Interpolatory Model Reduction of Large-Scale Dynamical Systems

Athanasios C. Antoulas, Christopher A. Beattie, and Serkan Gugercin

1 Introduction

Large scale dynamical systems are a common framework for the modeling and control of many complex phenomena of scientific interest and industrial value, with examples of diverse origin that include signal propagation and interference in electric circuits, storm surge prediction before an advancing hurricane, vibration suppression in large structures, temperature control in various media, neurotransmission in the nervous system, and behavior of micro-electro-mechanical systems. Direct numerical simulation of underlying mathematical models is one of few available means for accurate prediction and control of these complex phenomena. The need for ever greater accuracy compels inclusion of greater detail in the model and potential coupling to other complex systems leading inevitably to very large-scale and complex dynamical models. Simulations in such large-scale settings can make untenable demands on computational resources and efficient model utilization becomes necessary. *Model reduction* is one response to this challenge, wherein one seeks a simpler (typically lower order) model that nearly replicates the behavior of the original model. When high fidelity is achieved with a reduced-order model, it can then be used reliably as an efficient surrogate to the original, perhaps replacing it as a component in larger simulations or in allied contexts such as development of simpler, faster controllers suitable for real time applications.

A.C. Antoulas
Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005-1892, USA
e-mail: aca@rice.edu

C.A. Beattie
Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA
e-mail: beattie@vt.edu

S. Gugercin
Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA
e-mail: gugercin@math.vt.edu

J. Mohammadpour and K.M. Grigoriadis, *Efficient Modeling and Control of Large-Scale Systems*, DOI 10.1007/978-1-4419-5757-3_1,

Interpolatory model reduction methods have emerged as effective strategies for approximation of large-scale linear dynamical systems. These methods produce reduced models whose transfer function interpolates the original system transfer function at selected interpolation points. They are closely related to rational Krylov-based methods and indeed, are precisely that for single input/single output systems. One main reason why interpolatory model reduction has become the method of choice in true large-scale settings is that this class of methods is numerically stable and well-suited to large scale computations; the model reduction process does not require any dense matrix transformations such as singular value decompositions. Moreover, the approximation theory behind this approach is similar to that of rational interpolants used to approximate meromorphic functions. The goal of this contribution is to present a survey of model reduction by interpolation and provide a selection of the most recent developments in this direction.

2 Problem Setting

Linear dynamical systems are principally characterized through their input–output map $\mathcal{S} : \mathbf{u} \mapsto \mathbf{y}$, mapping inputs $\mathbf{u}$ to outputs $\mathbf{y}$ via a state-space realization given as:

$$\mathcal{S} : \begin{cases} \mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \quad \text{with } \mathbf{x}(0) = \mathbf{0}, \tag{1}$$

where $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$, and $\mathbf{D} \in \mathbb{R}^{p\times m}$ are constant matrices. In (1), $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{u}(t) \in \mathbb{R}^m$ and $\mathbf{y}(t) \in \mathbb{R}^p$ are, respectively, an internal variable (the state if $\mathbf{E}$ is non-singular), the input and the output of the system $\mathcal{S}$. We refer to $\mathcal{S}$ as a *single-input/single-output* (SISO) system when $m = p = 1$ (scalar-valued input, scalar-valued output) and as a *multi-input/multi-output* (MIMO) system otherwise.

Systems of the form (1) with extremely large state-space dimension n arise in many disciplines; see [2] and [57] for a collection of such examples. Despite large state-space dimension, in most cases, the state space trajectories, $\mathbf{x}(t)$, hew closely to subspaces with substantially lower dimensions and evolve in ways that do not fully occupy the state space. The original model $\mathcal{S}$ behaves nearly as if it had many fewer internal degrees-of-freedom, in effect, much lower state-space dimension. *Model reduction* seeks to produce a surrogate dynamical system that evolves in a much lower dimensional state space (say, dimension $r \ll n$), yet is able to mimic the original dynamical system, recovering very nearly the original input–output map. We want the reduced input–output map, $\mathcal{S}_r : \mathbf{u} \mapsto \mathbf{y}_r$, to be close to $\mathcal{S}$ in an appropriate sense.

Being a smaller version of the original dynamical model, the input–output map $\mathcal{S}_r$ is described by the reduced system in state-space form as:

$$\mathcal{S}_r : \begin{cases} \mathbf{E}_r\dot{\mathbf{x}}_r(t) = \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t) + \mathbf{D}_r\mathbf{u}(t) \end{cases} \quad \text{with } \mathbf{x}_r(0) = \mathbf{0}, \tag{2}$$

where $\mathbf{A}_r, \mathbf{E}_r \in \mathbb{R}^{r\times r}$, $\mathbf{B}_r \in \mathbb{R}^{r\times m}$, $\mathbf{C}_r \in \mathbb{R}^{p\times r}$ and $\mathbf{D}_r \in \mathbb{R}^{p\times m}$. Note that the number of inputs, m, and the number of outputs, p, are the same for both the original and reduced models; only the internal state-space dimensions differ: $r \ll n$. A successful reduced-order model should meet the following criteria:

Goals for Reduced Order Models

1. The reduced input–output map $\mathcal{S}_r$ should be uniformly close to $\mathcal{S}$ in an appropriate sense. That is, when presented with the same inputs, $\mathbf{u}(t)$, the difference between full and reduced system outputs, $\mathbf{y} - \mathbf{y}_r$, should be *small* with respect to a physically relevant norm over a wide range of system inputs, such as over all $\mathbf{u}$ with bounded energy (e.g., in the unit ball of $L^2([0,\infty),\mathbb{R}^m)$).
2. Critical system features and structure should be preserved in the reduced order system. This can include passivity, Hamiltonian structure, subsystem interconnectivity, or second-order structure.
3. Strategies for obtaining $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ should lead to robust, numerically stable algorithms and furthermore require minimal application-specific tuning with little to no expert intervention. It is important that model reduction methods be computationally efficient and reliable so that very large problems remain tractable; they should be robust and largely automatic to allow the broadest level of flexibility and applicability in complex multiphysics settings.

Whenever the input $\mathbf{u}(t)$ is exponentially bounded – that is, when there is a fixed $\gamma \in \mathbb{R}$ such that $\|\mathbf{u}(t)\| \sim \mathcal{O}(e^{\gamma t})$, then $\mathbf{x}(t)$ and $\mathbf{y}(t)$ from (1) and $\mathbf{x}_r(t)$ and $\mathbf{y}_r(t)$ from (2) will also be exponentially bounded and the Laplace transform can be applied to (1) and (2) to obtain

$$\widehat{\mathbf{y}}(s) = \left(\mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D}\right)\widehat{\mathbf{u}}(s), \tag{3}$$

$$\widehat{\mathbf{y}}_r(s) = \left(\mathbf{C}_r\left(s\mathbf{E}_r - \mathbf{A}_r\right)^{-1}\mathbf{B}_r + \mathbf{D}_r\right)\widehat{\mathbf{u}}(s), \tag{4}$$

where we have denoted Laplace transformed quantities with "$\widehat{\ }$". We define transfer functions accordingly – from (3) and (4)

$$\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{E} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{5}$$

$$\mathbf{H}_r(s) = \mathbf{C}_r\left(s\mathbf{E}_r - \mathbf{A}_r\right)^{-1}\mathbf{B}_r + \mathbf{D}_r \tag{6}$$

so that

$$\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s) = \left[\mathbf{H}(s) - \mathbf{H}_r(s)\right]\widehat{\mathbf{u}}(s)$$

For a given input, the loss in fidelity in passing from the full model to a reduced model can be associated with the difference between respective system transfer functions over a range of inputs. This is evaluated more precisely in Sect. 2.4.

It is a fact of life that many dynamical systems of critical interest are *nonlinear*. By contrast, we will consider here only linear dynamical systems and the methods we discuss fundamentally exploit this linearity. We do this unapologetically but note that by narrowing the scope of the problems that one considers in this way, one is able to be far more ambitious and demanding in the quality of outcomes that are achieved, and indeed, the techniques we discuss here are often dramatically successful. To the extent that many methods for approaching nonlinear systems build upon the analysis of carefully chosen linearizations of these systems, the methods we describe here can be expected to play a role in evolving strategies for the reduction of large scale nonlinear dynamical systems as well. See, for instance, [12].

2.1 The General Interpolation Framework

Interpolation is a very simple and yet effective approach that is used ubiquitously for the general approximation of complex functions using simpler ones. The construction of such approximations is easy: indeed, school children learn how to construct piecewise linear approximations and Taylor polynomial approximations to functions through interpolation. The accuracy of the resulting approximations and the connections with strategic placement of interpolating points has been studied in many broad contexts – indeed, in the case of interpolation of meromorphic functions by polynomials or rational functions, the associated error analysis constitutes a large body of work tied in closely with potential theory and classical complex analysis. Many of these recognized advantages of approximation via interpolation translate immediately into boons in our setting as well as we seek to approximate a "difficult" transfer function, $\mathbf{H}(s)$, with a simpler one, $\mathbf{H}_r(s)$.

Simply stated, our overarching goal is to produce a low order transfer function, $\mathbf{H}_r(s)$, that approximates a large order transfer function, $\mathbf{H}(s)$ with high fidelity: for order $r \ll n$, we want $\mathbf{H}_r(s) \approx \mathbf{H}(s)$ in a sense that will be made precise later. Interpolation is the primary vehicle with which we approach this problem. Naively, we could select a set of points $\{\sigma_i\}_{i=1}^r \subset \mathbb{C}$ and then seek a reduced order transfer function, $\mathbf{H}_r(s)$, such that $\mathbf{H}_r(\sigma_i) = \mathbf{H}(\sigma_i)$ for $i = 1, \ldots, r$. This is a good starting place for SISO systems but turns out to be overly restrictive for MIMO systems, since the condition $\mathbf{H}_r(\sigma_i) = \mathbf{H}(\sigma_i)$ in effect imposes $m \cdot p$ scalar conditions at each interpolation point. This may not be realizable.

A far more advantageous formulation is to consider interpolation conditions that are imposed at specified interpolation points as above but that only act in specified directions, that is, *tangential interpolation*. We state this formally as part of our first problem of focus:

Problem 1: Model Reduction Given State Space System Data

Given a full-order model (1) specified through knowledge of the system matrices, $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ and given

left interpolation points:		*right interpolation points*:
$\{\mu_i\}_{i=1}^{q} \subset \mathbb{C}$,		$\{\sigma_i\}_{i=1}^{r} \subset \mathbb{C}$
with corresponding	and	with corresponding
left tangent directions:		*right tangent directions*:
$\{\tilde{\mathsf{c}}_i\}_{i=1}^{q} \subset \mathbb{C}^p$,		$\{\tilde{\mathsf{b}}_i\}_{i=1}^{r} \subset \mathbb{C}^m$.

Find a reduced-order model (2) through specification of reduced system matrices $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ such that the associated transfer function, $\mathbf{H}_r(s)$, in (6) is a *tangential interpolant* to $\mathbf{H}(s)$ in (5):

$$\begin{array}{c} \tilde{\mathsf{c}}_i^T \mathbf{H}_r(\mu_i) = \tilde{\mathsf{c}}_i^T \mathbf{H}(\mu_i) \\ \text{for } i = 1, \ldots, q, \end{array} \quad \text{and} \quad \begin{array}{c} \mathbf{H}_r(\sigma_j)\tilde{\mathsf{b}}_j = \mathbf{H}(\sigma_j)\tilde{\mathsf{b}}_j, \\ \text{for } j = 1, \ldots, r, \end{array} \tag{7}$$

Interpolation points and tangent directions are selected to realize the model reduction goals described on page 5.

It is of central concern to determine effective choices for interpolation points and tangent directions. In Sect. 3, we discuss the selection of interpolation points and tangent directions that can yield *optimal* reduced order models with respect to $\mathcal{H}_2$ quality measures of system approximation. In Sect. 4, we describe selection of interpolation points and tangent directions that guarantee passivity is preserved in the reduced-order model. Preservation of other types of system structure such as second-order systems, systems that involve delays or memory terms, and systems having a structured dependence on parameters is discussed in Sects. 5 and 6.

Tangential interpolation is a remarkably flexible framework within which to consider model reduction when one considers the significance of the interpolation data. Note that if $\tilde{\mathsf{y}} = \mathbf{H}(\sigma)\tilde{\mathsf{b}}$ then $e^{\sigma t}\tilde{\mathsf{y}}$ is precisely the response of the full order system to a pure input given by $\mathbf{u}(t) = e^{\sigma t}\tilde{\mathsf{b}}$, so the tangential interpolation conditions that characterize $\mathbf{H}_r(s)$ could (at least in principle) be obtained from *measured input–output data* drawn directly from observations on the original system. For example, if $\sigma = \imath\omega_0$, one is observing in $\tilde{\mathsf{y}}$, the sinusoidal response of the system to a pure tone input of frequency ω_0. Similarly, if the *dual* dynamical system were driven by an input given by $e^{\mu t}\tilde{\mathsf{c}}$ producing an output $e^{\mu t}\tilde{\mathsf{z}}$ then $\tilde{\mathsf{c}}^T\mathbf{H}(\mu) = \tilde{\mathsf{z}}^T$. This creates the following alternative problem setting that we consider based entirely on observed input–output response data and requiring no other a priori information about system-related quantities.

Problem 2: Model Reduction Given Input–Output Data

Given a set of input–output response measurements on the full-order system specified by

left driving frequencies:		*right driving frequencies*:
$\{\mu_i\}_{i=1}^{q} \subset \mathbb{C}$,		$\{\sigma_i\}_{i=1}^{r} \subset \mathbb{C}$
using *left input directions*:	and	using *right input directions*:
$\{\tilde{c}_i\}_{i=1}^{q} \subset \mathbb{C}^p$,		$\{\tilde{b}_i\}_{i=1}^{r} \subset \mathbb{C}^m$
producing *left responses*:		producing *right responses*:
$\{\tilde{z}_i\}_{i=1}^{q} \subset \mathbb{C}^m$,		$\{\tilde{y}_i\}_{i=1}^{r} \subset \mathbb{C}^p$

Find a dynamical system model (2) by specifying (reduced) system matrices $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ such that the associated transfer function, $\mathbf{H}_r(s)$, in (6) is a *tangential interpolant* to the given data:

$$\begin{array}{ll} \tilde{c}_i^T \mathbf{H}_r(\mu_i) = \tilde{z}_i^T & \text{and} \qquad \mathbf{H}_r(\sigma_j)\tilde{b}_j = \tilde{y}_j, \\ \text{for } i = 1, \ldots, q, & \text{for } j = 1, \ldots, r, \end{array} \tag{8}$$

Interpolation points and tangent directions are determined (typically) by the availability of experimental data.

The structure of Problem 2 introduces difficulties involved in the requirement that noisy and inconsistent data must be accommodated appropriately – in this case, we describe how to construct a reduced order system described by the system matrices $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ so that experimentally observed response data (reformulated as approximate interpolation conditions) are matched at least approximately: $\mathbf{H}_r(\sigma_i)\tilde{b}_i \approx \tilde{y}_i$ and $\tilde{c}_j^T \mathbf{H}(\mu_j) \approx \tilde{z}_j^T$.

Numerous issues arise spontaneously for both problem types and we discuss here what we feel is a representative subset. Note for both problems, it is necessary to have a computationally stable method for constructing the (reduced) system matrices $\mathbf{E}_r$, $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, and $\mathbf{D}_r$ that produces an associated transfer function, $\mathbf{H}_r(s)$, satisfying the interpolation conditions. We discuss this in Sects. 2.2 and 2.3 for Problem 1 and Sect. 7.2 for Problem 2.

2.2 Model Reduction via Projection

When the full-order dynamical system $\mathcal{S}$ described in (1) has been specified via the state-space matrices $\mathbf{E}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, most model reduction methods proceed with some variation of a Petrov-Galerkin projective approximation to construct a reduced-order model $\mathcal{S}_r$. We motivate this approach by describing the evolution of the full order model (1) in an indirect way – although the systems described in

(1) and (2) typically are real-valued, we find it convenient to allow all quantities involved to be complex-valued:

$$\begin{aligned}&\text{Find } \mathbf{x}(t) \text{ contained in } \mathbb{C}^n \text{ such that}\\ &\qquad \mathbf{E}\dot{\mathbf{x}}(t) - \mathbf{A}\mathbf{x}(t) - \mathbf{B}\,\mathbf{u}(t) \quad \perp \quad \mathbb{C}^n \text{ (i.e., } = 0).\\ &\text{Then the associated output is } \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t).\end{aligned}$$

The Petrov-Galerkin projective approximation proceeds by replacing $\mathbb{C}^n$ in the above description with (nominally small) subspaces of $\mathbb{C}^n$: we choose an r-dimensional trial subspace, the *right modeling subspace*, $\mathcal{V}_r \subset \mathbb{C}^n$, and an r-dimensional test subspace, the *left modeling subspace*, $\mathcal{W}_r \subset \mathbb{C}^n$ – and then we describe the evolution of a reduced order model in this way:

$$\begin{aligned}&\text{Find } \mathbf{v}(t) \text{ contained in } \mathcal{V}_r \text{ such that}\\ &\qquad \mathbf{E}\dot{\mathbf{v}}(t) - \mathbf{A}\mathbf{v}(t) - \mathbf{B}\,\mathbf{u}(t) \quad \perp \quad \mathcal{W}_r.\\ &\text{Then the associated output is } \mathbf{y}_r(t) = \mathbf{C}\mathbf{v}(t) + \mathbf{D}\mathbf{u}(t).\end{aligned} \tag{9}$$

The dynamics described by (9) can be represented as a dynamical system evolving in a reduced order state-space as in (2) once bases are chosen for the two subspaces $\mathcal{V}_r$ and $\mathcal{W}_r$. Let $\mathsf{Ran}(\mathbf{R})$ denote the range of a matrix $\mathbf{R}$. Let $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ and $\mathbf{W}_r \in \mathbb{C}^{n\times r}$ be matrices defined so that $\mathcal{V}_r = \mathsf{Ran}(\mathbf{V}_r)$ and $\mathcal{W}_r = \mathsf{Ran}(\overline{\mathbf{W}}_r)$. We can represent the reduced system trajectories as $\mathbf{v}(t) = \mathbf{V}_r\mathbf{x}_r(t)$ with $\mathbf{x}_r(t) \in \mathbb{C}^r$ for each t and the Petrov-Galerkin approximation (9) can be rewritten as

$$\mathbf{W}_r^T\left(\mathbf{E}\mathbf{V}_r\dot{\mathbf{x}}_r(t) - \mathbf{A}\mathbf{V}_r\mathbf{x}_r(t) - \mathbf{B}\,\mathbf{u}(t)\right) = \mathbf{0} \quad \text{and} \quad \mathbf{y}_r(t) = \mathbf{C}\mathbf{V}_r\mathbf{x}_r(t) + \mathbf{D}\mathbf{u}(t),$$

leading to the reduced order state-space representation (2) with

$$\begin{aligned}\mathbf{E}_r &= \mathbf{W}_r^T\mathbf{E}\mathbf{V}_r, \quad & \mathbf{B}_r &= \mathbf{W}_r^T\mathbf{B},\\ \mathbf{A}_r &= \mathbf{W}_r^T\mathbf{A}\mathbf{V}_r, \quad & \mathbf{C}_r &= \mathbf{C}\mathbf{V}_r,\end{aligned} \qquad \text{and} \qquad \mathbf{D}_r = \mathbf{D}. \tag{10}$$

We note that the definitions of reduced-order quantities in (10) are invariant under change of basis for the original state space, so the quality of reduced approximations evidently will depend only on effective choices for the right modeling space $\mathcal{V}_r = \mathsf{Ran}(\mathbf{V}_r)$ and the left modeling space $\mathcal{W}_r = \mathsf{Ran}(\overline{\mathbf{W}}_r)$. We choose the modeling subspaces to enforce interpolation which allows us to shift our focus to how best to choose effective interpolation points and tangent directions.

Since $\mathbf{D} \in \mathbb{R}^{p\times m}$ and both p and m are typically of only modest size, usually $\mathbf{D}$ does not play a significant role in the cost of simulation and $\mathbf{D}_r = \mathbf{D}$ is both a common choice and a natural one arising in the context of a Petrov-Galerkin approximation as described in (10). Note that if $\mathbf{E}$ and $\mathbf{E}_r$ are both nonsingular then the choice $\mathbf{D}_r = \mathbf{D}$ also enforces interpolation at infinity: $\lim_{s\to\infty}\mathbf{H}(s) = \lim_{s\to\infty}\mathbf{H}_r(s) = \mathbf{D}$, facilitating, in effect, a good match between true and reduced system outputs for sufficiently high frequency inputs. The case that $\mathbf{E}$ is singular and other performance goals can motivate choices for $\mathbf{D}_r \neq \mathbf{D}$. This is discussed in some detail on p. 13.

