} := \operatorname{span}(V)$ and $\mathcal{W} := \operatorname{span}(W)$ of $\mathbb{R}^N$ such that the ROM (2) obtained via projection onto these spaces is a good approximation of (1). Even though it is not the focus of this paper, we note that there are data-driven approaches to PMOR in which $\widehat{\Sigma}(p)$ is constructed without access to the FOM dynamics in (1) and with only access to input-output data; see, e.g., [3, 9, 11, 12, 15], and the references therein.

For the linear parametric dynamical systems (1) and (2) we consider here, the concept of transfer function provides a natural framework to analyze the MOR problem. Let $Y(s,p)$ and $U(s,p)$ denote Laplace transforms of $y(t,p)$ and $u(t,p)$. Then, by taking the Laplace transform of (1), we obtain

$$Y(s,p) = H(s;p)U(s,p)$$

where

$$H(s;p) := C(p)\left(sA(p) - A(p)\right)^{-1} B(p) \tag{4}$$

is the transfer function of $\Sigma(p)$. Similarly, transfer function of the ROM $\widehat{\Sigma}(p)$ is given by

$$\widehat{H}(s;p) := \widehat{C}(p)\left(s\widehat{E}(p) - \widehat{A}(p)\right)^{-1} \widehat{B}(p). \tag{5}$$

In this paper, we will focus on interpolatory approaches to construct $\widehat{H}(s;p)$. Interpolatory MOR is one of the most commonly employed frameworks to MOR and yield (locally) optimal approximations in the $\mathcal{H}_2$-norm. We skip those details here and refer the reader to [2]. The interpolatory framework we develop here deviates from the usual approach in the literature as we explain next.

1.2 Interpolation problem to construct $\widehat{\Sigma}(p)$

The common approach to interpolatory PMOR chooses V and W so that $\widehat{H}(s,p)$ interpolates $H(s,p)$ at some *selected* right frequency samples $\{\lambda_i\}_{i=1}^{N_s}$, left frequency samples $\{\mu_i\}_{i=1}^{N_s}$, parameter samples $\{\pi_j\}_{j=1}^{N_p}$ along the right interpolation (tangent) directions $\{r_i\}_{i=1}^{N_s} \in \mathbb{C}^m$ and left interpolation (tangent) directions $\{\ell_i\}_{i=1}^{N_s} \in \mathbb{C}^\ell$; i.e.,

$$H(\lambda_i,\pi_j)r_i = \widehat{H}(\lambda_i,\pi_j)r_i \quad \text{and} \quad \ell_i^\top H(\mu_i,\pi_j) = \ell_i^\top \widehat{H}(\mu_i,\pi_j)$$

for $i = 1,2,\ldots,N_s$ and $j = 1,2,\ldots,N_p$. One can also enforce interpolating the derivatives of H with respect to s and p, and the discussion here directly extends. However, for brevity, we only focus on simple interpolation in this paper. We show in Theorem 2.1 how to construct V and W to satisfy the interpolation conditions listed above. These are *discretized* interpolation conditions in the sense that they hold over a discrete set of sampling points.

In this paper, we consider a more general problem of *interpolating $H(s,p)$ along parameter-dependent curves in the frequency domain.* More precisely, we are interested in solving the following problem.

Problem 1.1 *Consider the dynamical system (1) with transfer function $H(s;p)$. For given functions $\lambda\colon \mathbb{P} \to \mathbb{C}$, $\mu\colon \mathbb{P} \to \mathbb{C}$, $r\colon \mathbb{P} \to \mathbb{C}^m$, and $\ell\colon \mathbb{P} \to \mathbb{C}^\ell$, construct a ROM with transfer function $\widehat{H}(s;p)$ that tangentially interpolates H at λ along the right tangent directions r and at μ along the right tangent directions ℓ for all parameters, i.e., $\widehat{H}(s;p)$ satisfies*

$$H(\lambda(p);p)r(p) = \widehat{H}(\lambda(p);p)r(p), \quad \textit{and} \tag{6a}$$

$$\ell(p)^\top H(\lambda(p);p) = \ell(p)^\top \widehat{H}(\lambda(p);p), \quad \textit{for all } p \in \mathbb{P}. \tag{6b}$$

In general, we cannot expect to find constant matrices $V, W \in \mathbb{R}^{N\times n}$ with small n such that (6) is satisfied for all parameters $p \in \mathbb{P}$. Instead, motivated by the lower-bound for the Kolmogorov n-widths [23, Thm. 3], we propose to construct parameter dependent model reduction bases $\mathcal{V}(p)$ and $\mathcal{W}(p)$, exemplified by the matrix functions

$$V, W\colon \mathbb{P} \to \mathbb{R}^{N\times n}.$$

Our analysis is inspired by the ideas presented in [24], which studied the balanced truncation method for parametric system.