2.3 Interpolatory Projections

Given the full-order dynamical system $\mathbf{H}(s)$, a set of r right interpolation points $\{\sigma_i\}_{i=1}^{r} \in \mathbb{C}$, with right directions $\{\mathsf{b}_i\}_{i=1}^{r} \in \mathbb{C}^m$, and a set of q left interpolation points $\{\mu_i\}_{i=1}^{q} \in \mathbb{C}$, with left-tangential directions $\{\mathsf{c}_i\}_{i=1}^{q} \in \mathbb{C}^p$, interpolatory model reduction involves finding a reduced-order model $\mathbf{H}_r(s)$ that tangentially interpolates $\mathbf{H}(s)$ using this given data: find $\mathbf{H}_r(s)$ so that

$$\begin{aligned} \mathbf{H}(\sigma_i)\mathsf{b}_i &= \mathbf{H}_r(\sigma_i)\mathsf{b}_i, \qquad \text{for } i = 1,\ldots,r, \\ \mathsf{c}_j^T\mathbf{H}(\mu_j) &= \mathsf{c}_j^T\mathbf{H}_r(\mu_j), \qquad \text{for } j = 1,\ldots,r, \end{aligned} \tag{11}$$

We wish to interpolate $\mathbf{H}(s)$ without ever computing the quantities to be matched since these numbers are numerically ill-conditioned, as Feldman and Freund [33] illustrated for the special case of single-input/single-output dynamical systems. Remarkably, this can be achieved by employing Petrov-Galerkin projections with carefully chosen test and trial subspaces.

Interpolatory projections for model reduction were initially introduced by Skelton et. al. in [30, 90, 91]. Grimme [42], later, modified this approach into a numerically efficient framework by utilizing the rational Krylov subspace method of Ruhe [79]. The projection framework for the problem setting we are interested in, that is, for rational tangential interpolation of MIMO dynamical systems, has been recently developed by Gallivan et al. [39] for the case of Problem 1 and by Mayo and Antoulas [66] for the case of Problems 2 as laid out in §2.1. We will be using these frameworks throughout the manuscript.

We first illustrate and prove how to solve the rational tangential interpolation problem in (11) by projection. The general case for matching derivatives of the transfer function (in effect, generalized Hermite interpolation) will be presented in Theorem 1.2. As we point out above, the Petrov-Galerkin approximation described in (10) leads to the choice of $\mathbf{D}_r = \mathbf{D}$ in a natural way, but other choices are possible. We will first assume $\mathbf{D}_r = \mathbf{D}$, so that $\mathbf{H}_r(s)$ in Theorems 1.1 and 1.2 below is assumed to be obtained as in (10).

Theorem 1.1. *Let $\sigma, \mu \in \mathbb{C}$ be such that $s\mathbf{E} - \mathbf{A}$ and $s\mathbf{E}_r - \mathbf{A}_r$ are invertible for $s = \sigma, \mu$. Also let $\mathbf{V}_r, \mathbf{W}_r \in \mathbb{C}^{n\times r}$ in (10) have full-rank. If $\mathsf{b} \in \mathbb{C}^m$ and $\mathsf{c} \in \mathbb{C}^\ell$ are fixed nontrivial vectors then*

(a) if $(\sigma\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b} \in \mathsf{Ran}(\mathbf{V}_r)$, then $\mathbf{H}(\sigma)\mathsf{b} = \mathbf{H}_r(\sigma)\mathsf{b}$;
(b) if $\left(\mathsf{c}^T\mathbf{C}\,(\mu\,\mathbf{E}-\mathbf{A})^{-1}\right)^T \in \mathsf{Ran}(\mathbf{W}_r)$, then $\mathsf{c}^T\mathbf{H}(\mu) = \mathsf{c}^T\mathbf{H}_r(\mu)$; and
(c) if both (a) and (b) hold, and $\sigma = \mu$, then $\mathsf{c}^T\mathbf{H}'(\sigma)\mathsf{b} = \mathsf{c}^T\mathbf{H}_r'(\sigma)\mathsf{b}$ as well.

Proof. We follow the recent proofs provided in [49] and [16]. Define

$$\begin{aligned} &\mathcal{P}_r(z) = \mathbf{V}_r(z\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{W}_r^T(z\mathbf{E}_r - \mathbf{A}) \quad \text{and} \\ &\mathcal{Q}_r(z) = (z\mathbf{E}-\mathbf{A})\mathcal{P}_r(z)(z\mathbf{E}-\mathbf{A})^{-1} = (z\mathbf{E}-\mathbf{A})\mathbf{V}_r(z\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{W}_r^T. \end{aligned}$$

Both $\mathcal{P}_r(z)$ and $\mathcal{Q}_r(z)$ are analytic matrix-valued functions in neighborhoods of $z=\sigma$ and $z=\mu$. It is easy to verify that both $\mathcal{P}_r(z)$ and $\mathcal{Q}_r(z)$ are projectors, i.e. $\mathcal{P}_r^2(z)=\mathcal{P}_r(z)$ and $\mathcal{Q}_r^2(z)=\mathcal{Q}_r(z)$. Moreover, for all z in a neighborhood of σ, $\mathcal{V}_r=\mathsf{Ran}(\mathcal{P}_r(z))=\mathsf{Ker}\,(\mathbf{I}-\mathcal{P}_r(z))$ and $\mathcal{W}_r^{\perp}=\mathsf{Ker}(\mathcal{Q}_r(z))=\mathsf{Ran}\,(\mathbf{I}-\mathcal{Q}_r(z))$ where $\mathsf{Ran}(\mathbf{R})$ denotes the kernel of a matrix $\mathbf{R}$. First observe that

$$\mathbf{H}(z)-\mathbf{H}_r(z)=\mathbf{C}(z\mathbf{E}-\mathbf{A})^{-1}\left(\mathbf{I}-\mathcal{Q}_r(z)\right)(z\mathbf{E}-\mathbf{A})\Big(\mathbf{I}-\mathcal{P}_r(z)\Big)(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} \quad (12)$$

Evaluating this expression at $z=\sigma$ and post-multiplying by b yields the first assertion; evaluating (12) at $z=\mu$ and pre-multiplying by c^T yields the second. Note

$$((\sigma+\varepsilon)\mathbf{E}-\mathbf{A})^{-1}=(\sigma\mathbf{E}-\mathbf{A})^{-1}-\varepsilon(\sigma\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\sigma\mathbf{E}-\mathbf{A})^{-1}+\mathcal{O}(\varepsilon^2)$$
$$((\sigma+\varepsilon)\mathbf{E}_r-\mathbf{A}_r)^{-1}=(\sigma\mathbf{E}_r-\mathbf{A}_r)^{-1}-\varepsilon(\sigma\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{E}_r(\sigma\mathbf{E}_r-\mathbf{A}_r)^{-1}+\mathcal{O}(\varepsilon^2)$$

so evaluating (12) at $z=\sigma+\varepsilon$, premultiplying by c^T, and postmultiplying by b under the hypotheses of the third assertion yields

$$\mathsf{c}^T\mathbf{H}(\sigma+\varepsilon)\mathsf{b}-\mathsf{c}^T\mathbf{H}_r(\sigma+\varepsilon)\mathsf{b}=\mathcal{O}(\varepsilon^2).$$

Now observe that since $\mathsf{c}^T\mathbf{H}(\sigma)\mathsf{b}=\mathsf{c}^T\mathbf{H}_r(\sigma)\mathsf{b}$,

$$\frac{1}{\varepsilon}\left(\mathsf{c}^T\mathbf{H}(\sigma+\varepsilon)\mathsf{b}-\mathsf{c}^T\mathbf{H}(\sigma)\mathsf{b}\right)-\frac{1}{\varepsilon}\left(\mathsf{c}^T\mathbf{H}_r(\sigma+\varepsilon)\mathsf{b}-\mathsf{c}^T\mathbf{H}_r(\sigma)\mathsf{b}\right)\to 0,$$

as $\varepsilon\to 0$, which proves the third assertion. □

Theorem 1.1 shows that given a set of distinct right interpolation points ("right shifts") $\{\sigma_i\}_{i=1}^r$, a set of distinct interpolation left points ("left shifts") $\{\mu_j\}_{j=1}^r$, left tangential directions $\{\mathsf{c}_k\}_{k=1}^r\in\mathbb{C}^p$, and right tangential directions $\{\mathsf{b}_k\}_{k=1}^r\in\mathbb{C}^m$, the solution of the tangential rational Hermite interpolation problem is straightforward and constructive. One simply computes matrices $\mathbf{V}_r$ and $\mathbf{W}_r$ such that

$$\mathbf{V}_r=\left[(\sigma_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b}_1,\ldots,(\sigma_r\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b}_r\right] \text{ and} \quad (13)$$

$$\mathbf{W}_r^T=\begin{bmatrix}\mathsf{c}_1^T\mathbf{C}(\mu_1\mathbf{E}-\mathbf{A})^{-1}\\ \vdots\\ \mathsf{c}_r^T\mathbf{C}(\mu_r\mathbf{E}-\mathbf{A})^{-1}\end{bmatrix}. \quad (14)$$

The reduced order system $\mathbf{H}_r(s)=\mathbf{C}_r(s\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{B}_r$ defined by (10) then solves the tangential Hermite interpolation problem provided that $\sigma_i\mathbf{E}_r-\mathbf{A}_r$ and $\mu_i\mathbf{E}_r-\mathbf{A}_r$ are nonsingular for each $i=1,\ldots,r$.

For completeness, we describe how to solve the analogous generalized Hermite interpolation problem, which involves matching transfer function values and higher derivative values. The mth derivative of $\mathbf{H}(s)$ with respect to s evaluated at $s=\sigma$, will be denoted by $\mathbf{H}^{(m)}(\sigma)$.

Theorem 1.2. *Let* $\sigma, \mu \in \mathbb{C}$ *be such that* $s\mathbf{E} - \mathbf{A}$ *and* $s\mathbf{E}_r - \mathbf{A}_r$ *are invertible for* $s = \sigma, \mu$. *If* $\mathsf{b} \in \mathbb{C}^m$ *and* $\mathsf{c} \in \mathbb{C}^\ell$ *are fixed nontrivial vectors then*

(a) if $\left((\sigma\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}\right)^{j-1}(\sigma\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\mathsf{b} \in \mathsf{Ran}(\mathbf{V}_r)$ *for* $j = 1,.,N$

then $\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$ *for* $\ell = 0,1,\ldots,N-1$

(b) if $\left((\mu\mathbf{E}-\mathbf{A})^{-T}\mathbf{E}^T\right)^{j-1}(\mu\mathbf{E}-\mathbf{A})^{-T}\mathbf{C}^T\mathsf{c} \in \mathsf{Ran}(\mathbf{W}_r)$ *for* $j = 1,.,M$,

then $\mathsf{c}^T\mathbf{H}^{(\ell)}(\mu) = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\mu)\mathsf{b}$ *for* $\ell = 0,1,\ldots,M-1$;

(c) if both (a) and (b) hold, and if $\sigma = \mu$, *then* $\mathsf{c}^T\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$,
for $\ell = 1,\ldots,M+N+1$

Theorem 1.2 solves the rational tangential interpolation problem via projection. All one has to do is to construct the matrices $\mathbf{V}_r$ and $\mathbf{W}_r$ according to the theorem, and the reduced order model is guaranteed to satisfy the interpolation conditions. Note that Theorems 1.1 and 1.2 solve the interpolation problem without ever explicitly computing the values that are interpolated, since that computation is known to be prone to poor numerical conditioning.

Model reduction by interpolation (also called Krylov-based model reduction or moment matching in the single-input/single-output case) has been recognized previously, especially in the circuit simulation community; see, for example, [9–11, 33, 37, 73] and the references therein. [74, 76] also investigated the interpolation problem for efficient circuit analysis, however required explicit computation of the transfer function values and derivatives to be matched.

Despite their computational efficiency compared to Gramian-based model reduction techniques such as balanced-truncation [70, 72] and optimal Hankel norm approximation [40], interpolatory model reduction used to have the main drawback of depending on an ad hoc selection of interpolation points and tangential directions, i.e. how to choose the interpolation points and tangential directions optimally or at least effectively were not known until very recently. This has raised some criticism of interpolatory methods for not guaranteeing stability and not providing error bounds or error estimates for the resulting reduced-order model. To solve the stability issue, an implicit restart technique was proposed by Grimme et al. [43] which removed the unstable poles via an implicitly restarted Arnoldi process; however the reduced-model no longer interpolated the original one, but rather a nearby one. Gugercin and Antoulas [47] proposed combining interpolatory projection methods with Gramian-based approaches that guaranteed stability of the reduced-order model. Recently, reformulation of the Fourier model reduction approach of Willcox and Megretzki [87] in the interpolation framework by Gugercin and Willcox [50] has shown this method is indeed a stability preserving interpolatory model reduction method with an $\mathcal{H}_\infty$ upper bound, not necessarily easily computable. Bai [93], and Gugercin and Antoulas [46] have also focused on the error estimates and provided *error expressions* (not error bounds) for the interpolatory model reduction. However, despite these efforts, the optimal interpolation points and tangential direction selection remained unresolved until very recently.

This optimal point selection strategy together with some other very recent developments in interpolation-based model reduction will be presented in the remaining of the manuscript: Sect. 3 will illustrate how to choose the interpolation points to minimize the $\mathcal{H}_2$ norm of the error systems. Passivity of the reduced-model is vital when dealing with model coming from, especially, circuit theory. Section 4 will show how to construct high-quality passivity-preserving reduced-order models using interpolation. Obtaining reduced-order models solely from input-output measurements without the knowledge of the original state-space quantities has been the focus of recent research. In Sect. 7, we will illustrate how to achieve this in the interpolation framework. Even though the state-space formulation of the original system in (1) is quite general, and indeed, will include, with suitable reformulation, the majority of linear dynamical systems, some dynamical systems might have a natural description that takes a form quite different from this standard realization; such as systems with internal delays, memory terms, etc. Section 5 will extend the interpolatory model reduction theory described here to settings where the transfer functions $\mathbf{H}(s)$ take a much more general form than that of (5), allowing interpolatory model reduction for a much richer category of dynamical systems.

Constructing Interpolants with $\mathbf{D}_r \neq \mathbf{D}$. Certain circumstances in model reduction require that a reduced-order model have a $\mathbf{D}_r$ term different than $\mathbf{D}$. One such case is that of a singular $\mathbf{E}$ with a non-defective eigenvalue at 0. This occurs when the internal system dynamics has an auxiliary algebraic constraint that must always be satisfied (perhaps representing rigid body translations or rotations or fluid incompressibility, for example). The dynamical system description given in (1) is then a *differential algebraic equation* (DAE) and our condition that $\mathbf{E}$ have a non-defective eigenvalue at 0 amounts to the requirement that the DAE be of index 1 (see [60]). In this case, $\lim_{s\to\infty} \mathbf{H}(s) \neq \mathbf{D}$. If we want $\mathbf{H}_r(s)$ to match $\mathbf{H}(s)$ asymptotically at high frequencies then we should choose

$$\mathbf{D}_r = \lim_{s\to\infty} \left(\mathbf{H}(s) - \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r\right)$$

If $r < \text{rank}(\mathbf{E})$ then it will often happen that $\mathbf{E}_r$ will be nonsingular, in which case then $\lim_{s\to\infty} \mathbf{H}_r(s) = \mathbf{D}_r$ and we may assign $\mathbf{D}_r = \lim_{s\to\infty} \mathbf{H}(s)$ if we wish $\mathbf{H}_r(s)$ to match $\mathbf{H}(s)$ asymptotically at high frequencies. See, for example, [21] where a reduced-order model with $\mathbf{D}_r \neq \mathbf{D}$ constructed in case of a singular $\mathbf{E}$ to match $\mathbf{H}(s)$ and its higher derivatives around $s = \infty$.

One may be less concerned with $\mathbf{H}_r(s)$ matching $\mathbf{H}(s)$ well at high frequencies but instead may wish to minimize the maximum mismatch between $\mathbf{H}(s)$ and $\mathbf{H}_r(s)$ over the imaginary axis (best $\mathcal{H}_\infty$ approximation). Flexibility in choosing the $\mathbf{D}_r$ term is necessary in this case as well. How to incorporate arbitrary choices of $\mathbf{D}_r$ without losing interpolation properties, is described in the next theorem, first presented in [66] and later generalized in [16]. For simplicity, we assume $\mathbf{D} = \mathbf{0}$, i.e.

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}. \tag{15}$$

The general case with $\mathbf{D} \neq \mathbf{0}$ is recovered by replacing $\mathbf{D}_r$ with $\mathbf{D}_r - \mathbf{D}$.

Theorem 1.3. *Given* $\mathbf{H}(s)$ *as in (15), 2r distinct points,* $\{\mu_i\}_{i=1}^r \cup \{\sigma_j\}_{j=1}^r$, *in the right halfplane, together with 2r nontrivial vectors,* $\{\mathsf{c}_i\}_{i=1}^r \subset \mathbb{C}^p$ *and* $\{\mathsf{b}_j\}_{j=1}^r \subset \mathbb{C}^m$, *let* $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ *and* $\mathbf{W}_r \in \mathbb{C}^{n\times r}$ *be as in (13) and (14), respectively. Define* $\widetilde{\mathsf{B}}$ *and* $\widetilde{\mathsf{C}}$ *as*

$$\widetilde{\mathsf{B}} = [\mathsf{b}_1, \mathsf{b}_2, ..., \mathsf{b}_r] \quad \textit{and} \quad \widetilde{\mathsf{C}}^T = [\mathsf{c}_1, \mathsf{c}_2, \ldots, \mathsf{c}_r]^T$$

For any $\mathbf{D}_r \in \mathbb{C}^{p\times m}$, *define*

$$\begin{aligned} \mathbf{E}_r &= \mathbf{W}_r^T \mathbf{E} \mathbf{V}_r, \qquad \mathbf{A}_r = \mathbf{W}_r^T \mathbf{A} \mathbf{V}_r + \widetilde{\mathsf{C}}^T \mathbf{D}_r \widetilde{\mathsf{B}}, \\ \mathbf{B}_r &= \mathbf{W}_r^T \mathbf{B} - \widetilde{\mathsf{C}}^T \mathbf{D}_r, \qquad \text{and } \mathbf{C}_r = \mathbf{C}\mathbf{V}_r - \mathbf{D}_r \widetilde{\mathsf{B}} \end{aligned} \tag{16}$$

Then the reduced-order model $\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r + \mathbf{D}_r$ *satisfies*

$$\mathbf{H}(\sigma_i)\mathsf{b}_i = \mathbf{H}_r(\sigma_i)\mathsf{b}_i \quad \textit{and} \quad \mathsf{c}_i^T \mathbf{H}(\mu_i) = \mathsf{c}_i^T \mathbf{H}_r(\mu_i) \quad \textit{for } i = 1, ..., r.$$

Theorem 1.3 describes how to construct an interpolant with an *arbitrary* $\mathbf{D}_r$ term; hence the reduced-order model $\mathbf{H}_r$ solves the rational interpolation problem for any $\mathbf{D}_r$ that can be chosen to satisfy specific design goals.

It may happen that the state space representation of the reduced system is not minimal (for an example see [31]), in which case interpolation of some of the data is lost. Theorem 1.3 provides a remedy: there always exists $\mathbf{D}_r$, such that the realization given by (16) is minimal, i.e., both controllable and observable, and hence interpolation of all the data is guaranteed.

2.4 Error Measures

The loss in accuracy in replacing a full-order model with a reduced model is measured by the difference between respective system outputs over a range of inputs, which is best understood in the frequency domain. For the full-order system described by the transfer function $\mathbf{H}(s)$ as in (5) and the reduced-order system by $\mathbf{H}_r(s)$ as in (6), the *error system*, in the time domain, is defined by $\mathcal{S}_{error} : \mathbf{u}(t) \mapsto [\mathbf{y} - \mathbf{y}_r](t)$. Equivalently, we can represent the error in the frequency domain by $\widehat{\mathbf{y}}(s) - \widehat{\mathbf{y}}_r(s) = [\mathbf{H}(s) - \mathbf{H}_r(s)]\,\widehat{\mathbf{u}}(s)$; hence the error system can be considered having the transfer function $\mathbf{H}(s) - \mathbf{H}_r(s)$.

We consider the reduced-order system to be a good approximation of the full-order system if the error system $\mathcal{S}_{error}$ is small in a sense we wish to make precise. Ultimately, we want to produce a reduced order system that will yield outputs, $\mathbf{y}_r(t)$, such that $\mathbf{y}_r(t) \approx \mathbf{y}(t)$ uniformly well over a large class of inputs $\mathbf{u}(t)$. Different measures of approximation and different choices of input classes lead to different model reduction goals. Indeed, the "size" of linear systems can be measured in various ways – the two most common metrics are the so-called $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms.

The $\mathcal{H}_\infty$ Norm The $\mathcal{H}_\infty$ norm of a stable linear system associated with a transfer function, $\mathbf{H}(s)$, is defined as

$$\|\mathbf{H}\|_{\mathcal{H}_\infty} = \max_{\omega\in\mathbb{R}} \|\mathbf{H}(\imath\omega)\|_2\,,$$

where $\|\mathbf{R}\|_2$ here denotes the usual induced 2-norm of a matrix $\mathbf{R}$. If the matrix $\mathbf{E}$ is singular, we must require in addition that 0 is a nondefective eigenvalue of $\mathbf{E}$ so that $\mathbf{H}(s)$ is bounded as $s\to\infty$. Suppose one wants to ensure that the output error $\mathbf{y}(t)-\mathbf{y}_r(t)$ is small in a root mean square sense for $t>0$ (that is, we want $(\int_0^\infty \|\mathbf{y}(t)-\mathbf{y}_r(t)\|_2\,dt)^{1/2}$ to be small) uniformly over all inputs, $\mathbf{u}(t)$, having bounded "energy," $\int_0^\infty \|\mathbf{u}(t)\|_2^2\,dt\le 1$. Observe first that $\widehat{\mathbf{y}}(s)-\widehat{\mathbf{y}}_r(s) = [\mathbf{H}(s)-\mathbf{H}_r(s)]\,\widehat{\mathbf{u}}(s)$ so then by the Parseval relation:

$$\begin{aligned}\int_0^\infty \|\mathbf{y}(t)-\mathbf{y}_r(t)\|_2^2 dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\,\widehat{\mathbf{y}}(\imath\omega)-\widehat{\mathbf{y}}_r(\imath\omega)\|_2^2\;d\omega\\ &\le \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\mathbf{H}(\imath\omega)-\mathbf{H}_r(\imath\omega)\|_2^2\;\|\widehat{\mathbf{u}}(\imath\omega)\|_2^2\,d\omega\\ &\le \max_\omega \|\mathbf{H}(\imath\omega)-\mathbf{H}_r(\imath\omega)\|_2^2 \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\|\widehat{\mathbf{u}}(\imath\omega)\|_2^2\,d\omega\right)^{1/2}\\ &\le \max_\omega \|\mathbf{H}(\imath\omega)-\mathbf{H}_r(\imath\omega)\|_2^2 \;\stackrel{\text{def}}{=}\; \|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_\infty}^2\end{aligned}$$

Hence, proving that the $\|\mathbf{H}\|_{\mathcal{H}_\infty}$ norm is the L^2 induced operator norm of the associated system mapping, $\mathcal{S}:\mathbf{u}\mapsto\mathbf{y}$.

The $\mathcal{H}_2$ Norm. The $\mathcal{H}_2$ norm of a linear system associated with a transfer function, $\mathbf{H}(s)$, is defined as

$$\|\mathbf{H}\|_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\|\mathbf{H}(\imath\omega)\|_{\mathrm{F}}^2\right)^{1/2} \tag{17}$$

where now $\|\mathbf{R}\|_{\mathrm{F}}^2 = \mathit{trace}(\overline{\mathbf{R}}\,\mathbf{R}^T)$ denotes the Frobenius norm of a complex matrix $\mathbf{R}$. Note that we must require in addition that if $\mathbf{E}$ is singular then 0 is a nondefective eigenvalue of $\mathbf{E}$ and that $\lim_{s\to\infty}\mathbf{H}(s)=0$ in order that the $\mathcal{H}_2$ norm of the system to be finite. For more details, see [2] and [94].

Alternatively, suppose one wants to ensure that each component of the output error $\mathbf{y}(t)-\mathbf{y}_r(t)$ remains small for all $t>0$ (that is, we want $\max_{t>0}\|\mathbf{y}(t)-\mathbf{y}_r(t)\|_\infty$ to be small) again uniformly over all inputs, $\mathbf{u}(t)$ with $\int_0^\infty \|\mathbf{u}(t)\|_2^2\,dt\le 1$.

$$\begin{aligned}\max_{t>0}\|\mathbf{y}(t)-\mathbf{y}_r(t)\|_\infty &= \max_{t>0}\left\|\frac{1}{2\pi}\int_{-\infty}^{\infty}(\widehat{\mathbf{y}}(\imath\omega)-\widehat{\mathbf{y}}_r(\imath\omega))\,e^{\imath\omega t}\,d\omega\right\|_\infty\\ &\le \frac{1}{2\pi}\int_{-\infty}^{\infty}\|\widehat{\mathbf{y}}(\imath\omega)-\widehat{\mathbf{y}}_r(\imath\omega)\|_\infty\;d\omega\\ &\le \frac{1}{2\pi}\int_{-\infty}^{\infty}\|\mathbf{H}(\imath\omega)-\mathbf{H}_r(\imath\omega)\|_F\;\|\widehat{\mathbf{u}}(\imath\omega)\|_2\,d\omega\end{aligned}$$

$$\leq \left(\int_{-\infty}^{\infty} \|\mathbf{H}(\imath\omega) - \mathbf{H}_r(\imath\omega)\|_F^2 \; d\omega \right)^{1/2} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \|\widehat{\mathbf{u}}(\imath\omega)\|_2^2 \, d\omega \right)^{1/2}$$

$$\leq \left(\int_{-\infty}^{+\infty} \|\mathbf{H}(\imath\omega) - \mathbf{H}_r(\imath\omega)\|_F^2 \, d\omega \right)^{1/2} \stackrel{\text{def}}{=} \|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2}$$

It should be noted that in the single-input single-output case, the above relation holds with equality sign, because the $\mathcal{H}_2$ norm is equal to the $(2,\infty)$ induced norm of the convolution operator (see [2] for details).

$\mathcal{H}_2$ denotes the set of matrix-valued functions, $\mathbf{G}(z)$, with components that are analytic for z in the open right half plane, $Re(z) > 0$, such that for $Re(z) = x > 0$, $\mathbf{G}(x+\imath y)$ is square integrable as a function of $y \in (-\infty, \infty)$:

$$\sup_{x>0} \int_{-\infty}^{\infty} \|\mathbf{G}(x+\imath y)\|_F^2 \; dy < \infty.$$

$\mathcal{H}_2$ is a Hilbert space and holds our interest because transfer functions associated with stable finite dimensional dynamical systems are elements of $\mathcal{H}_2$. Indeed, if $\mathbf{G}(s)$ and $\mathbf{H}(s)$ are transfer functions associated with stable dynamical systems having the same input and output dimensions, the $\mathcal{H}_2$-inner product can be defined as

$$\langle \mathbf{G},\, \mathbf{H} \rangle_{\mathcal{H}_2} \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathsf{Tr}\left(\overline{\mathbf{G}(\imath\omega)} \mathbf{H}(\imath\omega)^T \right) d\omega = \int_{-\infty}^{\infty} \mathsf{Tr}\left(\overline{\mathbf{G}}(-\imath\omega) \mathbf{H}(\imath\omega)^T \right) d\omega, \tag{18}$$

with a norm defined as

$$\|\mathbf{G}\|_{\mathcal{H}_2} \stackrel{\text{def}}{=} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} \|\mathbf{G}(\imath\omega)\|_F^2 \, d\omega \right)^{1/2}. \tag{19}$$

where $\mathsf{Tr}(\mathbf{M})$ and $\|\mathbf{M}\|_F$ denote the trace and Frobenius norm of $\mathbf{M}$, respectively. Notice in particular that if $\mathbf{G}(s)$ and $\mathbf{H}(s)$ represent real dynamical systems then $\langle \mathbf{G},\, \mathbf{H} \rangle_{\mathcal{H}_2} = \langle \mathbf{H},\, \mathbf{G} \rangle_{\mathcal{H}_2}$ so that $\langle \mathbf{G},\, \mathbf{H} \rangle_{\mathcal{H}_2}$ must be real.

Suppose $f(s)$ is a meromorphic function (analytic everywhere except at isolated poles of finite order). λ is a simple pole of $f(s)$ if $\lim_{s\to\lambda}(s-\lambda)^\ell f(s) = 0$ for $\ell \geq 2$ and the residue is nontrivial: $\mathsf{res}[f(s),\lambda] = \lim_{s\to\lambda}(s-\lambda)f(s) \neq 0$. For matrix-valued meromorphic functions, $\mathbf{F}(s)$, we say that λ is a simple pole of $\mathbf{F}(s)$ if $\lim_{s\to\lambda}(s-\lambda)^\ell \mathbf{F}(s) = 0$ for $\ell \geq 2$ and $\mathsf{res}[\mathbf{F}(s),\lambda] = \lim_{s\to\lambda}(s-\lambda)f(s)$ has rank 1. λ is a semi-simple pole of $\mathbf{F}(s)$ if $\lim_{s\to\lambda}(s-\lambda)^\ell \mathbf{F}(s) = 0$ for $\ell \geq 2$ and $\mathsf{res}[\mathbf{F}(s),\lambda] = \lim_{s\to\lambda}(s-\lambda)\mathbf{F}(s)$ has rank larger than 1.

If $\|\mathbf{F}(s)\|_2$ remains bounded as $s \to \infty$ then $\mathbf{F}(s)$ only has a finite number of poles. If all these poles are either simple or semi-simple then we define the *order* or *dimension* of $\mathbf{F}(s)$ by $\dim \mathbf{F} = \sum_\lambda \mathrm{rank}\,(\mathsf{res}[\mathbf{F}(s),\lambda])$ where the sum is taken over all poles λ. In this case, we can represent $\mathbf{F}(s)$ as

$$\mathbf{F}(s) = \sum_{i=1}^{\dim \mathbf{F}} \frac{1}{s-\lambda_i} \mathsf{c}_i \mathsf{b}_i^T,$$

where λ_i are indexed according to multiplicity as indicated by *rank*$(\mathsf{res}[\mathbf{F}(s),\lambda_i])$.

Lemma 1.1. *Suppose that* $\mathbf{G}(s)$ *and* $\mathbf{H}(s)$ *are stable (poles contained in the open left halfplane) and suppose that* $\mathbf{H}(s)$ *has poles at* $\mu_1, \mu_2, \ldots \mu_m$. *Then*

$$\langle \mathbf{G},\, \mathbf{H} \rangle_{\mathcal{H}_2} = \sum_{k=1}^{m} \mathsf{res}[\mathsf{Tr}\big(\overline{\mathbf{G}}(-s)\mathbf{H}(s)^T\big)\,, \mu_k]. \tag{20}$$

In particular, if $\mathbf{H}(s)$ *has only simple or semi-simple poles at* $\mu_1, \mu_2, \ldots \mu_m$ *and* $m = \dim \mathbf{H}$ *then* $\mathbf{H}(s) = \sum_{i=1}^{m} \frac{1}{s-\mu_i} \mathsf{c}_i \mathsf{b}_i^T$ *and*

$$\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2} = \sum_{k=1}^{m} \mathsf{c}_k^T \overline{\mathbf{G}}(-\mu_k) \mathsf{b}_k$$

and

$$\|\mathbf{H}\|_{\mathcal{H}_2} = \left(\sum_{k=1}^{m} \mathsf{c}_k^T \overline{\mathbf{H}}(-\mu_k) \mathsf{b}_k \right)^{1/2}.$$

Proof. Notice that the function $\mathsf{Tr}\big(\overline{\mathbf{G}}(-s)\mathbf{H}(s)^T\big)$ has singularities in the left half plane only at $\mu_1, \mu_2, \ldots \mu_m$. For any $R > 0$, define the semicircular contour in the left halfplane:

$$\Gamma_R = \{z \,|\, z = \imath\omega \text{ with } \omega \in [-R, R]\,\} \cup \left\{ z \,\middle|\, z = R e^{\imath\theta} \text{ with } \theta \in [\frac{\pi}{2}, \frac{3\pi}{2}] \right\}.$$

Γ_R bounds a region that for sufficiently large R contains all the system poles of $\mathbf{H}(s)$ and so, by the residue theorem

$$\begin{aligned} \langle \mathbf{G},\, \mathbf{H} \rangle_{\mathcal{H}_2} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathsf{Tr}\big(\overline{\mathbf{G}}(-\imath\omega)\mathbf{H}(\imath\omega)^T\big)\, d\omega \\ &= \lim_{R\to\infty} \frac{1}{2\pi\imath} \int_{\Gamma_R} \mathsf{Tr}\big(\overline{\mathbf{G}}(-s)\mathbf{H}(s)^T\big)\, ds \\ &= \sum_{k=1}^{m} \mathsf{res}[\mathsf{Tr}\big(\overline{\mathbf{G}}(-s)\mathbf{H}(s)^T\big)\,, \mu_k]. \end{aligned}$$

The remaining assertions follow from the definition. □

Lemma 1.1 immediately yields a new expression, recently introduced by Beattie and Gugercin in [17], for the $\mathcal{H}_2$ error norm for MIMO dynamical systems based on the poles and residues of the transfer function. A similar expression for SISO systems was first introduced by Krajewski et al. [58], and rediscovered later by Gugercin and Antoulas [2, 44, 46]:

Theorem 1.4. *Given a full-order real system,* $\mathbf{H}(s)$ *and a reduced model,* $\mathbf{H}_r(s)$, *having the form* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i} \mathsf{c}_i \mathsf{b}_i{}^T$ ($\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$ *and rank-1 residues* $\mathsf{c}_1\mathsf{b}_1^T, \ldots, \mathsf{c}_r\mathsf{b}_r^T$.), *the* $\mathcal{H}_2$ *norm of the error system is given by*

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2}^2 = \|\mathbf{H}\|_{\mathcal{H}_2}^2 - 2\sum_{k=1}^{r} \mathsf{c}_k^T \mathbf{H}(-\hat{\lambda}_k)\mathsf{b}_k + \sum_{k,\ell=1}^{r} \frac{\mathsf{c}_k^T \mathsf{c}_\ell\, \mathsf{b}_\ell^T \mathsf{b}_k}{-\hat{\lambda}_k - \hat{\lambda}_\ell} \tag{21}$$

3 Interpolatory Optimal $\mathcal{H}_2$ Approximation

Given a full-order system $\mathbf{H}(s)$, the optimal $\mathcal{H}_2$ model reduction problem seeks to find a reduced-order model $\mathbf{H}_r(s)$ that minimizes the $\mathcal{H}_2$ error; i.e.

$$\|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_2} = \min_{\substack{\dim(\widetilde{\mathbf{H}}_r)=r \\ \widetilde{\mathbf{H}}_r:\text{ stable}}} \left\|\mathbf{H}-\widetilde{\mathbf{H}}_r\right\|_{\mathcal{H}_2}. \tag{22}$$

This optimization problem is nonconvex – obtaining a global minimizer can be intractable and is at best a hard task. The common approach is to find reduced order models that satisfy first-order necessary optimality conditions. Many researchers have worked on this problem. These efforts can be grouped into two categories: Lyapunov-based optimal $\mathcal{H}_2$ methods such as [53, 54, 81, 88, 89, 95]; and interpolation-based optimal $\mathcal{H}_2$ methods such as [15,17,25,45,48,49,59,68,83]. The main drawback of most of the Lyapunov-based methods is that they require solving a series of Lyapunov equations and so rapidly become infeasible as dimension increases. We note that even though recent developments have made approximate solution of large order Lyapunov equations possible, solving a sequence of them through an iteration remains daunting and the effect of only having approximate solutions throughout the iteration has not yet been assessed. On the other hand, interpolatory approaches can take advantage of the sparsity structure of $\mathbf{E}$ and $\mathbf{A}$ and have proved numerically very effective. Moreover, both problem frameworks are equivalent, as shown by Gugercin et al. [49], and this has further motivated expansion of interpolatory approaches to optimal $\mathcal{H}_2$ approximation.

Interpolation-based $\mathcal{H}_2$-optimality conditions were originally developed by Meier and Luenberger [68] for SISO systems. Based on these conditions, an effective numerical algorithm for interpolatory optimal $\mathcal{H}_2$ approximation, called the **Iterative Rational Krylov Algorithm** (IRKA), was introduced in [45, 48]. Analogous $\mathcal{H}_2$-optimality conditions for MIMO systems within a tangential interpolation framework have recently been developed by [25, 49, 83] leading to an analogous algorithm for the MIMO case; descriptions can be found in [25, 49]. Below, we present a new proof of the interpolatory first-order conditions for the optimal $\mathcal{H}_2$ model reduction problem. These conditions are the starting point for the numerical algorithm presented in the next section.

Theorem 1.5. *Suppose* $\mathbf{H}(s)$ *is a real stable dynamical system and that* $\mathbf{H}_r(s) = \sum_{i=1}^{r} \frac{1}{s-\hat{\lambda}_i}\mathsf{c}_i\mathsf{b}_i^T$ *is a real dynamical system that is the best stable* rth *order approximation of* $\mathbf{H}$ *with respect to the* $\mathcal{H}_2$ *norm. (*$\mathbf{H}_r$ *has simple poles at* $\hat{\lambda}_1, \hat{\lambda}_2, \ldots \hat{\lambda}_r$ *and rank-1 residues* $\mathsf{c}_1\mathsf{b}_1^T, \ldots, \mathsf{c}_r\mathsf{b}_r^T$*.) Then*

$$(a)\ \mathbf{H}(-\hat{\lambda}_k)\mathsf{b}_k = \mathbf{H}_r(-\hat{\lambda}_k)\mathsf{b}_k, \quad (b)\ \mathsf{c}_k^T\mathbf{H}(-\hat{\lambda}_k) = \mathsf{c}_k^T\mathbf{H}_r(-\hat{\lambda}_k), \tag{23}$$

$$\text{and} \quad (c)\ \mathsf{c}_k^T\mathbf{H}'(-\hat{\lambda}_k)\mathsf{b}_k = \mathsf{c}_k^T\mathbf{H}_r'(-\hat{\lambda}_k)\mathsf{b}_k \quad \text{for } k = 1, 2, ..., r.$$

Proof. Suppose $\widetilde{\mathbf{H}}_r(s)$ is a transfer function associated with a stable rth order dynamical system. Then

$$\|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_2}^2 \le \|\mathbf{H}-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 = \|\mathbf{H}-\mathbf{H}_r+\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2$$

$$= \|\mathbf{H}-\mathbf{H}_r\|_{\mathcal{H}_2}^2 + 2\mathcal{R}e\,\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} + \|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2$$

$$\text{so that} \qquad 0 \le 2\mathcal{R}e\,\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} + \|\mathbf{H}_r-\widetilde{\mathbf{H}}_r\|_{\mathcal{H}_2}^2 \tag{24}$$

By making judicious choices in how $\widetilde{\mathbf{H}}_r$ is made to differ from $\mathbf{H}_r$, we arrive at tangential interpolation conditions via Lemma 1.1.

Toward this end, pick an arbitrary unit vector $\xi \in \mathbb{C}^m$, $\varepsilon > 0$, and for some ℓ, define $\theta = \pi - \arg \xi^T \left(\mathbf{H}(-\hat{\lambda}_\ell) - \mathbf{H}_r(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell$, and so that

$$\mathbf{H}_r(s) - \widetilde{\mathbf{H}}_r(s) = \frac{\varepsilon e^{\imath\theta}}{s-\hat{\lambda}_\ell}\xi \mathsf{b}_\ell^T$$

and using Lemma 1.1, $\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} = -\varepsilon\,|\xi^T \left(\mathbf{H}(-\hat{\lambda}_\ell) - \mathbf{H}_r(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell|$.

Now (24) leads to

$$0 \le |\xi^T \left(\mathbf{H}(-\hat{\lambda}_\ell) - \mathbf{H}_r(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell| \le \varepsilon \frac{\|\mathsf{b}_\ell\|_2^2}{-2\mathbb{R}e(\hat{\lambda}_\ell)}$$

which by taking ε small implies first that

$$\xi^T \left(\mathbf{H}(-\hat{\lambda}_\ell) - \mathbf{H}_r(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell = 0$$

but then since ξ was chosen arbitrarily, we must have that

$$\left(\mathbf{H}(-\hat{\lambda}_\ell) - \mathbf{H}_r(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell = 0.$$

A similar argument yields (23b).