Once the parameter dependent bases are chosen, the ROM is constructed via projection onto the spaces given by $\mathcal{V}(p) := \operatorname{span}(V(p))$ and $\mathcal{W}(p) := \operatorname{span}(W(p))$, i.e.,

$$\begin{aligned} \widehat{E}(p) &:= W(p)^\top E(p)V(p), \quad \widehat{A}(p) := W(p)^\top A(p)V(p), \\ \widehat{B}(p) &:= W(p)^\top B(p), \qquad \widehat{C}(p) := C(p)V(p). \end{aligned} \tag{7}$$

Remark 1.1 *Time- and state-dependent projection matrices are currently heavily investigated in the efficient approx-*

imation of transport-dominated phenomena, where a nonlinear projection framework is used to overcome slowly decaying Kolmogorov n-widths, see, e.g., [8, 18, 21] and the references therein.

After this introduction, we recall some preliminary results in Section 2. Our main contribution is presented in Section 3 with additional computational details presented in Section 4.

Notation Besides standard notation, we use multi-indices, i.e., for $j = (j_1, \ldots, j_\nu) \in \mathbb{N}_0^\nu$ and $p = (p_1, \ldots, p_\nu)$ we write

$$p^j := \prod_{i=1}^{\nu} p_i^{j_i}.$$

2 Preliminaries

2.1 Interpolation conditions

Interpolatory model reduction [2] constructs reduced-order models whose transfer function interpolates the transfer function of the original model at selected interpolation points. For a fixed parameter $\pi \in \mathbb{P}$, interpolation via projection can be guaranteed as follows [2, 4].

Theorem 2.1 (Tangential interpolation) *For a fixed parameter $\pi \in \mathbb{P}$ consider the FOM (1) with transfer function $H(s;\pi)$ and the ROM (2) with transfer function $\widehat{H}(s;\pi)$ constructed as in (3) using $W, V \in \mathbb{R}^{n\times r}$. For interpolation points $\lambda_0, \mu_0 \in \mathbb{C}$, assume that $\lambda_0 E(\pi) - A(\pi)$ and $\mu_0 E(\pi) - A(\pi)$ are nonsingular. Let $r_0 \in \mathbb{R}^m$ and $\ell_0 \in \mathbb{R}^m$.*

1. *If $(\lambda_0 E(\pi) - A(\pi))^{-1} B(\pi) r_0 \in \operatorname{span}(V)$, then $H(\lambda_0;\pi) r_0 = \widehat{H}(\lambda_0;\pi) r_0$.*
2. *If $(\ell_0^\top C(\pi)(\mu_0 E(\pi) - A(\pi))^{-1})^\top \in \operatorname{span}(W)$, then $\ell_0^\top H(\mu_0;\pi) = \ell_0^\top \widehat{H}(\mu_0;\pi)$.*

It is easy to see (cf. [10]) that matrices satisfying the conditions in Theorem 2.1 for driving frequencies λ_i, μ_i and tangent directions r_i, ℓ_i $(i = 1, \ldots, n)$ can be constructed by solving the two Sylvester equations

$$A(\pi)V - E(\pi)V\Lambda + B(\pi)R = 0, \tag{8a}$$

$$W^\top A(\pi) - M^\top W^\top E(\pi) + L^\top C(\pi) = 0, \tag{8b}$$

for the unknowns V and W where

$$\Lambda := \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \qquad M := \operatorname{diag}(\mu_1, \ldots, \mu_n), \tag{9a}$$

$$R := \begin{bmatrix} r_1 & \ldots & r_n \end{bmatrix}, \qquad L := \begin{bmatrix} \ell_1 & \ldots & \ell_n \end{bmatrix}. \tag{9b}$$

If the driving frequencies and tangent directions are closed under complex conjugation, then one can use real versions of the matrices in (9). For fixed $\pi \in \mathbb{P}$, the condition guaranteeing the existence and uniqueness of solutions to (8) is well-known, see, e.g., [1, Cha. 6].

Lemma 2.1 *For $\pi \in \mathbb{P}$, the Sylvester equations (8) have a unique solution if and only if $\lambda_i, \mu_i \notin \sigma(E(\pi), A(\pi))$, where*

$$\sigma(E(\pi), A(\pi)) := \{s \in \mathbb{C} \mid \operatorname{rank}(sE(\pi) - A(\pi)) < N\}.$$

is the spectrum of the matrix pencil $sE(\pi) - A(\pi)$.

2.2 Holomorphic functions

Our analysis requires that the matrix functions in (1) can be expanded in a power series. If the parameter domain is one-dimensional, this is then equivalent to the matrix functions being holomorphic (resp. analytic). Since we do not intend to restrict our analysis to a single parameter, we recall the appropriate definitions and results for functions of several parameters. For our presentation we follow [16] and [24].

A function $f\colon \mathbb{C}^\nu \supseteq \mathbb{P} \to \mathbb{C}$ is called holomorphic in $p = [p_j] \in \mathbb{P}$ if the complex derivative

$$f'(p) = \lim_{h\to 0} \frac{f(p + hq) - f(p)}{h}$$

exists for any $q \in \mathbb{C}^\nu$. It is said to be holomorphic in $\mathbb{P}$, if it is holomorphic in every $p \in \mathbb{P}$. Many of the results for the one-dimensional case extend to a higher dimensional domain, such as the Cauchy integral formula. In particular, if f is holomorphic, it can locally be represented via a power series. For the analysis of its domain of convergence, we need the following definition, taken from [16].

Definition 2.2 (Reinhardt domain) *An open set $\Omega \subseteq \mathbb{C}^\nu$ is called* Reinhardt domain, *if $p = (p_1, \ldots, p_\nu) \in \Omega$ implies $(\exp(\imath\theta_1)p_1, \ldots, \exp(\imath\theta_\nu)p_\nu) \in \Omega$ for all $(\theta_1, \ldots, \theta_\nu) \in \mathbb{R}^\nu$, where $\imath$ denotes the imaginary unit.*

Theorem 2.3 *Let $\mathbb{P} \subseteq \mathbb{C}^\nu$ be a connected Reinhardt domain containing 0 and suppose that $f\colon \mathbb{P} \to \mathbb{C}$ is holomorphic in $\mathbb{P}$. Then there exist unique $f_i \in \mathbb{C}$ for $i \in \mathbb{N}_0^\nu$ such that*

$$f(p) = \sum_{i\in\mathbb{N}_0^\nu} f_i p^i \qquad \text{for each } p \in \mathbb{P}. \tag{10}$$

Note that for simplicity, we have presented Theorem 2.3 solely for the expansion point $\bar{p} = 0$. For practical applications, we may want to use a different expansion point or rescale the parameter domain and the system matrices such that 0 is included in $\mathbb{P}$.

A question that arises immediately is whether there is an holomorphic version of the implicit mapping theorem available. This is indeed the case. For our analysis, we use the following extension of the implicit mapping theorem [24].

Proposition 2.1 *Consider a function $\mathcal{F}\colon \mathbb{C}^{\nu} \times \mathbb{C}^{n_1 \times n_2} \to \mathbb{C}^{n_1 \times n_2}$ and suppose there exists $p_0 \in \mathbb{C}^{\mathbb{P}}$ and $X_0 \in \mathbb{C}^{n_1\times n_2}$ such that $\mathcal{F}(p_0, X_0) = 0$ and $\mathcal{F}$ is holomorphic around this point. If*

$$0 = \frac{\partial}{\partial h}\mathcal{F}(p_0, X_0 + hD)\big|_{h=0}$$

implies $D = 0$, then there exists an neighborhood $\mathcal{P} \subset \mathbb{C}^{\nu}$ around p_0 and a holomorphic function $X\colon \mathcal{P} \to \mathbb{C}^{n_1\times n_2}$ such that

$$\mathcal{F}(p, X(p)) = 0$$

for all $p \in \mathbb{P}$.

3 Rational interpolation along parameter-dependent curves

In this section, we establish the main result that guarantees existence of holomorphic functions $V(p)$ and $W(p)$ such that the reduced model in (7) solves the new interpolation problem defined in Problem 1.1.

Theorem 3.1 *Consider the dynamical system (1) and assume that E, A, B, C are holomorphic in the compact set $\mathbb{P} \subseteq \mathbb{C}^{\nu}$. Assume that for $i = 1, \ldots, n$ the holomorphic functions $\lambda_i, \mu_i\colon \mathbb{P} \to \mathbb{C}$ are such that*

$$\lambda_i(p), \mu_i(p) \notin \sigma(E(p), A(p))$$

for all $p \in \mathbb{P}$. If the tangent directions $r_i\colon \mathbb{P} \to \mathbb{C}^m$ and $\ell_i\colon \mathbb{P} \to \mathbb{C}^{\ell}$ are holomorphic, then there exists holomorphic functions $V, W\colon \mathbb{P} \to \mathbb{C}^{N\times n}$ satisfying

$$A(p)V(p) - E(p)V(p)\Lambda(p) + B(p)R(p) = 0, \quad (11)$$

$$W^{\top}(p)A(p) - M^{\top}(p)W^{\top}(p)E(p) + L^{\top}(p)C(p) = 0 \quad (12)$$

for all $p \in \mathbb{P}$, where $\Lambda(p)$, $M(p)$, $R(p)$, $L(p)$ are defined as in (9), but now with parametric dependence.

Proof: We show the assertion only for V. The proof for W follows similarly. Define the holomorphic function

$$\mathcal{F}\colon \mathbb{P} \times \mathbb{C}^{N\times n} \to \mathbb{C}^{N\times n},$$
$$(p, V) \mapsto A(p)V - E(p)V\Lambda(p) + B(p)R(p).$$

Let $p_0 \in \mathbb{P}$. Then, using Lemma 2.1, there exists $V_0 \in \mathbb{C}^{N\times n}$ satisfying the condition $\mathcal{F}(p_0, V_0) = 0$. In addition, for any $\widetilde{V} \in \mathbb{C}^{N\times n}$ we obtain

$$\frac{\partial}{\partial \varepsilon}\mathcal{F}(p, V_0 + \varepsilon\widetilde{V}) = A(p)\widetilde{V} - E(p)\widetilde{V}\Lambda(p).$$