If (23c) did not hold then $\mathsf{c}_\ell^T \mathbf{H}'(-\hat{\lambda}_\ell)\mathsf{b}_\ell \neq \mathsf{c}_\ell^T \mathbf{H}_r'(-\hat{\lambda}_\ell)\mathsf{b}_\ell$ and we may pick $0 < \varepsilon < |\mathcal{R}e(\hat{\lambda}_\ell)|$ and $\theta = -\arg \mathsf{c}_\ell^T \left(\mathbf{H}'(-\hat{\lambda}_\ell) - \mathbf{H}_r'(-\hat{\lambda}_\ell)\right) \mathsf{b}_\ell$ in such a way that $\mu = \hat{\lambda}_\ell + \varepsilon e^{\imath\theta}$ does not coincide with any reduced order poles $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_r$. Note that $\mathcal{R}e(\mu) < 0$. Define $\widetilde{\mathbf{H}}_r(s)$ so that

$$\mathbf{H}_r(s) - \widetilde{\mathbf{H}}_r(s) = \left(\frac{1}{s-\hat{\lambda}_\ell} - \frac{1}{s-\mu}\right) \mathsf{c}_\ell \mathsf{b}_\ell^T.$$

$\widetilde{\mathbf{H}}_r(s)$ has the same poles and residues as $\mathbf{H}_r(s)$ aside from μ which replaces $\hat{\lambda}_\ell$ as a pole in $\widetilde{\mathbf{H}}_r(s)$ with no change in the associated residue. Because of (23a-b), we calculate

$$\begin{aligned}\langle \mathbf{H}-\mathbf{H}_r, \mathbf{H}_r-\widetilde{\mathbf{H}}_r\rangle_{\mathcal{H}_2} &= \mathsf{c}_\ell^T\left(\mathbf{H}(-\hat{\lambda}_\ell)-\mathbf{H}_r(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell - \mathsf{c}_\ell^T\left(\mathbf{H}(-\mu)-\mathbf{H}_r(-\mu)\right)\mathsf{b}_\ell\\ &= -\mathsf{c}_\ell^T\left(\mathbf{H}(-\mu)-\mathbf{H}_r(-\mu)\right)\mathsf{b}_\ell\end{aligned}$$

Then (24) leads to

$$0 \le -2\mathcal{R}e(\mathsf{c}_\ell^T\left(\mathbf{H}(-\mu)-\mathbf{H}_r(-\mu)\right)\mathsf{b}_\ell) - \frac{|\mu-\hat{\lambda}_\ell|^2}{|\overline{\mu}+\hat{\lambda}_\ell|^2}\frac{\mathcal{R}e(\mu+\hat{\lambda}_\ell)}{2\mathcal{R}e(\hat{\lambda}_\ell)\mathcal{R}e(\mu)}\|\mathsf{c}_\ell\|_2^2\|\mathsf{b}_\ell\|_2^2 \tag{25}$$

Now, easy manipulations yield first a resolvent identity

$$(-\mu\mathbf{E}-\mathbf{A})^{-1} = (-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1} + (\mu-\hat{\lambda}_\ell)(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\mu\mathbf{E}-\mathbf{A})^{-1}$$

and then resubstituting,

$$\begin{aligned}(-\mu\mathbf{E}-\mathbf{A})^{-1} =& (-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1} + (\mu-\hat{\lambda}_\ell)(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\\ &+(\mu-\hat{\lambda}_\ell)^2(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\mu\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\end{aligned} \tag{26}$$

Premultiplying by $\mathbf{C}$ and postmultiplying by $\mathbf{B}$, (26) implies

$$\begin{aligned}\mathbf{H}(-\mu) =& \mathbf{H}(-\hat{\lambda}_\ell) + (\mu-\hat{\lambda}_\ell)\mathbf{H}'(-\hat{\lambda}_\ell)\\ &+(\mu-\hat{\lambda}_\ell)^2\mathbf{C}(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\mu\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(-\hat{\lambda}_\ell\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\end{aligned}$$

and analogous arguments yield

$$\begin{aligned}&\mathbf{H}_r(-\mu) = \mathbf{H}_r(-\hat{\lambda}_\ell) + (\mu-\hat{\lambda}_\ell)\mathbf{H}_r'(-\hat{\lambda}_\ell)\\ &+(\mu-\hat{\lambda}_\ell)^2\mathbf{C}_r(-\hat{\lambda}_\ell\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{E}_r(-\mu\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{E}_r(-\hat{\lambda}_\ell\mathbf{E}_r-\mathbf{A}_r)^{-1}\mathbf{B}_r\end{aligned}$$

Using these expressions in (25) yields

$$0 \le -2\varepsilon|\mathsf{c}_\ell^T\left(\mathbf{H}'(-\hat{\lambda}_\ell)-\mathbf{H}_r'(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell| + \mathcal{O}(\varepsilon^2).$$

As $\varepsilon \to 0$ a contradiction occurs unless $|\mathsf{c}_\ell^T\left(\mathbf{H}'(-\hat{\lambda}_\ell)-\mathbf{H}_r'(-\hat{\lambda}_\ell)\right)\mathsf{b}_\ell| = 0$. □

One byproduct of the proof of Theorem 1.5 is that if $\mathbf{H}_r(s)$ satisfies the necessary conditions (23), then $\mathbf{H}_r(s)$ is guaranteed to be an *optimal* approximation of $\mathbf{H}(s)$ relative to the $\mathcal{H}_2$ norm among all reduced order systems having the same reduced system poles $\{\hat{\lambda}_i\}_{i=1}^r$. This follows from the observation that the set of all systems having the all same poles $\{\hat{\lambda}_i\}_{i=1}^r$ comprise a subspace of $\mathcal{H}_2$ and the projection theorem is both necessary and sufficient for $\mathbf{H}_r(s)$ to be a minimizer out of a subspace of candidate minimizers.

Conversely, the set of all stable rth order dynamical systems is not only not a subspace but it is not even convex. Indeed, the original problem (22) may have

multiple minimizers and there may be "local minimizers" that do not solve (22) yet do satisfy the first order conditions (23). To clarify, a reduced order system $\mathbf{H}_r$, is said to be a *local minimizer* for (22), if for all $\varepsilon > 0$ sufficiently small,

$$\|\mathbf{H} - \mathbf{H}_r\|_{\mathcal{H}_2} \leq \|\mathbf{H} - \widetilde{\mathbf{H}}_r^{(\varepsilon)}\|_{\mathcal{H}_2}, \tag{27}$$

for all stable dynamical systems, $\widetilde{\mathbf{H}}_r^{(\varepsilon)}$ such that both $\dim(\widetilde{\mathbf{H}}_r^{(\varepsilon)}) = r$ and $\|\mathbf{H}_r - \widetilde{\mathbf{H}}_r^{(\varepsilon)}\|_{\mathcal{H}_2} \leq C\,\varepsilon$ for some constant C.

As a practical matter, global minimizers that solve (22) are difficult to obtain reliably and with certainty; current approaches favor seeking reduced order models that satisfy local (first-order) necessary conditions for optimality and so may produce models that are only certain to be local minimizers. Even though such strategies do not guarantee global minimization, they often produce very effective reduced order models, nonetheless.

3.1 An Algorithm for Interpolatory Optimal $\mathcal{H}_2$ Model Reduction

Theorem 1.5 states that first-order conditions for $\mathcal{H}_2$ optimality are tangential interpolation conditions at mirror images of reduced-order poles. However, one cannot simply construct the corresponding matrices $\mathbf{V}_r$ and $\mathbf{W}_r$ since the interpolation points and the tangential directions depend on the reduced-model; hence are not known a priori. The *Iterative Rational Krylov Algorithm* (**IRKA**) introduced in [49] resolves this problem by iteratively correcting the interpolation points and the directions as outlined below.

Let $\mathbf{Y}^*$ and $\mathbf{X}$ denote, respectively, the left and right eigenvectors for $\lambda\mathbf{E}_r - \mathbf{A}_r$ so that $\mathbf{Y}^*\mathbf{A}_r\mathbf{X} = \mathrm{diag}(\hat{\lambda}_i)$ and $\mathbf{Y}^*\mathbf{E}_r\mathbf{X} = \mathbf{I}_r$. Denote the columns of $\mathbf{C}_r\mathbf{X}$ as $\hat{\mathsf{c}}_i$ and the rows of $\mathbf{Y}^*\mathbf{B}_r$ as $\hat{\mathsf{b}}_i^T$. Then in **IRKA** interpolation points used in the next step are chosen to be the mirror images of the current reduced order poles, i.e., the eigenvalues, $\lambda(\mathbf{A}_r, \mathbf{E}_r)$, of the pencil $\lambda\mathbf{E}_r - \mathbf{A}_r$ in the current step. The tangent directions are corrected in a similar way, using the residues of the previous reduced-ordered model until (23) is satisfied. A brief sketch of **IRKA** is described in Algorithm 1:

Upon convergence, $\{\sigma_i\}$ will be the mirror images of the eigenvalues of the reduced pencil $\lambda\mathbf{E}_r - \mathbf{A}_r$; and $\hat{\mathsf{b}}_i^T = \mathbf{e}_i^T\mathbf{Y}^*\mathbf{B}_r$ and $\hat{\mathsf{c}}_i = \mathbf{C}_r\mathbf{X}\mathbf{e}_i$ for $i = 1, \ldots, r$. Consequently from Steps 4.(d) and 4.(e), the reduced-order transfer function satisfies (23), first-order conditions for $\mathcal{H}_2$ optimality. The main computational cost of this algorithm involves solving $2r$ linear systems to generate $\mathbf{V}_r$ and $\mathbf{W}_r$. Computing the eigenvectors $\mathbf{Y}$ and $\mathbf{X}$, and the eigenvalues of the reduced pencil $\lambda\mathbf{E}_r - \mathbf{A}_r$ are cheap since the dimension r is small.

IRKA has been remarkably successful in producing high fidelity reduced-order approximations; it is numerically effective and has been successfully applied to finding $\mathcal{H}_2$-optimal reduced models for systems of high order, $n > 160,000$, see [56].

MIMO $\mathcal{H}_2$ Optimal Tangential Interpolation Method

1. Make an initial r-fold shift selection: $\{\sigma_1,\ldots,\sigma_r\}$ that is closed under conjugation (i.e., $\{\sigma_1,\ldots,\sigma_r\}\equiv\{\overline{\sigma_1},\ldots,\overline{\sigma_r}\}$ viewed as sets)
   and initial tangent directions $\hat{b}_1,\ldots,\hat{b}_r$ and $\hat{c}_1,\ldots,\hat{c}_r$, also closed under conjugation.
2. $\mathbf{V}_r=\left[(\sigma_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\hat{b}_1\ \cdots\ (\sigma_r\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\hat{b}_r\right]$
   $\mathbf{W}_r=\left[(\sigma_1\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\hat{c}_1\cdots(\sigma_r\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\hat{c}_1\right]$.
3. while (not converged)
   a. $\mathbf{A}_r=\mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r=\mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r=\mathbf{W}_r^T\mathbf{B}$, and $\mathbf{C}_r=\mathbf{C}\mathbf{V}_r$
   b. Compute $\mathbf{Y}^*\mathbf{A}_r\mathbf{X}=diag(\tilde{\lambda}_i)$ and $\mathbf{Y}^*\mathbf{E}_r\mathbf{X}=\mathbf{I}_r$ where $\mathbf{Y}^*$ and $\mathbf{X}$ are the left and right eigenvectors of $\lambda\mathbf{E}_r-\mathbf{A}_r$.
   c. $\sigma_i\longleftarrow-\lambda_i(\mathbf{A}_r,\mathbf{E}_r)$ for $i=1,\ldots,r$, $\hat{b}_i^*\longleftarrow\mathbf{e}_i^T\mathbf{Y}^*\mathbf{B}_r$ and $\hat{c}_i\longleftarrow\mathbf{C}_r\mathbf{X}\mathbf{e}_i$.
   d. $\mathbf{V}_r=\left[(\sigma_1\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\hat{b}_1\ \cdots\ (\sigma_r\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\hat{b}_r\right]$
   e. $\mathbf{W}_r=\left[(\sigma_1\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\hat{c}_1\cdots(\sigma_r\mathbf{E}-\mathbf{A}^T)^{-1}\mathbf{C}^T\hat{c}_1\right]$.
4. $\mathbf{A}_r=\mathbf{W}_r^T\mathbf{A}\mathbf{V}_r$, $\mathbf{E}_r=\mathbf{W}_r^T\mathbf{E}\mathbf{V}_r$, $\mathbf{B}_r=\mathbf{W}_r^T\mathbf{B}$, $\mathbf{C}_r=\mathbf{C}\mathbf{V}_r$

Remark 1.1. Note that the $\mathcal{H}_2$ error formulae in Theorem 1.4 expresses the error as a function of reduced-order system poles and residues. Based on this expression, Beattie and Gugercin [17] recently developed a trust-region based descent algorithm where the poles and residues of $\mathbf{H}_r(s)$ are the optimization parameters as opposed to the interpolation points and tangential directions in **IRKA**. The algorithm reduces the $\mathcal{H}_2$ error at each step of the iteration and globally converges to a local minimum. Even though the method does not use the interpolation point as the parameters, it is worth noting that the resulting reduced-order model indeed is an interpolatory approximation satisfying the interpolatory $\mathcal{H}_2$ conditions of Theorem 1.5. Hence, the method of [17] achieves the interpolation conditions while using the reduced-order poles and residues as the variables.

3.2 Numerical Results for IRKA

This problem arises during a cooling process in a rolling mill and is modeled as boundary control of a two dimensional heat equation. A finite element discretization results in a descriptor system state-dimension $n=79,841$, i.e., $\mathbf{A},\mathbf{E}\in\mathbb{R}^{79841\times79841}$, $\mathbf{B}\in\mathbb{R}^{79841\times7}$, $\mathbf{C}\in\mathbb{R}^{6\times79841}$. For details regarding this model, see [18, 20].

Using **IRKA**, we reduce the order of the full-order system, $\mathbf{H}(s)$, to $r=20$ to obtain the $\mathcal{H}_2$ optimal reduced model, $\mathbf{H}_{\mathrm{IRKA}}(s)$. Figure 1 illustrates the convergence behavior of **IRKA** through the history of the relative $\mathcal{H}_\infty$ error, $\dfrac{\|\mathbf{H}-\mathbf{H}_{\mathrm{IRKA}}\|_{\mathcal{H}_\infty}}{\|\mathbf{H}\|_{\mathcal{H}_\infty}}$, as it evolves through the course of the iteration. Even though the figure shows the method converging after approximately 11-12 iterations, one may also see that it nearly achieves the final error value much earlier in the iteration; already around the sixth or seventh step.

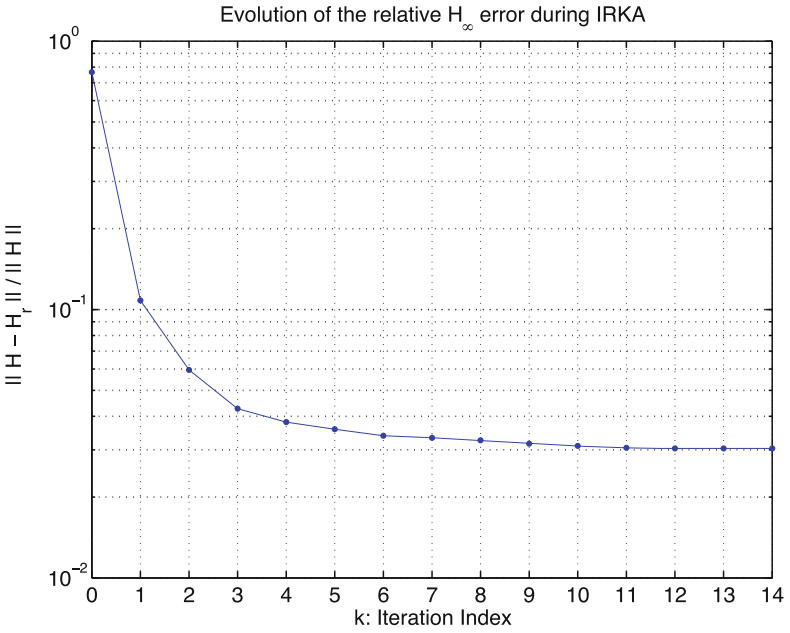


Fig. 1 Evolution of **IRKA**

Starting from a large initial relative error (nearly 100%), the method automatically and without any user intervention corrects both interpolation points and tangent directions, reaching within a few iterations a near-optimal solution with a relative error of around 3×10^{-2}.

We compare our approach with other commonly used model reduction techniques.

1. Modal Approximation: We reduce the order $r = 20$ using 20 dominant modes of $\mathbf{H}(s)$. The reduced model is denoted by $\mathbf{H}_{\text{modal}}$.
2. Interpolation using points on the imaginary axis: Based on the bode plot of $\mathbf{H}(s)$, we have placed interpolation points on the imaginary axis where $\|\mathbf{H}(\imath\omega)\|$ is dominant. This was a common approach for choosing interpolation points before the $\mathcal{H}_2$-optimal shift selection strategy was developed. The reduced model is denoted by $\mathbf{H}_{\imath\omega}$.
3. 20 real interpolation points are chosen as the mirror images of the poles of $\mathbf{H}(s)$. This selection is a good initialization for **IRKA** in the SISO case, see [49]. The reduced model is denoted by $\mathbf{H}_{\text{real}}$.

In our experiments, we found that in order to be able to find reasonable interpolation points and directions for $\mathbf{H}_{\imath\omega}$ and $\mathbf{H}_{\text{real}}$, we needed first to run several experiments. Many of them either resulted in unstable systems or very poor performance

Table 1 Error norms

	$\mathbf{H}_{\text{IRKA}}$	$\mathbf{H}_{\text{modal}}$	$\mathbf{H}_{\iota\omega}$	$\mathbf{H}_{\text{real}}$
Relative $\mathcal{H}_\infty$ error	3.03×10^{-2}	1.03×10^{-1}	5.42×10^{-1}	2.47×10^{-1}

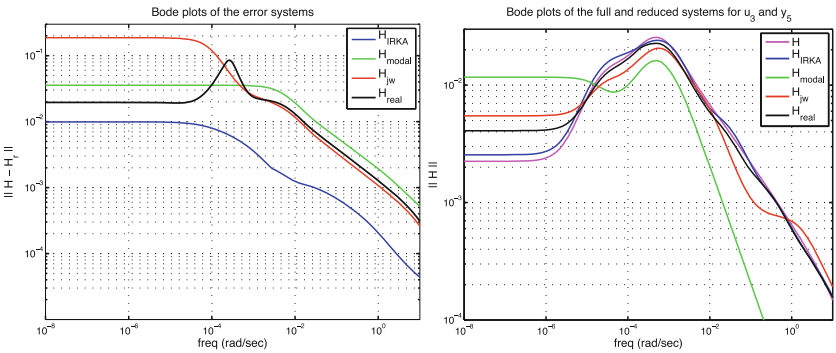


Fig. 2 Comparison of reduced-order models. (**a**) Error system bode plots. (**b**) Reduced model bode plots for $\mathbf{u}_3$ and $\mathbf{y}_5$

results. Here we are presenting the best selection we were able to find. This is the precise reason why **IRKA** is superior. We initiate it once, randomly in this case, and the algorithm automatically finds the optimal points and directions. There is no need for an ad hoc search. Table 1 shows the relative $\mathcal{H}_\infty$ error norms for each reduced model. Clearly, **IRKA** is the best one, yielding an error one order of magnitude smaller than those the other four approaches yield. Note from Fig. 1 that the initial guess for **IRKA** has a higher error than all the other methods. However, even after only two steps of the iteration long before convergence, the **IRKA** iterate has already a smaller error norm than all other three approaches. Note that $\mathbf{H}(s)$ has 7 inputs and 6 outputs; hence there are 42 input–output channels. $\mathbf{H}_{\text{IRKA}}$ with order $r = 20$, less than the total number of input-output channels, is able to replicate these behaviors with a relative accuracy of order 10^{-2}. Even though **IRKA** is an $\mathcal{H}_2$-based approach, superior $\mathcal{H}_\infty$ performance is also observed and is not surprising. It is an efficient general purpose $\mathcal{H}_2$ and $\mathcal{H}_\infty$ model reduction method, not only $\mathcal{H}_2$ optimal. Figure 2a depicts the bode plots of the error systems. It is clear from the error plots that $\mathbf{H}_{\text{IRKA}}$ outperforms the rest of the methods. To give an example how the reduced models match the full-order model for a specific input/output channel, we show the bode plots for the transfer function between the third input $\mathbf{u}_3$ and the fifth output $\mathbf{y}_5$. Clearly, $\mathbf{H}_{\text{IRKA}}$ yields the best match. We note that while Table 1 lists the relative error norms, the error norms in Fig. 2 shown are the absolute error values.

Since **IRKA** yielded a much lower error value, we have checked what lowest order model from **IRKA** would yield similar error norms as the other approaches. We have found that **IRKA** for order $r = 2$ yields a relative $\mathcal{H}_\infty$ error of 2.26×10^{-1}; already better than 20th order $\mathbf{H}_{\iota\omega}$ and $\mathbf{H}_{\text{real}}$. For $r = 6$, for example, **IRKA** yielded a relative $\mathcal{H}_\infty$ error of 1.64×10^{-1}, a number close to the one obtained by 20th order $\mathbf{H}_{\text{modal}}$. These numbers vividly illustrate the strength of the optimal $\mathcal{H}_2$ shift selection.