From Lemma 2.1 we conclude that $\frac{\partial}{\partial\varepsilon}\mathcal{F}(p, V_0 + \varepsilon\widetilde{V}) = 0$ if and only if $\widetilde{V} = 0$. Thus, Proposition 2.1 implies that there exists a neighborhood $\mathcal{P} \subseteq \mathbb{C}^{\nu}$ around p_0 and a holomorphic function $V\colon \mathcal{P} \cap \mathbb{P} \to \mathbb{C}^{N\times n}$ satisfying $\mathcal{F}(p, V(p)) = 0$. Let $\mathcal{P}$ denote the maximal neighborhood such that the previous construction holds. It remains to show that $\mathcal{P} \cap \mathbb{P} = \mathbb{P}$. Assume $\mathcal{P} \cap \mathbb{P} \neq \mathbb{P}$ and let $\pi \in \mathbb{P} \setminus \mathcal{P}$. Repeating the construction, we obtain a neighborhood $\tilde{\mathcal{P}} \subseteq \mathbb{C}^{\nu}$ and holomorphic function $\tilde{V}\colon \tilde{\mathcal{P}} \cap \mathbb{P} \to \mathbb{C}^{N\times n}$ satisfying $\mathcal{F}(p, \tilde{V}(p)) = 0$. Assume first $\mathcal{P} \cap \tilde{\mathcal{P}} \neq \varnothing$. Then there exists $\tilde{\pi} \in \mathcal{P} \cap \tilde{\mathcal{P}}$. Due to Lemma 2.1 and the assumptions we conclude $V(\tilde{\pi}) = \tilde{V}(\tilde{\pi})$. From the holomorphic identity theorem [17, Thm. 1.2.14] we infer $V = \tilde{V}$, a contradiction. If, on the other hand, $\mathcal{P} \cap \tilde{\mathcal{P}} = \varnothing$, we can select further points in $\mathbb{P}$ until we obtain an open covering of $\mathbb{P}$. Since $\mathbb{P}$ is compact, we can choose a finite covering and proceed as before. We conclude $\mathcal{P} \cap \mathbb{P} = \mathbb{P}$.

Corollary 3.2 *Suppose that the assumptions enforced in Theorem 3.1 are satisfied and construct a ROM as in (7). Then the ROM satisfies the interpolation conditions (6) for all $p \in \mathbb{P}$, thus solving Problem 1.1.*

Remark 3.3 *Using [22, Prop. 3.24], Theorem 3.1 can be extended to structured systems with a transfer function of the form $H(s;p) = C(p)(\sum_{k=1}^{K} h_k(s;p)A_k(p))^{-1}B(p)$, which includes, for instance, delay equations, fractional systems, and viscoelastic dynamics.*

4 Computational details

Even though we have established the theoretical framework for constructing $V(p)$ and $W(p)$ to solve the new parametric interpolation problem, for a numerically efficient PMOR framework we need to consider the computational aspects in solving (11) and (12), and performing the projection (7).

For the brevity of presentation we restrict ourselves in this section to standard state-space systems with $E(p) \equiv I_N$. For a parameter-dependent E matrix, the construction is similar, but the formulas are more involved.

4.1 Numerical construction of $V(p)$ and $W(p)$

Assuming that holomorphic matrix functions, Theorem 2.3 ensures that we can decompose these matrices as

$$A(p) = \sum_{i\in\mathbb{N}_0^{\nu}} p^i A_i, \quad \Lambda(p) = \sum_{i\in\mathbb{N}_0^{\nu}} p^i \Lambda_i, \quad V(p) = \sum_{i\in\mathbb{N}_0^{\nu}} p^i V_i$$
$$B(p) = \sum_{i\in\mathbb{N}_0^{\nu}} p^i B_i, \quad \text{and} \quad R(p) = \sum_{i\in\mathbb{N}_0^{\nu}} p^i R_i.$$

In many practical applications, the system matrices are directly available in such a form with a finite number of terms. Then the Sylvester equation (11) becomes

$$\begin{aligned}0 &= \sum_{j\in\mathbb{N}_0^\nu}\sum_{i\in\mathbb{N}_0^\nu}(A_iV_j - V_j\Lambda_i + B_iR_j)p^{i+j}\\ &= \sum_{\rho\in\mathbb{N}_0^\nu}\sum_{\substack{i+j=\rho\\ i,j\in\mathbb{N}_0^\nu}}(A_iV_j - V_j\Lambda_i + B_iR_j)p^{\rho}.\end{aligned}$$

Using the holomorphic identity theorem [17, Thm. 1.2.14], we conclude that for $\rho\in\mathbb{N}_0^\nu$ we have

$$\begin{aligned}0 &= \sum_{\substack{i+j=\rho\\ i,j\in\mathbb{N}_0^\nu}}(A_iV_j - V_j\Lambda_i + B_iR_j)\\ &= A_0V_\rho - V_\rho\Lambda_0 + \sum_{\substack{i+j=\rho\\ j\neq\rho}}A_iV_j - V_j\Lambda_i + \sum_{i+j=\rho}B_iR_j,\end{aligned}\tag{13}$$

which provides an iterative method to solve for the coefficients V_i. A similar strategy can be obtained for the coefficients for W, which we omit here to avoid redundancy.

Corollary 4.1 *Under the assumptions of Theorem 3.1 the Sylvester equation (13) is uniquely solvable for each $\rho\in\mathbb{N}_0^\nu$.*

Proof: This follows immediately from $A_0 = A(0)$, $\Lambda_0 = \Lambda(0)$, and Lemma 2.1.

Note that if the coefficients of A, Λ, B, and R are real (i.e., the interpolation frequencies and tangent directions are closed under conjugation), then the V_j are real thus yielding a real-valued matrix $V(p)$ for each real parameter $p\in\mathbb{P}$.

In numerical computations, we cannot compute all the coefficients V_i and thus have to truncate the power-series expansion at an index based on a tolerance. In other words, for a given tolerance τ, we truncate the power series expansion when $\max_{p\in\mathbb{P}}|p^i|\|V_i\|\leq\tau$, and similarly for $W(p)$. As a consequence, we cannot ensure exact interpolation any longer. A similar issue arises in the usual interpolatory model reduction framework when the required subspace vectors in Theorem 2.1, namely

$$(\lambda_0E(\pi)-A(\pi))^{-1}B(\pi)r_0,\ \text{and}\ (\ell_0^\top C(\pi)(\mu_0E(\pi)-A(\pi))^{-1})^\top,$$

are computed via iterative solves; see, e.g., [5]. We revisit this issue in Section 6.

4.2 Constructing the reduced matrices

For simplicity, we only focus on $\widehat{A}(p)$ in (7); but the discussion extends directly to other reduced order quantities.

We will work with the truncated quantities, i.e.,

$$\begin{aligned}A(p) &= \sum_{\|k\|\leq\rho_a}p^kA_k,\\ W(p) = \sum_{\|j\|\leq\rho_w}p^iW_i,&\qquad V(p) = \sum_{\|i\|\leq\rho_v}p^iV_i.\end{aligned}\tag{14}$$

For every new parameter vector $\pi\in\mathbb{P}$, forming $V(\pi)$ (and $W(\pi)$) can be efficiently done using the truncated form as in (14). However constructing $\widehat{A}(\pi)$ requires computing $\widehat{A}(\pi) = W(\pi)^\top A(\pi)V(\pi)$, which involves two matrix multiplications in the original dimension N. We resolve this issue using the truncated forms (14):

$$\widehat{A}(\pi) = \sum_{\|j\|\leq\rho_w}\sum_{\|k\|\leq\rho_a}\sum_{\|i\|\leq\rho_v}\left(W_j^\top A_kV_i\right)\pi^{i+j+k}.\tag{15}$$

Note that the reduced coefficients $W_j^\top A_kV_i\in\mathbb{R}^n$ in (15) can be precomputed (in the offline stage). Assuming ρ_a,ρ_v and ρ_w are modest integers, storing all the coefficients $W_j^\top A_kV_i$ and then forming the overall sum can be efficiently computed in the online stage.

5 Numerical examples

We illustrate the theoretical analysis on three models.

5.1 A toy example

Consider a simple example for which the dimension of the parameter set is $\nu = 1$ (and the parameter enters only in the vector B). The matrices are as follows

$$\begin{aligned}&A(p)\equiv-\mathrm{diag}(1,1,2), && C(p)\equiv\begin{bmatrix}2 & 1 & 1\end{bmatrix},\\ &B(p) = \begin{bmatrix}p & 1-p & 1\end{bmatrix}^\top, && E(p)\equiv I_3.\end{aligned}\tag{16}$$

Hence, it follows that

$$\begin{aligned}&A_0 = -\mathrm{diag}(1,1,2),\ C_1 = \begin{bmatrix}1 & 1 & 2\end{bmatrix},\ A_i, C_i = 0\ \forall i\geq 1\\ &B_0 = \begin{bmatrix}0 & 1 & 1\end{bmatrix}^\top,\ B_1 = \begin{bmatrix}1 & -1 & 1\end{bmatrix}^\top,\ B_i = 0\ \forall i\geq 2.\end{aligned}$$

Originally, note that $N = 3$ and that we choose $n = 2$ as the reduction order. Choose interpolation points and tangent directions that are independent of the parameter, e.g.,

$$\lambda_1 = 1,\ \lambda_2 = 3,\ \mu_1 = 2,\ \mu_2 = 4,\tag{17}$$

and also $r = \ell = [1; 1]$. Choose the following:

$$\begin{aligned}\Lambda_0 = \begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix},\ M_0 = \begin{bmatrix}2 & 0\\ 0 & 4\end{bmatrix},\ \Lambda_i = 0,\ M_i = 0\ \forall i\geq 1\\ R_0 = L_0^\top = \begin{bmatrix}1 & 1\end{bmatrix},\ R_i = L_i^\top = 0, \forall i\geq 1.\end{aligned}$$