4 Interpolatory Passivity Preserving Model Reduction

Passive systems are an important class of dynamical systems that can only absorb power – for which the work done by the inputs in producing outputs must always be positive. A system is called *passive* if

$$\mathcal{R}e\int_{-\infty}^{t}\mathbf{u}(\tau)^{*}\mathbf{y}(\tau)\mathrm{d}\tau\geq 0,$$

for all $t\in\mathbb{R}$ and for all $\mathbf{u}\in\mathcal{L}_2(\mathbb{R})$. In this section, we will show how to create interpolatory projections so that passivity of reduced models is preserved. For further details we refer to [3].

A rational (square) matrix function $\mathbf{H}(s)$ is called *positive real* if:

1. $\mathbf{H}(s)$, is analytic for $\mathcal{R}e(s)>0$,
2. $\overline{\mathbf{H}(s)}=\mathbf{H}(\overline{s})$ for all $s\in\mathbb{C}$, and
3. the Hermitian part of $\mathbf{H}(s)$, i.e., $\mathbf{H}(s)+\mathbf{H}^{T}(\overline{s})$, is positive semi-definite for all $\mathcal{R}e(s)\geq 0$.

Dynamical systems as given in (1) are passive if and only if the associated transfer function $\mathbf{H}(s)=\mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}+\mathbf{D}$, is positive real. An important consequence of positive realness is the existence of a *spectral factorization*: a matrix function $\boldsymbol{\Phi}(s)$ for which $\mathbf{H}(s)+\mathbf{H}^{T}(-s)=\boldsymbol{\Phi}(s)\boldsymbol{\Phi}^{T}(-s)$, and the poles as well as the (finite) zeros of $\boldsymbol{\Phi}(s)$ are all stable. The *spectral zeros* of the system represented by $\mathbf{H}(s)$ are defined to be those values, λ, for which $\boldsymbol{\Phi}(\lambda)$ (and hence $\mathbf{H}(\lambda)+\mathbf{H}^{T}(-\lambda)$) loses rank.

Assume for simplicity that λ is a spectral zero with multiplicity one (so that nullity$(\boldsymbol{\Phi}(\lambda))=1$). Then there is a right spectral zero direction, $\mathbf{z}$, such that $(\mathbf{H}(\lambda)+\mathbf{H}^{T}(-\lambda))\mathbf{z}=\mathbf{0}$. Evidently, if $(\lambda,\,\mathbf{z})$ is a right spectral zero pair for the system represented by $\mathbf{H}(s)$, then $(-\overline{\lambda},\,\mathbf{z}^{*})$ is a left spectral zero pair: $\mathbf{z}^{*}(\mathbf{H}(-\overline{\lambda})+\mathbf{H}^{T}(\overline{\lambda}))=\mathbf{0}$.

The key result that is significant here was shown in [3] that if interpolation points are chosen as spectral zeros, passivity is preserved.

Theorem 1.6. *Suppose the dynamical system given in (1) and represented by the transfer function* $\mathbf{H}(s)=\mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}+\mathbf{D}$ *is stable and passive. Suppose that for some index* $r\geq 1$, $\lambda_1,\ldots,\lambda_r$ *are stable spectral zeros of* $\mathbf{H}$ *with corresponding right spectral zero directions* $\mathbf{z}_1,\ldots,\mathbf{z}_r$.

If a reduced order system $\mathbf{H}_r(s)$ *tangentially interpolates* $\mathbf{H}(s)$ *as in (11) with* $\sigma_i=\lambda_i$, $\mathsf{b}_i=\mathbf{z}_i$, $\mu_i=-\overline{\lambda_i}$, *and* $\mathsf{c}_i^{T}=\mathbf{z}^{*}$ *for* $i=1,\ldots,r$, *then* $\mathbf{H}_r(s)$ *is stable and passive.*

As a practical matter, the first task that we face is computation of the spectral zeros of the system represented by $\mathbf{H}(s)$. This can be formulated as a structured

eigenvalue problem. Following the development of [80], we define the following pair of matrices which has *Hamiltonian* structure:

$$\mathcal{H} = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & -\mathbf{A}^T & -\mathbf{C}^T \\ \mathbf{C} & \mathbf{B}^T & \mathbf{D}+\mathbf{D}^T \end{bmatrix} \text{ and } \mathcal{E} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

The *spectral zeros* of the system are the generalized eigenvalues of the pencil: $\mathcal{H}\mathbf{x}_i = \lambda_i \mathcal{E}\mathbf{x}_i$. To see this, partition the eigenvector $\mathbf{x}_i$ into components $\mathbf{v}_i$, $\overline{\mathbf{w}_i}$, and $\mathbf{z}_i$ such that

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & -\mathbf{A}^T & -\mathbf{C}^T \\ \mathbf{C} & \mathbf{B}^T & \mathbf{D}+\mathbf{D}^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_i \\ \overline{\mathbf{w}_i} \\ \mathbf{z}_i \end{bmatrix} = \lambda_i \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_i \\ \overline{\mathbf{w}_i} \\ \mathbf{z}_i \end{bmatrix},$$

Then note that as a consequence,

$$\mathbf{v}_i = (\lambda_i \mathbf{E} - \mathbf{A})^{-1}\mathbf{B}\mathbf{z}_i, \qquad \mathbf{w}_i^T = \mathbf{z}_i^* \mathbf{C}(-\overline{\lambda_i}\mathbf{E} - \mathbf{A})^{-1},$$
$$\text{and} \quad (\mathbf{H}(\lambda_i) + \mathbf{H}^T(-\lambda_i))\mathbf{z}_i = \mathbf{0}.$$

Thus, λ_i are spectral zeros of $\mathbf{H}(s)$ associated with the right spectral zero directions, $\mathbf{z}_i$, for $i = 1, \ldots, r$. Furthermore, basis vectors for the right and left modeling subspaces used to impose the tangential interpolation conditions required by Theorem 1.6, are determined immediately by the remaining two components of $\mathbf{x}_i$: $\mathbf{V}_r = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$ and $\mathbf{W}_r = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r]$.

Since $\mathcal{H}$ and $\mathcal{E}$ are real, the eigenvalues of $\mathcal{H}\mathbf{x} = \lambda\,\mathcal{E}\mathbf{x}$ occur in complex conjugate pairs. Furthermore, if we define $\mathbf{J} = \begin{bmatrix} \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}$ Then $(\mathbf{J}\mathcal{H})^* = \mathbf{J}\mathcal{H}$ and $(\mathbf{J}\mathcal{E})^* = -\mathbf{J}\mathcal{E}$ and it is easy to see that if $\mathbf{x}^T = [\mathbf{v}^T\ \overline{\mathbf{w}}^T\ \mathbf{z}^T]$ is a (right) eigenvector associated with λ: $\mathcal{H}\mathbf{x} = \lambda\,\mathcal{E}\mathbf{x}$, then $\mathbf{y}^* = [\mathbf{w}^T\ -\mathbf{v}^*\ \mathbf{z}^*]$ is a left eigenvector of the pencil associated with $-\overline{\lambda}$: $\mathbf{y}^*\mathcal{H} = -\overline{\lambda}\,\mathbf{y}^*\mathcal{E}$. Thus, if λ is an eigenvalue of $\mathcal{H}\mathbf{x} = \lambda\,\mathcal{E}\mathbf{x}$ then so is each of $\overline{\lambda}$, $-\overline{\lambda}$, and $-\lambda$.

Thus, the spectral zeros, associated spectral zero directions, and bases for the left and right modeling subspaces can be obtained by means of the above Hamiltonian eigenvalue problem.

The question remains of which r spectral zeros to choose. The concept of *dominance* arising in modal approximation is useful in distinguishing effective choices of spectral zero sets. For details we refer to [55].

Assume for simplicity that $\mathbf{D}+\mathbf{D}^T$ is invertible, take $\Delta = (\mathbf{D}+\mathbf{D}^T)^{-1}$ and define

$$\mathcal{B} = \begin{bmatrix} \mathbf{B} \\ -\mathbf{C}^T \\ \mathbf{0} \end{bmatrix} \Delta, \quad \mathcal{C} = -\Delta[\mathbf{C},\ \mathbf{B}^T,\ \mathbf{0}],$$

It can be checked (with some effort) that

$$\mathbf{G}(s) \stackrel{\Delta}{=} [\mathbf{H}(s) + \mathbf{H}^T(-s)]^{-1} = \Delta + \mathcal{C}(s\mathcal{E} - \mathcal{H})^{-1}\mathcal{B}.$$

Let the partial fraction expansion of $\mathbf{G}(s)$ be

$$\mathbf{G}(s) = \sum_{j=1}^{2n} \frac{\mathbf{R}_j}{s-\lambda_j}, \quad \text{with} \quad \mathbf{R}_j = \frac{1}{\mathbf{y}_j^* \mathcal{E} \mathbf{x}_j} \mathcal{C} \mathbf{x}_j \mathbf{y}_j^* \mathcal{B},$$

where λ_i are the spectral zeros of the original system (poles of the associated Hamiltonian system) and $\mathbf{R}_j$ are the *residues*. The left and right eigenvectors of $\mathbf{y}_j$, $\mathbf{x}_j$, are computed from $\mathcal{H}\mathbf{x}_j = \lambda_j \mathcal{E}\mathbf{x}_j$ and $\mathbf{y}_j^* \mathcal{H} = \lambda_j \mathbf{y}_j^* \mathcal{E}$.

A spectral zero λ_i is *dominant* over another spectral zero λ_j, if

$$\frac{\|\mathbf{R}_i\|_2}{|\mathcal{R}e(\lambda_i)|} > \frac{\|\mathbf{R}_j\|_2}{|\mathcal{R}e(\lambda_j)|}.$$

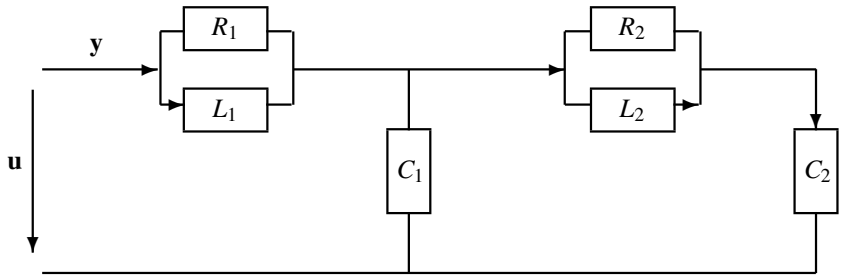
To efficiently compute the *r most dominant* spectral zeros of a dynamical system represented by $\mathbf{H}(s)$, we use an algorithm introduced in [77], and referred to as SADPA (Subspace Accelerated Dominant Pole Algorithm).

To summarize, the benefits of the spectral zero method are:

(1) The dominant spectral zeros λ_j can be computed automatically
(2) The resulting eigenvectors are readily available to construct projection matrices, $\mathbf{V}_r$ and $\mathbf{W}_r$, for the right and left modeling subspaces
(3) The dominance criterion is flexible and easily adjustable and
(4) Iterative methods are suitable for large scale applications For a detailed description of the resulting algorithm and all the missing details see [55].

4.1 An Example of Passivity Preserving Model Reduction

We conclude this section with a simple example which illustrates the concepts discussed above. Consider the following RLC circuit:



Using the voltages across C_1, C_2, and the currents through L_1, L_2, as state variables, $\mathbf{x}_i$, $= 1,2,3,4$, respectively, we end up with equations of the form $\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$, $\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, where

$$\mathbf{E} = \begin{pmatrix} C_1 & 0 & -G_1L_1 & G_2L_2 \\ 0 & C_2 & 0 & -G_2L_2 \\ 0 & 0 & L_1 & 0 \\ 0 & 0 & 0 & L_2 \end{pmatrix}, \ \mathbf{A} = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{C} = [-G_1, \ 0, \ 1, \ 0], \quad \mathbf{D} = G_1,$$

and $G_i = \frac{1}{R_i}$, $i = 1,2$, are the corresponding conductances. By construction, the system is passive (for non-negative values of the parameters), and it is easy to see that its transfer function has a zero at $s = 0$. Hence the system has a double spectral zero at $s = 0$. According to the definition of dominance mentioned above, among all finite spectral zeros, those on the imaginary axis are dominant. Hence we will compute a second order reduced system by using the the eigenpairs of $(\mathcal{H}, \ \mathcal{E})$, corresponding to the double zero eigenvalue. The Hamiltonian pair is:

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & G_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ -G_1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 2G_1 \end{pmatrix},$$

and

$$\mathcal{E} = \begin{pmatrix} C_1 & 0 & -G_1L_1 & G_2L_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & C_2 & 0 & -G_2L_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & L_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & L_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & C_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & C_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -G_1L_1 & 0 & L_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & G_2L_2 & -G_2L_2 & 0 & L_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

It follows that although the algebraic multiplicity of this eigenvalue is two, its geometric multiplicity is only one. Hence we need a Jordan chain of eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2$, corresponding to this eigenvalue. In particular, $\mathbf{x}_1$ satisfies $\mathcal{H}\mathbf{x}_1 = 0$, while $\mathbf{x}_2$, satisfies $\mathcal{H}\mathbf{x}_2 = \mathcal{E}\mathbf{x}_1$. These eigenvectors are:

$$[\mathbf{x}_1,\ \mathbf{x}_2] = \left(\begin{array}{c|c} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & \frac{C_2}{C_1+C_2} \\ \hline -1 & 0 \\ -1 & 0 \\ -G_1 & -1 \\ 0 & \frac{-C_2}{C_1+C_2} \\ \hline 1 & 0 \end{array}\right).$$

Thus the projection is defined by means of $\mathbf{V}_r = [\mathbf{x}_1(1:4,:),\ \mathbf{x}_2(1:4,:)]$ and $\mathbf{W}_r = -[\mathbf{x}_1(5:8,:),\ \mathbf{x}_2(5:8,:)]$, that is

$$\mathbf{V}_r = \left(\begin{array}{c|c} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & \frac{C_2}{C_1+C_2} \end{array}\right),\ \mathbf{W}_r = \left(\begin{array}{c|c} 1 & 0 \\ 1 & 0 \\ G_1 & 1 \\ 0 & \frac{C_2}{C_1+C_2} \end{array}\right).$$

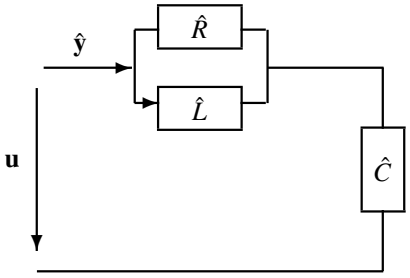
Therefore by (10) the reduced quantities are:

$$\mathbf{E}_r = \mathbf{W}_r^*\mathbf{E}\mathbf{V}_r = \begin{pmatrix} C_1+C_2 & 0 \\ 0 & L_1+\frac{{C_2}^2}{(C_1+C_2)^2}L_2 \end{pmatrix},\ \mathbf{A}_r = \mathbf{W}_r^*\mathbf{A}\mathbf{V}_r = \begin{pmatrix} -G_1 & 1 \\ -1 & 0 \end{pmatrix},$$
$$\mathbf{B}_r = \mathbf{W}_r^*\mathbf{B} = \begin{pmatrix} G_1 \\ 1 \end{pmatrix},\ \mathbf{C}_r = \mathbf{C}\mathbf{V}_r = \begin{pmatrix} -G_1 & 1 \end{pmatrix},\ \mathbf{D}_r = \mathbf{D}.$$

The corresponding transfer function $\mathbf{H}_r(s) = \mathbf{D} + \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r$, can be expressed as follows:

$$\mathbf{H}_r^{-1}(s) = \frac{1}{s(C_1+C_2)} + \frac{1}{G_1 + \dfrac{\kappa}{s}} \quad \text{where} \quad \kappa = \frac{L_1L_2(C_1+C_2)^2}{L_1{C_2}^2 + L_2(C_1+C_2)^2}.$$

From this expression or from the state space matrices, we can read-off an RLC realization. The reduced order circuit contains namely, a capacitor $\hat{C} = C_1 + C_2$, an inductor $\hat{L} = \frac{1}{\kappa} = L_1 + \frac{{C_2}^2}{(C_1+C_2)^2}L_2$, and a resistor of value $\hat{R} = R_1$. Thereby the capacitor is in series with a parallel connection of the inductor and the resistor as shown below.



Hence in this particular case, after reduction besides passivity, the structure (topology) of the circuit using the spectral zero reduction method, is preserved.

5 Structure-Preserving Model Reduction Using Generalized Coprime Factorizations

The model reduction framework we have described for reducing an original system of the form (1) to a reduced-order model of similar form (2) is quite general. With suitable generalizations and reformulations, this framework will encompass many linear dynamical systems of practical importance. However, there exist many linear dynamical systems whose natural descriptions take a form rather different than the standard canonical form (1).

We motivate the by considering a model of the dynamic response of a viscoelastic body. One possible continuum model for forced vibration of an isotropic incompressible viscoelastic solid could be described as

$$\begin{aligned} \partial_{tt}\mathbf{w}(x,t) - \eta\,\Delta\mathbf{w}(x,t) - \int_0^t \rho(t-\tau)\,\Delta\mathbf{w}(x,\tau)\,d\tau + \nabla\varpi(x,t) &= \mathbf{b}(x)\cdot\mathbf{u}(t), \\ \nabla\cdot\mathbf{w}(x,t) = 0 \text{ determining} \quad \mathbf{y}(t) &= [\mathbf{w}(\hat{x}_1,t),\ldots\mathbf{w}(\hat{x}_p,t)]^T, \end{aligned} \tag{28}$$

where $\mathbf{w}(x,t)$ is the displacement field for the body, $\varpi(x,t)$ is the associated pressure field and $\nabla\cdot\mathbf{w}=0$ represents the incompressibility constraint (see e.g., [63]). Here, $\rho(\tau)$ is a presumed known "relaxation function" satisfying $\rho(\tau)\geq 0$ and $\int_0^\infty \rho(\tau)\,d\tau<\infty$; $\eta>0$ is a known constant associated with the "initial elastic response". The term $\mathbf{b}(x)\cdot\mathbf{u}(t)=\sum_{i=1}^m b_i(x)\,u_i(t)$ is a superposition of the m inputs $\mathbf{u}(t)=\{u_1(t),\ldots,u_m(t)\}^T\in\mathbb{R}^m$. Displacements at $\hat{x}_1,\ldots,\hat{x}_p$ are the outputs. Semi-discretization of (28) with respect to space produces a large order linear dynamical system of the form:

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}}(t) + \eta\,\mathbf{K}\mathbf{x}(t) + \int_0^t \rho(t-\tau)\,\mathbf{K}\mathbf{x}(\tau)\,d\tau + \mathbf{D}\,\varpi(t) &= \mathbf{B}\mathbf{u}(t), \\ \mathbf{D}^T\mathbf{x}(t)=\mathbf{0}, \qquad \text{which determines} \quad \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t). \end{aligned} \tag{29}$$

where $\mathbf{x}\in\mathbb{R}^{n_1}$ is the discretization of the continuous displacement field $\mathbf{w}$; $\varpi\in\mathbb{R}^{n_2}$ is the discretization of the continuous pressure field ϖ. The matrices $\mathbf{M}$ and $\mathbf{K}$ are $n_1\times n_1$ real, symmetric, positive-definite matrices, $\mathbf{D}$ is an $n_1\times n_2$ matrix, $\mathbf{B}\in\mathbb{R}^{n_1\times m}$, and $\mathbf{C}\in\mathbb{R}^{p\times n_2}$. The state space dimension is $n=n_1+n_2$, an aggregate of $\mathbf{x}(t)$ and $\varpi(t)$.

Applying the Laplace transform to (29) yields

$$\widehat{\mathbf{y}}(s) = [\mathbf{C}\quad\mathbf{0}]\begin{bmatrix} s^2\mathbf{M}+(\widehat{\rho}(s)+\eta)\,\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{B}\\ \mathbf{0}\end{bmatrix}\widehat{\mathbf{u}}(s) \;=\; \mathcal{H}(s)\,\widehat{\mathbf{u}}(s), \tag{30}$$

defining the transfer function, $\mathcal{H}(s)$. The form of the transfer function is clearly different from the standard form (5). The system described in (29) is a *descriptor* system described by differential-algebraic equations with a hereditary damping; and a reformulation of (29) into the standard realization as in (1) is generally not possible (unless $\widehat{\rho}(s)$ has a simple form) and even when possible may not be desirable.