For $\rho = 0$, it follows that the equation (13) simplifies to $A_0V_0 - V_0\Lambda_0 + B_0R_0 = 0$. Similarly, based also on (13), V_1 satisfies the following Sylvester equation

$$A_0V_1 - V_1\Lambda_0 + A_1V_0 - V_0\Lambda_1 + B_1R_0 + B_0R_1 = 0,\tag{18}$$

which simplifies to $A_0V_1 - V_1\Lambda_0 + B_1R_0 = 0$. Hence, explicitly compute the first two Taylor coefficients

$$V_0 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{5} \end{bmatrix}, \; V_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{4} \\ 0 & 0 \end{bmatrix}, \tag{19}$$

and $V_i = 0, \forall i \geq 2$. Next, compute matrix $W^\top = W_0^\top \in \mathbb{C}^{2\times 3}$ by solving $W_0^T A_0 - M_0 W_0^T + L_0 C_0 = 0$, as

$$W^\top = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{4} \\ \frac{2}{5} & \frac{1}{5} & \frac{1}{6} \end{bmatrix}, \tag{20}$$

and put together the following reduced realization that does indeed depend on the parameter p as follows

$$\begin{aligned} \hat{E}^{(1)}(p) &= W^\top E V(p) = \begin{bmatrix} \frac{p}{6} + \frac{1}{4} & \frac{p}{12} + \frac{2}{15} \\ \frac{p}{10} + \frac{7}{45} & \frac{p}{20} + \frac{1}{12} \end{bmatrix}, \\ \hat{A}^{(1)}(p) &= W^\top A(p) V(p) = \begin{bmatrix} -\frac{p}{6} - \frac{1}{3} & -\frac{p}{12} - \frac{11}{60} \\ -\frac{p}{10} - \frac{19}{90} & -\frac{p}{20} - \frac{7}{60} \end{bmatrix}, \\ \hat{B}^{(1)}(p) &= W^\top B(p) = \begin{bmatrix} \frac{p}{3} + \frac{7}{12} \\ \frac{p}{5} + \frac{11}{30} \end{bmatrix}, \\ \hat{C}^{(1)}(p) &= C(p)V(p) = \begin{bmatrix} \frac{p}{2} + \frac{5}{6} & \frac{p}{4} + \frac{9}{20} \end{bmatrix}. \end{aligned} \tag{21}$$

We note that the system in (21) interpolates the original one in (16) at the selected frequencies for *every* value of the parameter p. We also note that the system in (21) is equivalent to a minimal realization of (16) for $p \in \{0, 1\}$.

5.2 Another toy example

Consider the following example:

$$\begin{aligned} A(p) &= \begin{bmatrix} -2 & p & 0 \\ -p & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad C(p) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \\ B(p) &= \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top, \quad E(p) = I_3. \end{aligned} \tag{22}$$

Hence, it follows that:

$$A_0 = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \; A_1 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and $A_i = 0$ for all $i \geq 2$. Then, we have also that

$$B_0^\top = C_0 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}, \text{ and } B_i^\top = C_i = 0, \forall i \geq 1.$$

For this case, consider two right interpolation points as:

$$\Lambda(p) = \begin{bmatrix} 0.1 & 0 \\ 0 & 5 \end{bmatrix} = \Lambda_0, \text{ and } \Lambda_i = 0, \; \forall i \geq 1.$$

Note that in this case we use $W^\top(p) = V^\top(p)$. The right directions are all ones and the Sylvester equations in (13) simplify to the following collection:

$$\begin{aligned} A_0V_0 - V_0\Lambda_0 + B_0R_0 &= 0, \\ A_0V_i - V_i\Lambda_0 + A_1V_{i-1} &= 0, \; \forall i \geq 1. \end{aligned} \tag{23}$$

Hence, one can iteratively compute V_i for any positive value of i. We do that for all values of i until $\|V_i\| < \tau$, for a tolerance value of $\tau = 10^{-5}$. This corresponds to a number of 26 Taylor coefficients that need to be computed. Finally, as described in Section 4.2, we put together the reduced-order matrices and evaluate the approximation errors for a 2D grid consisting in values $p \in [0, 1]$, and $s \in [10^{-2}, 10^1]$. The results are presented in Figure 1.

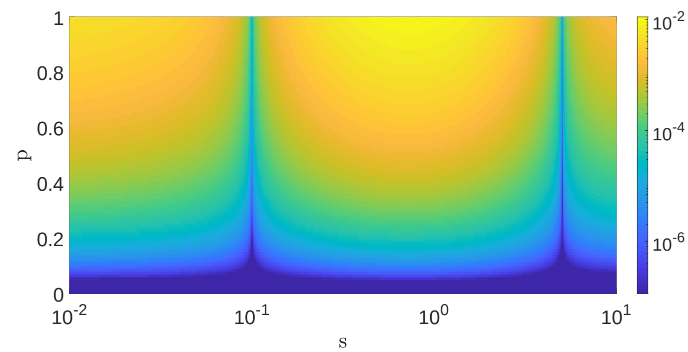

Figure 1: Approximation errors on a 2D grid (s, p).

5.3 A more involved numerical example

We analyze the dynamical system originally proposed in [19] and later modified in [9,15] to add a parameter dependence. The dynamics are characterized by the following equations:

$$\Sigma(p) \colon \begin{cases} \dot{x}(t;p) = A(p)x(t;p) + B(p)u(t), \\ y(t;p) = C(p)x(t;p), \\ x(0;p) = 0 \end{cases} \tag{24}$$

where $p \in \mathbb{P} = [0, 1]$ and $A \colon \mathbb{P} \to \mathbb{R}^{1006\times 1006}$

$$\begin{aligned} &A(p) = \operatorname{diag}\left(T_1(p), T_2, T_3, T_4\right), \text{ with} \\ &T_1(p) = \begin{bmatrix} -1 & p+100 \\ -100-p & -1 \end{bmatrix}, T_2 = \begin{bmatrix} -1 & 200 \\ -200 & -1 \end{bmatrix}, \\ &T_3 = \begin{bmatrix} -1 & 400 \\ -400 & -1 \end{bmatrix}, \quad T_4 = -\operatorname{diag}\left(1, 2, \ldots, 1000\right). \end{aligned}$$

Additionally, the constant vectors B and C are given by

$$B = C^T = [10e_6; e_{1000}],$$

where e_k denotes the k-dimensional vector of ones.