An effective reduced model for (29) should take into account the structure associated with distributed material properties. Therefore, we wish to consider reduced models similar to (29):

$$\begin{aligned} \mathbf{M}_r\,\ddot{\mathbf{x}}_r(t) + \eta\,\mathbf{K}_r\,\mathbf{x}_r(t) + \int_0^t \rho(t-\tau)\,\mathbf{K}_r\,\mathbf{x}_r(\tau)\,d\tau + \mathbf{D}_r\,\boldsymbol{\varpi}_r &= \mathbf{B}_r\,\mathbf{u}(t) \\ \mathbf{D}_r^T\,\mathbf{x}_r(t) = \mathbf{0} \qquad \text{which determines} \quad \mathbf{y}_r(t) &= \mathbf{C}_r\,\mathbf{x}_r(t) \end{aligned} \tag{31}$$

where $\mathbf{M}_r$ and $\mathbf{K}_r$ are now $r_1 \times r_1$ real, symmetric, positive-definite matrices, $\mathbf{D}_r$ is an $r_1 \times r_2$ matrix, $\mathbf{B}_r \in \mathbb{R}^{r_1 \times m}$ and $\mathbf{C}_r \in \mathbb{R}^{p \times r_2}$ with $r_1 \ll n_1$ and $r_2 \ll n_2$ and the reduced state space dimension is $r = r_1 + r_2$. Note that the reduced-model (31) has the same structure with the same hereditary damping as the original model. Systems with such structure may be expected to produce dynamic responses that are physically plausible and consistent with behaviors of viscoelastic bodies.

We proceed to construct reduced-order matrices in a similar way as we did in the standard case: We define matrices of trial vectors $\mathbf{U}_r \in \mathbb{R}^{n_1 \times r_1}$ and $\mathbf{Z}_r \in \mathbb{R}^{n_2 \times r_2}$; use the ansatz $\mathbf{x}(t) \approx \mathbf{U}_r\mathbf{x}_r(t)$ and $\boldsymbol{\varpi}(t) \approx \mathbf{Z}_r\boldsymbol{\varpi}_r(t)$; and force a Petrov-Galerkin orthogonality condition on the reduced state-space trajectory residuals to obtain finally reduced coefficient matrices for (31) as

$$\begin{aligned} \mathbf{M}_r = \mathbf{U}_r^T\mathbf{M}\mathbf{U}_r, \quad \mathbf{K}_r = \mathbf{U}_r^T\mathbf{K}\mathbf{U}_r, \quad \mathbf{D}_r = \mathbf{U}_r^T\mathbf{D}\mathbf{Z}_r, \\ \mathbf{B}_r = \mathbf{U}_r^T\mathbf{B}, \quad \text{and} \quad \mathbf{C}_r = \mathbf{C}\mathbf{U}_r. \end{aligned} \tag{32}$$

Note that this construction prevents mixing displacement state variables and pressure state variables. Also notice that both symmetry and positive-definiteness of $\mathbf{M}_r$ and $\mathbf{K}_r$ are preserved automatically as long as $\mathbf{U}_r$ has full column rank. The transfer function $\mathbf{H}_r(s)$ for the reduced-model can be obtained similarly as in (30). Now, in addition to preserving the structure we want to choose $\mathbf{U}_r$ and $\mathbf{Z}_r$ so that the reduced model $\mathbf{H}_r(s)$ interpolates $\mathbf{H}(s)$ as in the generic case. The ideas of Theorem 1.2 of §2.3 must be extended to this more general setting.

To proceed, we follow the discussion of Beattie and Gugercin [16] and consider MIMO systems with the following state space description in the Laplace transform domain (which includes the form of (30)):

$$\begin{aligned} \text{Find } \widehat{\mathbf{v}}(s) \text{ such that} \quad & \mathcal{K}(s)\,\widehat{\mathbf{v}}(s) = \mathcal{B}(s)\widehat{\mathbf{u}}(s) \\ \text{then} \quad & \widehat{\mathbf{y}}(s) \stackrel{\text{def}}{=} \mathcal{C}(s)\widehat{\mathbf{v}}(s) + \mathbf{D}\,\widehat{\mathbf{u}}(s) \\ \text{yielding a transfer function:} \quad & \mathbf{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s) + \mathbf{D}, \end{aligned} \tag{33}$$

where $\mathbf{D} \in \mathbb{R}^{p\times m}$, both $\mathcal{C}(s) \in \mathbb{C}^{p\times n}$ and $\mathcal{B}(s) \in \mathbb{C}^{n\times m}$ are analytic in the right half plane; and $\mathcal{K}(s) \in \mathbb{C}^{n\times n}$ is analytic and full rank throughout the right half plane. The goal is to find a reduced transfer function of the same form using a Petrov-Galerkin projection:

$$\mathbf{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s) + \mathbf{D}_r \tag{34}$$

where $\mathbf{W}_r, \mathbf{V}_r \in \mathbb{C}^{n\times r}$, $\mathcal{C}_r(s) = \mathcal{C}(s)\mathbf{V}_r \in \mathbb{C}^{p\times r}$; $\mathcal{B}_r(s) = \mathbf{W}_r^T\mathcal{B}(s) \in \mathbb{C}^{r\times m}$, and $\mathcal{K}_r(s) = \mathbf{W}_r^T\mathcal{K}(s)\mathbf{V}_r \in \mathbb{C}^{n\times n}$.

For simplicity we only consider the case $\mathbf{D} = \mathbf{D}_r$, which is equivalent with respect to the interpolation conditions to $\mathbf{D} = \mathbf{D}_r = \mathbf{0}$. The general case with $\mathbf{D} \neq \mathbf{D}_r$ (similar to Theorem 1.3) can be found in the original source [16]. Following the notation in [16], we write $\mathcal{D}_\sigma^\ell f$ to denote the ℓ^{th} derivative of the univariate function $f(s)$ evaluated at $s = \sigma$ with the usual convention for $\ell = 0$, $\mathcal{D}_\sigma^0 f = f(\sigma)$.

Theorem 1.7. *Suppose that $\mathcal{B}(s)$, $\mathcal{C}(s)$, and $\mathcal{K}(s)$ are analytic at $\sigma \in \mathbb{C}$ and $\mu \in \mathbb{C}$. Also let $\mathcal{K}(\sigma)$, $\mathcal{K}(\mu)$, $\mathcal{K}_r(\sigma) = \mathbf{W}_r^T\mathcal{K}(\sigma)\mathbf{V}_r$, and $\mathcal{K}_r(\mu) = \mathbf{W}_r^T\mathcal{K}(\mu)\mathbf{V}_r$ have full rank. Let nonnegative integers M and N be given as well as nontrivial vectors, $\mathsf{b} \in \mathbb{R}^m$ and $\mathsf{c} \in \mathbb{R}^p$.*

(a) If $\mathcal{D}_\sigma^i[\mathcal{K}(s)^{-1}\mathcal{B}(s)]\mathsf{b} \in \mathsf{Ran}(\mathbf{V}_r)$ for $i = 0,\ldots,N$
then $\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$ for $\ell = 0,\ldots,N$.
(b) If $\left(\mathsf{c}^T\mathcal{D}_\mu^j[\mathcal{C}(s)\mathcal{K}(s)^{-1}]\right)^T \in \mathsf{Ran}(\mathbf{W}_r)$ for $j = 0,\ldots,M$
then $\mathsf{c}^T\mathbf{H}^{(\ell)}(\mu) = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\mu)$ for $\ell = 0,\ldots,M$.
(c) If both (a) and (b) hold and if $\sigma = \mu$,
then $\mathsf{c}^T\mathbf{H}^{(\ell)}(\sigma)\mathsf{b} = \mathsf{c}^T\mathbf{H}_r^{(\ell)}(\sigma)\mathsf{b}$ for $\ell = 0,\ldots,M+N+1$.

Using this result, one can easily construct a reduced-model satisfying the desired interpolation conditions. For example, for $\mathbf{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)$, let r interpolation points $\{\sigma_i\}_{i=1}^r$, the left-directions $\{\mathsf{c}_i\}_{i=1}^r$ and the right-directions $\{\mathsf{b}_i\}_{i=1}^r$ be given. Construct

$$\mathbf{V}_r = \left[\mathcal{K}(\sigma_1)^{-1}\mathcal{B}(\sigma_1)\mathsf{b}_1, \ldots, \mathcal{K}(\sigma_r)^{-1}\mathcal{B}(\sigma_r)\mathsf{b}_r\right] \tag{35}$$

$$\text{and} \qquad \mathbf{W}_r = \left[\mathcal{K}(\sigma_1)^{-T}\mathcal{C}(\sigma_1)^T\mathsf{c}_1, \ldots, \mathcal{K}(\sigma_r)^{-T}\mathcal{C}(\sigma_r)^T\mathsf{c}_r\right]^T. \tag{36}$$

The reduced transfer function, $\mathbf{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s)$ satisfies bi-tangential interpolation conditions, i.e., $\mathbf{H}(\sigma_i)\mathsf{b}_i = \mathbf{H}_r(\sigma_i)\mathsf{b}_i$, $\mathsf{c}_i^T\mathbf{H}(\sigma_i) = \mathsf{c}_i^T\mathbf{H}_r(\sigma_i)$ and $\mathsf{c}_i^T\mathbf{H}'(\sigma_i)\mathsf{b}_i = \mathsf{c}_i^T\mathbf{H}_r'(\sigma_i)\mathsf{b}_i$ for $i = 1,\ldots,r$.

Theorem 1.7 achieves the desired result; it extends the interpolation theory of Theorem 1.2 from the canonical first-order setting to a general setting where we have a general coprime factorization, $\mathbf{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s) + \mathbf{D}$ and guarantees that the reduced transfer function will have a similarly structured coprime factorization: $\mathbf{H}_r(s) = \mathcal{C}_r(s)\mathcal{K}_r(s)^{-1}\mathcal{B}_r(s) + \mathbf{D}_r$.

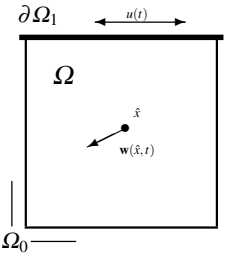
5.1 A Numerical Example: Driven Cavity Flow

We illustrate the concepts of Theorem 1.7 with a model of driven cavity flow in two dimensions, a slightly modified version of a model in (28). Consider a square domain $\Omega = [0,1]^2$, as shown below, representing a volume filled with a viscoelastic material with boundary separated into a movable top edge ("lid"), $\partial\Omega_1$, and the complement, $\partial\Omega_0$ (stationary bottom, left, and right edges). The material in the cavity is excited through shearing forces on the material caused by transverse displacement of the lid, $u(t)$. We are interested in the displacement response of the material, $\mathbf{w}(\hat{x},t)$, at the center of Ω, i.e. $\hat{x} = (0.5, 0.5)$.

$$\partial_{tt}\mathbf{w}(x,t) - \eta_0 \Delta \mathbf{w}(x,t) - \eta_1 \partial_t \int_0^t \frac{\Delta \mathbf{w}(x,\tau)}{(t-\tau)^\alpha}\, d\tau + \nabla \varpi(x,t) = 0 \quad \text{for } x \in \Omega$$

$$\begin{array}{c} \nabla \cdot \mathbf{w}(x,t) = 0 \\ \text{for } x \in \Omega \end{array} \quad \text{and} \quad \begin{array}{l} \mathbf{w}(x,t) = 0 \text{ for } x \in \partial\Omega_0 \\ \mathbf{w}(x,t) = u(t) \text{ for } x \in \partial\Omega_1 \end{array}$$

which determines the output $\mathbf{y}(t) = \mathbf{w}(\hat{x},t)$,



We use a model of a viscoelastic material given by Bagley and Torvik in [6] and take as material constants,

$$\eta_0 = 1.05 \times 10^6, \quad \eta_1 = 2.44 \times 10^5, \text{ and } \alpha = 0.519,$$

corresponding to experimentally derived values for the polymer butyl B252.

We discretize the continuous model using Taylor-Hood finite elements defined over a uniform triangular mesh on Ω. The transfer function of the discretized system is given by $\mathbf{H}(s) = \mathcal{C}(s)\mathcal{K}(s)^{-1}\mathcal{B}(s)$ where

$$\mathcal{K}(s) = \begin{bmatrix} s^2\mathbf{M} + \widehat{\rho}(s)\mathbf{K} & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix}, \quad \mathcal{C}(s) = [\mathbf{C}\ \mathbf{0}], \text{ and } \mathcal{B}(s) = \begin{bmatrix} s^2\mathbf{m} + \widehat{\rho(s)}\mathbf{k} \\ \mathbf{0} \end{bmatrix}.$$

$\mathbf{C} \in \mathbb{R}^{n\times 2}$ corresponds to measuring the horizontal and vertical displacement at $\hat{x} = (0.5, 0.5)$, $\mathbf{m}$ and $\mathbf{k}$ are sums of the columns of the free-free mass and stiffness matrix associated with x-displacement degrees of freedom on the top lid boundary and $\widehat{\rho}(s) = \eta_0 + \eta_1 s^\alpha$. This produces a frequency dependent input-to-state map, $\mathcal{B}(s)$, since the system input is a boundary *displacement* as opposed to a boundary *force*. Note the nonlinear frequency dependency in the state-to-input map; however Theorem 1.7 can be still applied to produce an interpolatory reduced-order model of the same form as shown next: Suppose an interpolation point $\sigma \in \mathbb{C}$ and a direction $\mathsf{c} \in \mathbb{C}^p$ are given. Since the input is a scalar, there is no need for a right tangential direction. We consider bitangential Hermite interpolating conditions:

$$\mathcal{H}_r(\sigma) = \mathcal{H}(\sigma),\ \mathsf{c}^T\mathcal{H}_r(\sigma) = \mathsf{c}^T\mathcal{H}(\sigma),\ \text{and } \mathsf{c}^T\mathcal{H}_r'(\sigma) = \mathsf{c}^T\mathcal{H}'(\sigma). \qquad (37)$$

Following Theorem 1.7, we solve the following two linear systems of equations:

$$\begin{bmatrix} \mathbf{F}(\sigma) & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{N}(\sigma) \\ \mathbf{0} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{F}(\sigma) & \mathbf{D} \\ \mathbf{D}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_2 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{C}^T\mathsf{c} \\ \mathbf{0} \end{bmatrix}.$$

where $\mathbf{F}(\sigma) = \sigma^2\mathbf{M} + \widehat{\rho}(\sigma)\mathbf{K}$ and $\mathbf{N}(\sigma) = s^2\mathbf{m} + \widehat{\rho}(\sigma)\mathbf{k}$. Define the matrices $\mathbf{U}_r = [\mathbf{u}_1, \mathbf{u}_2]$ and $\mathbf{Z}_r = [\mathbf{z}_1, \mathbf{z}_2]$. Then the reduced system matrices defined in (32) with the modification that $\mathcal{B}_r(s) = s^2\mathbf{U}_r^T\mathbf{m} + \widehat{\rho}(s)\mathbf{U}_r^T\mathbf{k}$ corresponds to the choice $\mathbf{V}_r = \mathbf{W}_r = \mathbf{U}_r \oplus \mathbf{Z}_r$ and satisfies the bitangential Hermite interpolation conditions of (37).

We compare three different models:

1. $\mathbf{H}_{\text{fine}}$, using a fine mesh Taylor-Hood FEM discretization with 51,842 displacement degrees of freedom and 6,651 pressure degrees of freedom (mesh size $h = \frac{1}{80}$). We treat this as the "full-order model".
2. $\mathbf{H}_{\text{coarse}}$, for a coarse mesh discretization with 29,282 displacement degrees of freedom and 3721 pressure degrees of freedom (mesh size $h = \frac{1}{60}$);
3. $\mathbf{H}_{30}$, a generalized interpolatory reduced order model as defined in (31)–(32) with $r = 30$, corresponding to 30 reduced displacement degrees of freedom and 30 reduced pressure degrees of freedom satisfying the bitangential Hermite interpolation conditions at each interpolation point.

The resulting frequency response plots shown in Figure 3:

In this example, this new framework allows interpolatory model reduction of a descriptor system with a hereditary damping term. Moreover, the reduced model is also a descriptor system with the same damping structure. The reduced model

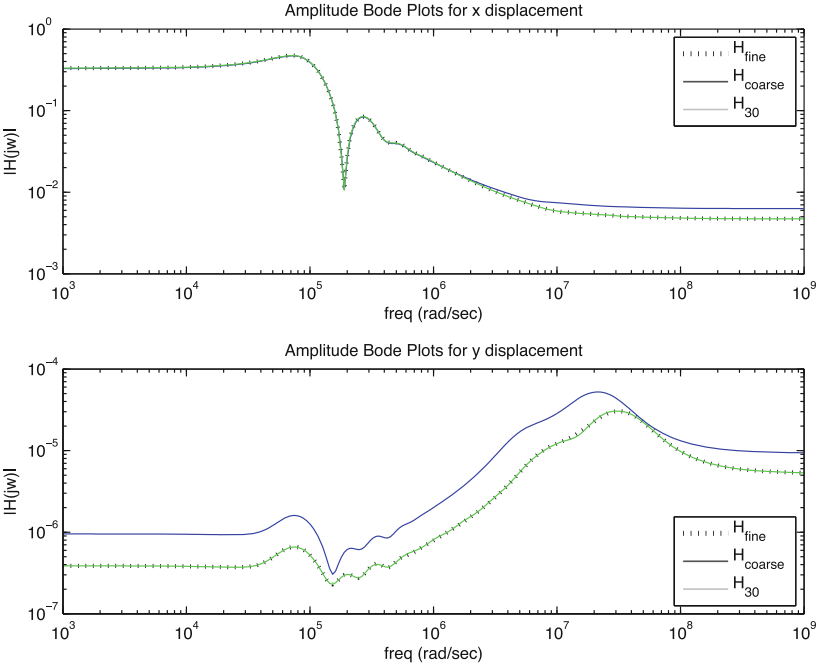


Fig. 3 Bode plots of $\mathbf{H}_{\text{fine}}$, $\mathbf{H}_{\text{coarse}}$ and reduced models $\mathbf{H}_{20}$ and $\mathbf{H}_{30}$

$\mathbf{H}_{30}$ having a total of 60 degrees of freedom has an input-output response that is virtually indistinguishable from the high fidelity model $\mathbf{H}_{\text{fine}}$ having nearly 58,500 total degrees of freedom. Moreover, $\mathbf{H}_{30}$ outperforms $\mathbf{H}_{\text{coarse}}$, which corresponds to 29,282 displacement and 3,721 pressure degrees of freedom (more than 500 times the order of $\mathbf{H}_{30}$).

5.2 Second-Order Dynamical Systems

One important variation on the dynamical system described in (33) arises in modeling the forced vibration of an n degree–of–freedom mechanical structure:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{G}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \tag{38}$$

$$\text{with the transfer function } \mathbf{H}(s) = \mathbf{C}(s^2\mathbf{M} + s\mathbf{G} + \mathbf{K})^{-1}\mathbf{B} \tag{39}$$

where $\mathbf{M}$, $\mathbf{G}$, and $\mathbf{K} \in \mathbb{R}^{n\times n}$ are positive (semi)-definite symmetric matrices describing, respectively, inertial mass distribution, energy dissipation, and elastic strain energy (stiffness) distribution throughout the structure. The input $\mathbf{u}(t) \in \mathbb{R}^m$ is a time-dependent force applied along degrees-of-freedom specified in $\mathbf{B} \in \mathbb{R}^{n\times m}$ and $\mathbf{y}(t) \in \mathbb{R}^p$ is a vector of output measurements defined through observation matrix $\mathbf{C} \in \mathbb{R}^p$. Second order systems of the form (38) arise naturally in analyzing other phenomena aside from structural vibration such as electrical circuits and micro-electro-mechanical systems, as well; see [7, 8, 26–28, 57, 75, 84], and references therein.

We wish to generate, for some $r \ll n$, an r^{th} order reduced *second-order system* of the same form:

$$\mathbf{M}_r\ddot{\mathbf{x}}_r(t) + \mathbf{G}_r\dot{\mathbf{x}}_r(t) + \mathbf{K}_r\mathbf{x}_r(t) = \mathbf{B}_r\mathbf{u}(t), \qquad \mathbf{y}_r(t) = \mathbf{C}_r\mathbf{x}_r(t),$$

where $\mathbf{M}_r, \mathbf{G}_r, \mathbf{K}_r \in \mathbb{R}^{r\times r}$ are positive (semi)-definite symmetric matrices, $\mathbf{B} \in \mathbb{R}^{r\times m}$, and $\mathbf{C} \in \mathbb{R}^{p\times r}$. In order to preserve the symmetry and positive definiteness of $\mathbf{M}$, $\mathbf{G}$ and $\mathbf{K}$ in the course of model reduction, one-sided reduction is applied, i.e. one takes $\mathbf{W}_r = \mathbf{V}_r$, resulting in

$$\mathbf{M}_r = \mathbf{V}_r^T\mathbf{M}\mathbf{V}_r,\ \mathbf{G}_r = \mathbf{V}_r^T\mathbf{G}\mathbf{V}_r,\ \mathbf{K}_r = \mathbf{V}_r^T\mathbf{K}\mathbf{V}_r, \mathbf{B}_r = \mathbf{V}_r^T\mathbf{B},,\ \mathbf{C}_r = \mathbf{C}\mathbf{V}_r.$$

One standard approach to model reduction of second-order systems involves the application of standard (first-order) model reduction techniques to a standard first-order realization of the system. However, this destroys the original second-order system structure and obscures the physical meaning of the reduced-order states. Once the reduction is performed in the first-order framework, it will not always be possible to convert this back to a corresponding second-order system of the form (38), see [69]. Even when this is possible, one typically cannot guarantee that structural properties such as positive definite symmetric reduced mass, damping, and stiffness

matrices, will be retained. Keeping the original structure is crucial both to preserve physical meaning of the states and to retain physically significant properties such as stability and passivity.