Next, we choose 40 logarithmically-spaced interpolation points $\lambda_1, \ldots, \lambda_{40}$ in the interval $[10^{-1}, 10^3]\imath$ (we are using a one-sided interpolation scheme). Additionally, let the tolerance value be $\tau = 10^{-7}$. It follows that we need to compute the first 11 Taylor coefficients of $V(p)$, i.e., $V_1, V_2, \ldots, V_{10}$, since $V_{11} < \tau$. As in the previous example, use $W(p) = V\top(p)$ as left projection matrix, and follow the formulas presented in Section 4.2, to compute the corresponding reduced-order matrices.

First we fix the frequency parameter as $s = \lambda_{20} = 8.8862\imath$ and vary p in between 0 and 1 (50 linearly-spaced points). We depict the approximation errors for different values of p in Figure 2. We note that the interpolation errors due to the truncation of the power series are small, in the interval $(10^{-9}, 10^{-6})$, in accordance with the tolerance $\tau = 10^{-7}$.

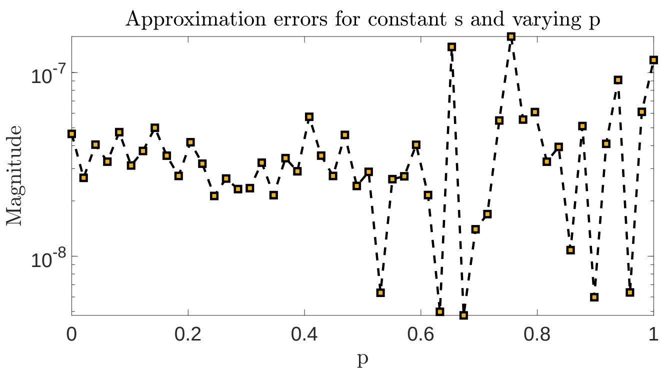

Figure 2: Approximation errors for $s = \lambda_{20}$ and varying p in $[0, 1]$.

For the next experiment, we fix the p parameter, i.e., choose $p = 0.5$ and vary the frequency parameter s in the interval $[10^{-1}, 10^3]\imath$ (200 logarithmically-spaced points). We depict the magnitudes of the two transfer functions (original and reduced) evaluated for different values of s in Figure 3, illustrating that FOM response is indeed well matched.

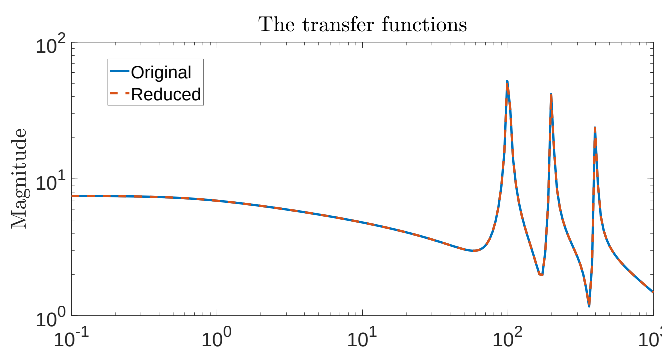

Figure 3: The two transfer functions for $p = 0.5$ and s in $[10^{-1}, 10^3]\imath$.

Finally, we construct a 2D grid consisting in pairs of parameters (s, p) evaluated on the Cartesian product of the two previously-mentioned discrete sets. Then, for all the $200 \times 50 = 10^4$ pairs, we compute the approximation error. The results are presented in Figure 4.

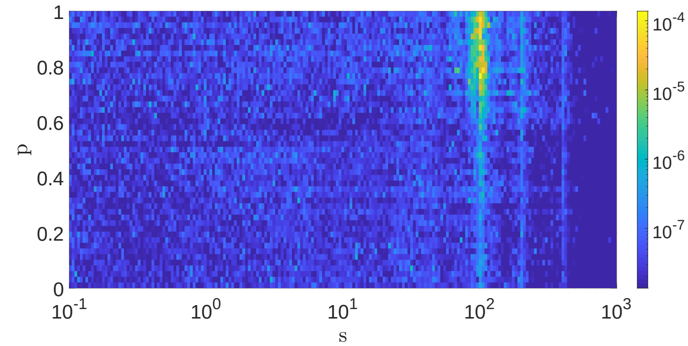

Figure 4: Approximation errors on a 2D grid (s, p).

6 Conclusions and future work

We have presented a theoretical framework that allows to construct a ROM whose transfer function interpolates the transfer function of the original high-dimensional system at parameter-dependent interpolation frequencies along some parameter-dependent directions. The associated parametric projection spaces are proven to have a holomorphic dependency on the parameter and the coefficients of its power series can be computed iteratively using standard methods.