There have been significant efforts in this direction. Building on the earlier work of [82], Bai and Su [13] introduced "second-order" Krylov subspaces and showed how to obtain a reduced-order system directly in a second-order framework while still satisfying interpolation conditions at selected points for single-input/single-output systems. Other second-order structure preserving interpolatory reduction techniques were introduced by Chahlaoui *et al.* [26] and Freund [38], though these approaches require a first-order state-space realization.

The second-order transfer function $\mathbf{H}(s)$ in (39) fits perfectly in the framework of Theorem 1.7. Indeed, due to the simple structures of $\mathcal{C}(s) = \mathbf{C}$, $\mathcal{B}(s) = \mathbf{B}$ and $\mathcal{K}(s) = s^2\mathbf{M} + s\mathbf{G} + \mathbf{K}$, one obtains a simple three-term recurrence for the rational tangential interpolation of MIMO second-order systems directly in a second-order framework. This is presented in Algorithm 4.1.

From Theorem 1.7, the resulting reduced second-order model tangentially interpolates the original model up through derivatives of order $J_i - 1$ at the selected (complex) frequencies σ_i while preserving the original second-order structure. The second-order recursion of Bai and Su [13] is a special case of this algorithm for a single-input/single-output system using a single interpolation point σ.

Algorithm 4.1. Second-order Tangential MIMO Order Reduction
Given interpolation points $\{\sigma_1, \sigma_2, \ldots, \sigma_N\}$, directions $\{\mathsf{b}_1, \mathsf{b}_2, \ldots, \mathsf{b}_N\}$, and interpolation orders $\{J_1, J_2, \ldots, J_N\}$ (with $r = \sum_{i=1}^{N} J_i$).

1. For $i = 1, ..., N$
   a. For each shift σ_i and tangent direction b_i, define $\mathcal{K}_0^{(i)} = \sigma_i^2\mathbf{M} + \sigma_i\mathbf{G} + \mathbf{K}$, $\mathcal{K}_1^{(i)} = 2\sigma_i\mathbf{M} + \mathbf{G}$, and $\mathcal{K}_2 = \mathbf{M}$.
   b. Solve $\mathcal{K}_0^{(i)}\mathbf{f}_{i1} = \mathbf{B}\mathsf{b}_i$ and set $\mathbf{f}_{i0} = 0$.
   c. For $j = 2 : J_i$
      - Solve $\mathcal{K}_0^{(i)}\mathbf{f}_{ij} = -\mathcal{K}_1^{(i)}\mathbf{f}_{i,j-1} - \mathcal{K}_2\mathbf{f}_{i,j-2}$,
2. Take $\mathbf{V}_r = [\mathbf{f}_{11}, \mathbf{f}_{12}, \ldots, \mathbf{f}_{1J_1}, \mathbf{f}_{21}, \ldots, \mathbf{f}_{2J_2}, \ldots, \mathbf{f}_{N1}, \ldots, \mathbf{f}_{NJ_N}]$ and then $\mathbf{M}_r = \mathbf{V}_r^T\mathbf{M}\mathbf{V}_r$, $\mathbf{G}_r = \mathbf{V}_r^T\mathbf{G}\mathbf{V}_r$, $\mathbf{K}_r = \mathbf{V}_r^T\mathbf{K}\mathbf{V}_r$, $\mathbf{B}_r = \mathbf{V}_r^T\mathbf{B}$, and $\mathbf{C}_r = \mathbf{C}\mathbf{V}_r$.

Remark 1.2. The general framework presented here can handle second-order systems of the form

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{G}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}_1\dot{\mathbf{u}}(t) + \mathbf{B}_0\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}_1\dot{\mathbf{x}}(t) + \mathbf{C}_0\mathbf{x}(t), \tag{40}$$

where the input $\mathbf{u}(t) \in \mathbb{R}^m$ is a time-dependent force or displacement, the output $\mathbf{y}(t) \in \mathbb{R}^p$ depends potentially not only on displacements but also the velocities. We note that second-order systems of the form (38) cannot be converted to generic first-order form (1) in a straightforward way when $\mathbf{B}_1 \neq 0$.

Remark 1.3. The discussion above and the algorithm can be easily generalized to higher order constant coefficient ordinary differential equations as well where the system dynamics follow

$$\mathbf{A}_0 \frac{d^\ell \mathbf{x}}{dt^\ell} + \mathbf{A}_1 \frac{d^{\ell-1}\mathbf{x}}{dt^{\ell-1}} + \cdots + \mathbf{A}_\ell \mathbf{x}(t) = \mathbf{B}\mathbf{u}(t) \quad \text{and} \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t). \tag{41}$$

6 Model Reduction of Parametric Systems

Frequently dynamical systems under study are varied by adjusting a comparatively small number of parameters repeatedly during the course of a set of simulations. Parameters can enter a model representing material properties, variations in shape of the structure or more generally of the solution domain, and strength of coupling in various types of boundary conditions. Micro-electromechanical systems, chip design and interconnect modeling are two of the most prominent examples where parameterized systems naturally arise; see, e.g., [29, 36, 51, 78]. It will be of value to have methods that will produce high fidelity reduced order models retaining the parametric structure of the original system. This leads to the concept of parameterized model reduction. The goal is to construct a high fidelity parametric reduced order models which recover the response of the original full order parametric system throughout the range of variation of interest of the design parameters.

Parametric model order reduction is at an early stage of the development. For interpolatory model reduction of parametric systems, see [14, 19, 22, 29, 32, 34–36, 51, 52, 64, 65, 71, 85, 86], and reference therein. These methods treat parametric dynamical systems in a standard first-order state-space form. Below, we extend these results to the generalized setting similar to that of Sect. 5.

We consider a multi-input/multi-output linear dynamical system that is parameterized with q parameters $\mathsf{p} = [\mathsf{p}_1, \ldots, \mathsf{p}_q]$:

$$\mathbf{H}(s,\mathsf{p}) = \mathcal{C}(s,\mathsf{p})\mathcal{K}(s,\mathsf{p})^{-1}\mathsf{b}(s,\mathsf{p}) \tag{42}$$

with $\mathcal{K}(s,\mathsf{p}) \in \mathbb{C}^{n\times n}$ and $\mathcal{B}(s,\mathsf{p}) \in \mathbb{C}^{n\times m}$ and $\mathcal{C}(s,\mathsf{p}) \in \mathbb{C}^{p\times n}$. We assume that

$$\begin{aligned}
\mathcal{K}(s,\mathsf{p}) &= \mathcal{K}^{[0]}(s) + a_1(\mathsf{p})\,\mathcal{K}^{[1]}(s) + \ldots + a_\nu(\mathsf{p})\,\mathcal{K}^{[\nu]}(s)\\
\mathcal{B}(s,\mathsf{p}) &= \mathcal{B}^{[0]}(s) + b_1(\mathsf{p})\,\mathcal{B}^{[1]}(s) + \ldots + b_\nu(\mathsf{p})\,\mathcal{B}^{[\nu]}(s),\\
\mathcal{C}(s,\mathsf{p}) &= \mathcal{C}^{[0]}(s) + c_1(\mathsf{p})\,\mathcal{C}^{[1]}(s) + \ldots + c_\nu(\mathsf{p})\,\mathcal{C}^{[\nu]}(s).
\end{aligned} \tag{43}$$

where $a_1(\mathsf{p}), a_2(\mathsf{p}) \ldots, b_1(\mathsf{p}), ..., c_\nu(\mathsf{p})$ are scalar-valued parameter functions that could be linear or non-linear. Equation (42) represents a system structure that we may wish to retain – that is, our goal is to to generate, for some $r \ll n$, a reduced-order system with dimension r having the same parametric structure. Suppose matrices $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ and $\mathbf{W}_r \in \mathbb{C}^{n\times r}$ are specified and consider:

$$\mathbf{H}_r(s,\mathsf{p}) = \mathcal{C}_r(s,\mathsf{p})\mathcal{K}_r(s,\mathsf{p})^{-1}\mathcal{B}_r(s,\mathsf{p}) \tag{44}$$

with $\mathcal{K}_r(s,\mathsf{p}) = \mathbf{W}_r^T\mathcal{K}(s,\mathsf{p})\mathbf{V}_r$, $\mathcal{B}_r(s,\mathsf{p}) = \mathbf{W}_r^T\mathcal{B}(s,\mathsf{p})$, and $\mathcal{C}_r(s,\mathsf{p}) = \mathcal{C}(s,\mathsf{p})\mathbf{V}_r$. We say that the reduced model has *the same parametric structure* in the sense that

$$\begin{aligned}
\mathcal{K}_r(s,\mathsf{p}) &= \left(\mathbf{W}_r^T\mathcal{K}^{[0]}(s)\mathbf{V}_r\right) + a_1(\mathsf{p})\left(\mathbf{W}_r^T\mathcal{K}^{[1]}(s)\mathbf{V}_r\right) + \ldots + a_\nu(\mathsf{p})\left(\mathbf{W}_r^T\mathcal{K}^{[\nu]}(s)\mathbf{V}_r\right)\\
\mathcal{B}_r(s,\mathsf{p}) &= \left(\mathbf{W}_r^T\mathcal{B}^{[0]}(s)\right) + b_1(\mathsf{p})\left(\mathbf{W}_r^T\mathcal{B}^{[1]}(s)\right) + \ldots + b_\nu(\mathsf{p})\left(\mathbf{W}_r^T\mathcal{B}^{[\nu]}(s)\right),\\
\mathcal{C}_r(s,\mathsf{p}) &= \left(\mathcal{C}^{[0]}(s)\mathbf{V}_r\right) + c_1(\mathsf{p})\left(\mathcal{C}^{[1]}(s)\mathbf{V}_r\right) + \ldots + c_\nu(\mathsf{p})\left(\mathcal{C}^{[\nu]}(s)\mathbf{V}_r\right).
\end{aligned} \tag{45}$$

with exactly the *same* parameter functions $a_1(\mathsf{p}),\ldots,c_\nu(\mathsf{p})$ as in (43), but with smaller coefficient matrices. Significantly, all reduced coefficient matrices can be *precomputed* before the reduced model is put into service. The next result extends Theorem 1.7 to the parameterized dynamical system setting:

Theorem 1.8. *Suppose $\mathcal{K}(s,\mathsf{p})$, $\mathcal{B}(s,\mathsf{p})$, and $\mathcal{C}(s,\mathsf{p})$ are analytic with respect to s at $\sigma \in \mathbb{C}$ and $\mu \in \mathbb{C}$, and are continuously differentiable with respect to p in a neighborhood of $\hat{\mathsf{p}} = [\hat{\mathsf{p}}_1,\ldots,\hat{\mathsf{p}}_q]$. Suppose further that both $\mathcal{K}(\sigma,\hat{\mathsf{p}})$ and $\mathcal{K}(\mu,\hat{\mathsf{p}})$ are nonsingular and matrices $\mathbf{V}_r \in \mathbb{C}^{n\times r}$ and $\mathbf{W}_r \in \mathbb{C}^{n\times r}$ are given such that both $\mathcal{K}_r(\sigma,\hat{\mathsf{p}}) = \mathbf{W}_r^T\mathcal{K}(\sigma,\hat{\mathsf{p}})\mathbf{V}_r$ and $\mathcal{K}_r(\mu,\hat{\mathsf{p}}) = \mathbf{W}_r^T\mathcal{K}(\mu,\hat{\mathsf{p}})\mathbf{V}_r$ are also nonsingular. For nontrivial tangential directions $\mathsf{b} \in \mathbb{C}^m$ and $\mathsf{c} \in \mathbb{C}^p$:*

(a) If $\mathcal{K}(\sigma,\hat{\mathsf{p}}))^{-1}\mathcal{B}(\sigma,\hat{\mathsf{p}})\mathsf{b} \in \mathsf{Ran}(\mathbf{V}_r)$ then $\mathbf{H}(\sigma,\hat{\mathsf{p}}))\mathsf{b} = \mathbf{H}_r(\sigma,\hat{\mathsf{p}}))\mathsf{b}$
(b) If $\left(\mathsf{c}^T\mathcal{C}(\mu,\hat{\mathsf{p}}))\mathcal{K}(\mu,\hat{\mathsf{p}}))^{-1}\right)^T \in \mathsf{Ran}(\mathbf{W}_r)$ then $\mathsf{c}^T\mathbf{H}(\mu,\hat{\mathsf{p}})) = \mathsf{c}^T\mathbf{H}_r(\mu,\hat{\mathsf{p}}))$
(c) If both (a) and (b) hold and if $\sigma = \mu$, then

$$\nabla_{\mathsf{p}}\mathsf{c}^T\mathbf{H}(\sigma,\hat{\mathsf{p}})\mathsf{b} = \nabla_{\mathsf{p}}\mathsf{c}^T\mathbf{H}_r(\sigma,\hat{\mathsf{p}})\mathsf{b} \quad \text{and} \quad \mathsf{c}^T\mathbf{H}'(\sigma,\hat{\mathsf{p}})\mathsf{b} = \mathsf{c}^T\mathbf{H}_r'(\sigma,\hat{\mathsf{p}})\mathsf{b}$$

Proof. We prove the first two assertions only. The proof of the third assertion is rather technical. We refer the reader to [14] for the proof of this fact for the special case of $\mathcal{K}(s,\mathsf{p}) = s\mathbf{E}(\mathsf{p}) - \mathbf{A}(\mathsf{p})$.

Define the projections

$$\mathbf{P}_r(s,\mathsf{p}) = \mathbf{V}_r\mathcal{K}_r(s,\mathsf{p})^{-1}\mathbf{W}_r^T\mathcal{K}(s,\mathsf{p}) \text{ and}$$

$$\mathbf{Q}_r(s,\mathsf{p}) = \mathcal{K}(s,\mathsf{p})\mathbf{V}_r\mathcal{K}_r(s,\mathsf{p})^{-1}\mathbf{W}_r^T$$

Also, define the vector $\mathbf{f}_0 = \mathcal{K}(\sigma,\hat{\mathsf{p}})^{-1}\mathcal{B}(s,\hat{\mathsf{p}})\mathsf{b}$. Note that the assumption of the first assertion implies that $\mathbf{f}_0 \in \mathsf{Ran}(\mathbf{P}_r(\sigma,\hat{\mathsf{p}}))$. Then, a direct computation yields

$$\mathbf{H}(\sigma,\hat{\mathsf{p}})\mathsf{b} - \mathbf{H}_r(\sigma,\hat{\mathsf{p}})\mathsf{b} = \mathcal{C}(s,\hat{\mathsf{p}})\left(\mathbf{I} - \mathcal{P}_r\right)\mathbf{f}_0 = 0,$$

which proves the first assertion. Similarly, define $\mathbf{g}_0^T = \mathsf{c}^T\mathcal{C}(s,\hat{\mathsf{p}})\mathcal{K}(\mu,\hat{\mathsf{p}})^{-1}$ and observe that $\mathbf{g}_0 \perp \mathrm{Ker}(\mathbf{Q}_r(\mu,\hat{\mathsf{p}}))$ due to the assumption of the second assumption. Then, one can directly obtain

$$\mathsf{c}^T\mathbf{H}(\mu,\hat{\mathsf{p}}) - \mathsf{c}^T\mathbf{H}_r(\mu,\hat{\mathsf{p}}) = \mathbf{g}_0^T\left(\mathbf{I} - \mathbf{Q}_r\right)\mathcal{B}(s,\hat{\mathsf{p}}) = 0,$$

which proves the second assertion. □

Theorem 1.8 extends the interpolation theory for the non-parametric systems to the parameterized ones. It can be easily extended to match higher order derivatives as well. This result is much more general than others in the literature in the sense that we allow transfer function to be in the generalized form; hence not constraining it to be in standard state-space form. Also, the result allows for arbitrary linear and nonlinear dependency in $\mathcal{K}$, $\mathcal{C}$, and $\mathcal{B}$, which themselves can reflect the greater generality of the structured dynamical systems setting considered earlier.

Given $\mathbf{H}(s,\mathsf{p}) = \mathcal{C}(s,\mathsf{p})\mathcal{K}(s,\mathsf{p})^{-1}\mathbf{B}(s,\mathsf{p})$ as defined in (43), let us assume that we want to construct a parametric reduced-model that matches $\mathbf{H}(s)$ at frequency points $\{\sigma_i\}_{i=1}^{K} \in \mathbb{C}$ and $\{\mu_i\}_{i=1}^{K} \in \mathbb{C}$, the parameter points $\{\mathsf{p}^{(j)}\}_{j=1}^{L} \in \mathbb{C}^q$ along the right tangential directions $\{\mathsf{b}_{ij}\}_{i=1,j=1}^{K,L} \in \mathbb{C}^m$ and the right tangential directions $\{\mathsf{c}_{ij}\}_{i=1,j=1}^{K,L} \in \mathbb{C}^p$. Define, for $i = 1,\ldots,K$ and $j = 1,\ldots,L$,

$$\mathbf{v}_{ij} = \mathcal{K}(\sigma_i,\mathsf{p}^{(j)})^{-1}\mathcal{B}(\sigma_i,\mathsf{p}^{(j)})\mathsf{b}_{i,j} \quad \text{and} \quad \mathbf{w}_{ij} = \mathcal{K}(\mu_i,\mathsf{p}^{(j)})^{-T}\mathcal{C}(\mu_i,\mathsf{p}^{(j)})^{T}\mathsf{c}_{i,j}$$

Then, construct $\mathbf{V}_r$ and $\mathbf{W}_r$ such that

$$\mathbf{V}_r = [\mathbf{v}_{11},\ldots,\mathbf{v}_{1L},\mathbf{v}_{21},\ldots,\mathbf{v}_{2L},\ldots,\mathbf{v}_{K1},\ldots,\mathbf{v}_{KL}] \in \mathbb{C}^{n\times(KL)}$$

and

$$\mathbf{W}_r = [\mathbf{w}_{11},\ldots,\mathbf{w}_{1L},\mathbf{w}_{21},\ldots,\mathbf{w}_{2L},\ldots,\mathbf{w}_{K1},\ldots,\mathbf{w}_{KL}] \in \mathbb{C}^{n\times(KL)}.$$

Then, the resulting projection-based reduced-order parametric model of (45) satisfies the interpolation condition stated in Theorem 1.8.

As the discussion above illustrates, *once the parameter points* $\{\mathsf{p}^{(j)}\}$ *are selected*, the machinery for interpolatory model reduction of parametric systems is very similar to that of non-parametric systems. However, in this case one faces the problem choosing not only the frequency points $\{\sigma_i\}$ but also the parameter points $\{\mathsf{p}^{(j)}\}$. Hence, the question of effective parameter point selection arises. In some application, the designer specifies important parameter sets to be used in the model reduction. However, the task becomes much harder if one only has the information as the range of parameter space without any knowledge of what parameter sets might be important. In this case, Bui-Thanh et al. [23] proposed the so-called greedy selection algorithm as a possible solution. Even though the method [23] proves to yield high quality approximations, the optimization algorithm that needs to be solved at each step could be computationally expensive. Another strategy for an efficient choice of parameters points in a higher dimensional parameter sparse is the use of sparse grid points [24, 41, 92]. An *optimal* parameter selection strategy was recently proposed by Baur et al. [14] for the special case of the parametric single-input/single-output dynamical systems of the form $\mathbf{H}(s) = (\mathbf{c}_0 + p_1\mathbf{c}_1)^T(s\mathbf{I}-\mathbf{A})^{-1}(\mathbf{b}_0 + p_2\mathbf{b}_1)$ where the scalars p_1 and p_2 are independent parameters and the system dynamic $\mathbf{A}$ is non-parametric. The optimal parameter set is obtained by converting the problem into an equivalent non-parametric optimal $\mathcal{H}_2$ MIMO approximation.