There are many natural avenues to investigate further. For example, interpolation of the higher-order derivatives is a natural next step. In this paper, we did not consider an optimality measure for choosing the projection spaces. One might consider combining our framework with the recent work on optimal parametric model reduction in a joint $\mathcal{H}_2 \otimes \mathcal{L}_2$ measure [14]. Even though we have considered here the projection-based approaches, data-driven methods have been also considered for parametric systems [15]. Interpreting our reduced model in that framework could provide further hints for data-driven modeling.

As we stated in Section 4.1, when the power series expansions are truncated, we can no longer guarantee exact interpolation. We will investigate in a future work how the perturbation results from interpolatory model reduction with inexact solves [5] can be used to quantify the interpolation error due to the truncation.

Acknowledgments

The work of S. Gugercin was supported in parts by National Science Foundation under Grant No. DMS-1923221 and DMS-1819110. The work of B. Unger is funded by the German Research Foundation (DFG) under Germany's Excellence Strategy - EXC 2075 – 390740016. In addition, B. Unger acknowledges support by the Stuttgart Center for Simulation Science (SimTech).

References

[1] A. C. Antoulas. *Approximation of large-scale dynamical systems*. Advances in Design and Control. SIAM, Philadelphia, 2005.

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

[3] A. C. Antoulas, A. C. Ionita, and S. Lefteriu. On two-variable rational interpolation. *Linear Algebra Appl.*, 436(8):28890–2915, apr 2012.

[4] U. Baur, P. Benner, C.A. Beattie, and S. Gugercin. Interpolatory projection methods for parameterized model reduction. *SIAM Journal on Scientific Computing*, 33:2489–2518, 2011.

[5] C.A. Beattie, S. Gugercin, and S. Wyatt. Inexact solves in interpolatory model reduction. *Linear Algebra Appl.*, 436(8):2916–2943, 2012.

[6] P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, editors. *Model Reduction and Approximation: Theory and Algorithms*. Computational Science & Engineering. SIAM Publications, Philadelphia, PA, 2017.

[7] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.

[8] F. Black, P. Schulze, and B. Unger. Projection-based model reduction with dynamically transformed modes. *ESAIM: Math. Model. Numer. Anal.*, 54(6):2011–2043, 2020.

[9] A. Carracedo Rodriguez and S. Gugercin. The p-AAA algorithm for data driven modeling of parametric dynamical systems. *ArXiv e-print 2003.06536*, 2020.

[10] K. Gallivan, A. Vandendorpe, and P. Van Dooren. Model reduction of MIMO systems via tangential interpolation. *SIAM J. Matrix Anal. Appl.*, 26(2):328–349, 2004.

[11] A. R. Grimm. *Parametric Dynamical Systems: Transient Analysis and Data Driven Modeling*. PhD thesis, Virginia Tech, 2018.

[12] S. Grivet-Talocia and E. Fevola. Compact parameterized black-box modeling via fourier-rational approximations. *IEEE Transactions on Electromagnetic Compatibility*, 59(4):1133–1142, 2017.

[13] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. Springer Briefs in Mathematics. Springer, Switzerland, 2016.

[14] M. Hund, T. Mitchell, P. Mlinaric, and J. Saak. Optimization-based parametric model order reduction via $\mathcal{H}_2 \otimes \mathcal{L}_2$ first-order necessary conditions. *ArXiv e-print 2103.03136*, 2021.

[15] A. C. Ionita and A.C. Antoulas. Data-driven parametrized model reduction in the Loewner framework. *SIAM J. Sci. Comput.*, 36(3):A984–A1007, 2014.

[16] B. Malgrange. *Lectures on The Theory of Functions of Several Complex Variables*, volume 1958. Springer-Verlag Berlin Heidelberg, 1984.

[17] J. Noguchi. *Analytic Function Theory of Several Variables*. Springer, Singapore, 2016.

[18] M. Ohlberger and S. Rave. Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing. *C. R. Math. Acad. Sci. Paris*, 351(23–24):901–906, 2013.

[19] T. Penzl. Algorithms for model reduction of large dynamical systems. *Linear Algebra Appl.*, 415(2–3):322–343, 2006.

[20] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: an introduction*. UNITEXT. Springer Cham, 2016.

[21] D. Rim, B. Peherstorfer, and K.T. Mandli. Manifold approximations via transported subspaces: Model reduction for transport-dominated problems. *ArXiv e-print 1912.13024*, 2019.

[22] P. Schulze, B. Unger, C. Beattie, and S. Gugercin. Data-driven structured realization. *Linear Algebra Appl.*, 537:250–286, 2018.

[23] B. Unger and S. Gugercin. Kolmogorov n-widths for linear dynamical systems. *Adv. Comput. Math.*, 2019.

[24] P. Wittmuess, C. Tarin, A. Keck, E. Arnold, and O. Sawodny. Parametric Model Order Reduction via Balanced Truncation with Taylor Series Representation. *IEEE Trans. Automat. Contr.*, 61(11):3438–3451, 2016.