6.1 Numerical Example

We illustrate the concept of parametric model reduction using a model representing thermal conduction in a semiconductor chip as described in [61]. An efficient model of thermal conduction should allow flexibility in specifying boundary conditions so that designers are able to evaluate the effects of environmental changes on the temperature distribution in the chip. The problem is modeled as homogenous heat diffusion with heat exchange occurring at three device interfaces modeled with convection boundary conditions that introduce film coefficients, p_1, p_2, and p_3, describing the heat exchange on the three device interfaces. Discretization of the underlying partial differential equations leads to a system of ordinary differential equations

$$\mathbf{E}\dot{\mathbf{x}}(t) = (\mathbf{A} + \sum_{i=1}^{3} \mathsf{p}_i \mathbf{A}_i)\mathbf{x}(t) + \mathbf{B}u(t), \;\; \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

where $\mathbf{E} \in \mathbb{R}^{4257\times 4257}$ and $\mathbf{A} \in \mathbb{R}^{4257\times 4257}$ are system matrices, $\mathbf{A}_i \in \mathbb{R}^{4257\times 4257}$, $i = 1,\ldots,3$, are diagonal matrices arising from the discretization of the convection boundary condition on the ith interface, and $\mathbf{B} \in \mathbb{R}^{4257}$ and $\mathbf{C} \in \mathbb{R}^{7\times 4257}$. We note even though Theorem 1.8 allows parametric model reduction in a much more general setting, we use a standard state-space model to illustrate the theory.

Each parameter value varies in the range of $[1, 10^4]$. Important parameter set values are listed in [61] of which we use the following two to apply model reduction:

$$\mathsf{p}^{(1)} = (10^4, 10^4, 1) \;\text{ and }\; \mathsf{p}^{(2)} = (1,1,1)$$

We follow an $\mathcal{H}_2$-inspired approach recently proposed in [14]: Note that Once the parameter set $\mathsf{p}^{(i)}$ for $i = 1,2$ are plugged into the state-space matrices, a non-parametric standard transfer function is obtained, denoted by $\mathbf{H}(s,\mathsf{p}^{(i)})$. Then, we apply the optimal $\mathcal{H}_2$ model reduction technique **IRKA** as outlined in Algorithm 1 to $\mathbf{H}(s,\mathsf{p}^{(i)})$ to reduce the order to r_i leading to projection subspaces $\mathbf{V}^{(i)} \in \mathbb{R}^{4257\times r_i}$ and $\mathbf{W}^{(i)} \in \mathbb{R}^{4257\times r_i}$ for $i = 1,2$ with $r_1 = 6$ and $r_2 = 5$. We, then, concatenate these matrices to build the final projection matrices

$$\mathbf{V}_r = [\mathbf{V}^{(1)},\, \mathbf{V}^{(2)}] \in \mathbb{R}^{4257\times 11} \;\text{ and }\; \mathbf{W}_r = [\mathbf{W}^{(1)},\, \mathbf{W}^{(2)}] \in \mathbb{R}^{4257\times 11}.$$

We obtain a final parameterized reduced-order model of order$r = r_1 + r_2 = 11$:

$$\begin{aligned}\mathbf{W}_r^T\mathbf{E}\mathbf{V}_r\dot{\mathbf{x}}_r(t) &= (\mathbf{W}_r^T\mathbf{A}\mathbf{V}_r + \sum_{i=1}^{3} \mathsf{p}_i \mathbf{W}_r^T\mathbf{A}_i\mathbf{V}_r)\mathbf{x}_r(t) + \mathbf{W}_r^T\mathbf{B}u(t),\\ \mathbf{y}_r(t) &= \mathbf{C}\mathbf{V}_r\mathbf{x}_r(t).\end{aligned}$$

A high-quality parameterized reduced-order model must represent the full parameterized model with high fidelity for a wide range of parameter values. To illustrate this is the case, we fix $\mathsf{p}_3 = 1$ and vary both p_1 and p_2 over the full parameter range, i.e., between 1 and 10^4. Then, for each mesh point (i.e. for each triple of parameter values in this range), we compute the corresponding full-order model

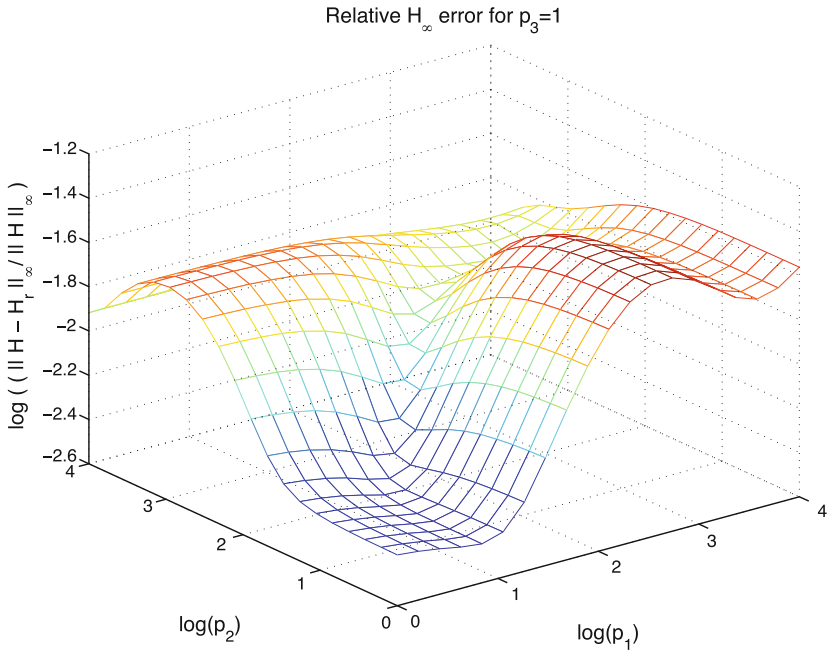


Fig. 4 Relative $\mathcal{H}_\infty$-error as p_1 and p_2 vary

and the reduced-order model; and compute the corresponding relative $\mathcal{H}_\infty$-errors. Figure 4 depicts the resulting mesh plot. As the figure illustrates, the reduced-model approximate the full-model uniformly well throughout the full parameter range. The maximum relative $\mathcal{H}_\infty$-error is 3.50×10^{-2}. Therefore, the parameterized reduced-order model $\mathbf{H}_r(s, \mathsf{p})$ of order $r = 11$ yields a relative accuracy of 10^{-2} for the complete range of p_1 and p_2.

7 Model Reduction from Measurements

In many instances, input/output measurements replace an explicit model of a to-be-simulated system. In such cases it is of great interest to be able to efficiently construct models and reduced models from the available data. In this section we will address this problem by means of *rational interpolation*. We will show that the natural tool is the *Loewner matrix pencil*, which is closely related to the *Hankel matrix*. For details we refer to [62, 66].

7.1 Motivation: S-Parameters

The growth in communications and networking and the demand for high data bandwidth requires streamlining of the simulation of entire complex systems from chips

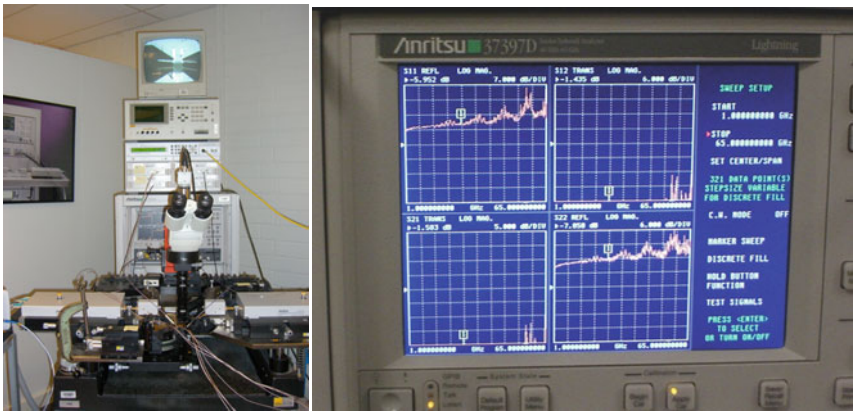


Fig. 5 VNA (Vector Network Analyzer) and VNA screen showing the magnitude of the *S*-parameters for a two port

to packages to boards. Moreover, in circuit simulation, signal intergrity (lack of signal distortion) of high speed electronic components requires that interconnect models be valid over a wide bandwidth.

An important tool for dealing with such problems is the *S*- or *scattering-parameter* system representation. The *S*-parameters represent a system as a black box. An important advantage is that these parameters can be measured using VNAs (Vector Network Analyzers) (see Fig. 5). Also, their norm for passive components, is not greater than one, and in high frequencies *S*-parameters are important because wave phenomena become dominant.

We recall that given a system in input/output representation: $\hat{\mathbf{y}}(s) = \mathbf{H}(s)\hat{\mathbf{u}}(s)$, the associated *S-parameter representation* is

$$\hat{\mathbf{y}}_s = \underbrace{[\mathbf{H}(s)+\mathbf{I}][\mathbf{H}(s)-\mathbf{I}]^{-1}}_{\mathbf{S}(s)}\,\hat{\mathbf{u}}_s(s),$$

where $\hat{\mathbf{y}}_s = \frac{1}{2}(\hat{\mathbf{y}}+\hat{\mathbf{u}})$ are the *transmitted waves* and, $\hat{\mathbf{u}}_s = \frac{1}{2}(\hat{\mathbf{y}}-\hat{\mathbf{u}})$ are the *reflected waves*. Thus the *S-parameter measurements* $\mathbf{S}(j\omega_k)$, are samples of the frequency response of the *S*-parameter system representation.

In the sections that follow we will describe a framework which applies also to the efficient modeling from *S*-parameter measurements. An example is provided in Sect. 7.4.3. For more details and examples see [62].

7.2 The Loewner Matrix Pair and Construction of Interpolants

Suppose we have observed response data as described in Problem 2 in Sect. 2.1. We are given r (right) driving frequencies: $\{\sigma_i\}_{i=1}^r \subset \mathbb{C}$ that use input directions $\{\tilde{\mathsf{b}}_i\}_{i=1}^r \subset \mathbb{C}^m$, to produce system responses, $\{\tilde{\mathsf{y}}_i\}_{i=1}^r \subset \mathbb{C}^p$ and q (left) driving fre-

quencies: $\{\mu_i\}_{i=1}^q \subset \mathbb{C}$ that use dual (left) input directions $\{\tilde{\mathsf{c}}_i\}_{i=1}^q \subset \mathbb{C}^p$, to produce dual (left) responses, $\{\tilde{\mathsf{z}}_i\}_{i=1}^q \subset \mathbb{C}^m$. We assume that there is an underlying dynamical system as defined in (1) so that

$$\begin{array}{ll} \tilde{\mathsf{c}}_i^T \mathbf{H}(\mu_i) = \tilde{\mathsf{z}}_i^T & \text{and} \qquad \mathbf{H}(\sigma_j)\tilde{\mathsf{b}}_j = \tilde{\mathsf{y}}_j, \\ \text{for } i = 1, \ldots, q, & \text{for } j = 1, \ldots, r. \end{array}$$

yet we are given access only to the $r+q$ response observations listed above and have no other information about the underlying system $\mathbf{H}(s)$. In this section we will sketch the solution of Problem 2 described in Sect. 2.1. Towards this goal we will introduce the Loewner matrix pair in the tangential interpolation case. The *Loewner matrix* is defined as follows:

$$\mathbb{L} = \begin{bmatrix} \dfrac{\tilde{\mathsf{z}}_1^T \tilde{\mathsf{b}}_1 - \tilde{\mathsf{c}}_1^T \tilde{\mathsf{y}}_1}{\mu_1 - \sigma_1} & \cdots & \dfrac{\tilde{\mathsf{z}}_1^T \tilde{\mathsf{b}}_r - \tilde{\mathsf{c}}_1^T \tilde{\mathsf{y}}_r}{\mu_1 - \sigma_r} \\ \vdots & \ddots & \vdots \\ \dfrac{\tilde{\mathsf{z}}_q^T \tilde{\mathsf{b}}_1 - \tilde{\mathsf{c}}_q^T \tilde{\mathsf{y}}_1}{\mu_q - \sigma_1} & \cdots & \dfrac{\tilde{\mathsf{z}}_q^T \tilde{\mathsf{b}}_r - \tilde{\mathsf{c}}_q^T \tilde{\mathsf{y}}_r}{\mu_q - \sigma_r} \end{bmatrix} \in \mathbb{C}^{q \times r}.$$

If we define matrices associated with the system observations as:

$$\widetilde{\mathsf{B}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \tilde{\mathsf{b}}_1 & \tilde{\mathsf{b}}_2 & \ldots & \tilde{\mathsf{b}}_r \\ \vdots & \vdots & & \vdots \end{bmatrix} \qquad \widetilde{\mathsf{Y}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \tilde{\mathsf{y}}_1 & \tilde{\mathsf{y}}_2 & \ldots & \tilde{\mathsf{y}}_r \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$\widetilde{\mathsf{Z}}^T = \begin{bmatrix} \ldots & \tilde{\mathsf{z}}_1^T & \ldots \\ \ldots & \tilde{\mathsf{z}}_2^T & \ldots \\ & \vdots & \\ \ldots & \tilde{\mathsf{z}}_q^T & \ldots \end{bmatrix} \qquad \widetilde{\mathsf{C}}^T = \begin{bmatrix} \ldots & \tilde{\mathsf{c}}_1^T & \ldots \\ \ldots & \tilde{\mathsf{c}}_2^T & \ldots \\ & \vdots & \\ \ldots & \tilde{\mathsf{c}}_q^T & \ldots \end{bmatrix}$$

$\mathbb{L}$ satisfies the Sylvester equation

$$\mathbb{L}\Sigma - M\mathbb{L} = \widetilde{\mathsf{C}}^T \widetilde{\mathsf{Y}} - \widetilde{\mathsf{Z}}^T \widetilde{\mathsf{B}}. \tag{46}$$

where $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) \in \mathbb{C}^{r \times r}$ and $M = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_q) \in \mathbb{C}^{q \times q}$ Suppose that state space data $(\mathbf{E}, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$, of minimal degree n are given such that $\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B}$+$\mathbf{D}$. If the generalized eigenvalues of $(\mathbf{A}, \mathbf{E})$ are distinct from σ_i and μ_j, we define $\mathbf{V}_r$ and $\mathbf{W}_q^T$ as

$$\mathbf{V}_r = \left[(\sigma_1 \mathbf{E} - \mathbf{A})^{-1} \mathbf{B} \tilde{\mathsf{b}}_1, \cdots, (\sigma_r \mathbf{E} - \mathbf{A})^{-1} \mathbf{B} \tilde{\mathsf{b}}_r \right] \in \mathbb{C}^{n \times r} \text{ and}$$

$$\mathbf{W}_q^T = \begin{bmatrix} \mathsf{c}_1^T \mathbf{C}(\mu_1 \mathbf{E} - \mathbf{A})^{-1} \\ \mathsf{c}_2^T \mathbf{C}(\mu_2 \mathbf{E} - \mathbf{A})^{-1} \\ \vdots \\ \mathsf{c}_q^T \mathbf{C}(\mu_q \mathbf{E} - \mathbf{A})^{-1} \end{bmatrix} \in \mathbb{C}^{q \times n}.$$

It follows that

$$\mathbb{L} = -\mathbf{W}_q^T \mathbf{E} \mathbf{V}_r$$

and we call $\mathbf{V}_r$ and $\mathbf{W}_q^T$ *generalized tangential* reachability and observability matrices, respectively.

Next we introduce a new object which is pivotal in our approach. This is the *shifted Loewner matrix*, defined as follows:

$$\mathbb{M} = \begin{bmatrix} \dfrac{\mu_1 \tilde{z}_1^T \tilde{b}_1 - \sigma_1 \tilde{c}_1^T \tilde{y}_1}{\mu_1 - \sigma_1} & \cdots & \dfrac{\mu_1 \tilde{z}_1^T \tilde{b}_r - \sigma_r \tilde{c}_1^T \tilde{y}_r}{\mu_1 - \sigma_r} \\ \vdots & \ddots & \vdots \\ \dfrac{\mu_q \tilde{z}_q^T \tilde{b}_1 - \sigma_1 \tilde{c}_q^T \tilde{y}_1}{\mu_q - \sigma_1} & \cdots & \dfrac{\mu_q \tilde{z}_q^T \tilde{b}_r - \sigma_r \tilde{c}_q^T \tilde{y}_r}{\mu_q - \sigma_r} \end{bmatrix} \in \mathbb{C}^{q \times r}$$

$\mathbb{M}$ satisfies the Sylvester equation

$$\mathbb{M}\boldsymbol{\Sigma} - M\mathbb{M} = \widetilde{\mathsf{C}}^T \widetilde{\mathsf{Y}} \boldsymbol{\Sigma} - M \widetilde{\mathsf{Z}}^T \widetilde{\mathsf{B}}. \tag{47}$$

If an interpolant $\mathbf{H}(s)$ is associated with the interpolation data, the shifted Loewner matrix is the Loewner matrix associated to $s\mathbf{H}(s)$. If a state space representation is available, then like for the Loewner matrix, the shifted Loewner matrix can be factored as

$$\mathbb{M} = -\mathbf{W}_q^T \mathbf{A} \mathbf{V}_r.$$

It therefore becomes apparent that $\mathbb{L}$ contains information about $\mathbf{E}$ while $\mathbb{M}$ contains information about $\mathbf{A}$. These observations are formalized in one of the main results of this section which shows how straightforward the solution of the interpolation problem becomes, in the Loewner matrix framework.

Theorem 1.9. *Assume that $r = q$ and that $\mu_i \neq \sigma_j$ for all $i, j = 1, \ldots, r$. Suppose that $\mathbb{M} - s\mathbb{L}$ is invertible for all $s \in \{\sigma_i\} \cup \{\mu_j\}$. Then, with*

$$\mathbf{E}_r = -\mathbb{L}, \quad \mathbf{A}_r = -\mathbb{M}, \quad \mathbf{B}_r = \widetilde{\mathsf{Z}}^T, \quad \mathbf{C}_r = \widetilde{\mathsf{Y}}, \quad \mathbf{D}_r = 0,$$

$$\mathbf{H}_r(s) = \mathbf{C}_r(s\mathbf{E}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r = \widetilde{\mathsf{Z}}^T(\mathbb{M} - s\mathbb{L})^{-1}\widetilde{\mathsf{Y}}$$

interpolates the data and furthermore is a minimal realization.

Next we will outline two proofs of this important result. They are both straightforward and hence reveal the main attributes of this approach.

Proof. Multiplying (46) by s and subtracting it from (47) we get

$$(\mathbb{M} - s\mathbb{L})\boldsymbol{\Sigma} - M(\mathbb{M} - s\mathbb{L}) = \widetilde{\mathsf{C}}^T \widetilde{\mathsf{Y}}(\boldsymbol{\Sigma} - s\mathbf{I}) - (M - s\mathbf{I})\widetilde{\mathsf{Z}}^T \widetilde{\mathsf{B}}. \tag{48}$$

Multiplying (48) by $\mathbf{e}_j$ (the jth unit vector) on the right and setting $s = \sigma_j$, we obtain

$$\begin{aligned} &(\sigma_j \mathbf{I} - M)(\mathbb{M} - \sigma_j \mathbb{L})\mathbf{e}_j = (\sigma_j \mathbf{I} - M)\widetilde{\mathsf{Z}}^T \tilde{b}_j \;\Rightarrow \\ &(\mathbb{M} - \sigma_i \mathbb{L})\mathbf{e}_j = \widetilde{\mathsf{Z}}^T \tilde{b}_j \;\Rightarrow\; \widetilde{\mathsf{Y}}\mathbf{e}_j = \widetilde{\mathsf{Y}}(\mathbb{M} - \sigma_j \mathbb{L})^{-1}\widetilde{\mathsf{Z}}^T \mathsf{b}_j \end{aligned}$$

Therefore $\mathbf{H}(\sigma_j)\tilde{\mathsf{b}}_j = \tilde{\mathsf{y}}_j$. This proves the right tangential interpolation property. To prove the left tangential interpolation property, we multiply (48) by $\mathbf{e}_i^T$ (the transpose of the ith unit vector) on the left and set $s = \mu_i$:

$$\begin{aligned}&\mathbf{e}_i^T(\mathbb{M}-\mu_i\mathbb{L})(\Sigma-\mu_i\mathbf{I}) = \widetilde{\mathsf{c}}_i^T\widetilde{\mathsf{Y}}(\Sigma-\mu_i\mathbf{I}) \;\;\Rightarrow\\ &\mathbf{e}_i^T(\mathbb{M}-\mu_i\mathbb{L}) = \widetilde{\mathsf{c}}_i^T\widetilde{\mathsf{Y}} \;\;\Rightarrow\;\; \mathbf{e}_i^T\widetilde{\mathsf{Z}}^T = \widetilde{\mathsf{c}}_i^T\widetilde{\mathsf{Y}}(\mathbb{M}-\mu_i\mathbb{L})^{-1}\widetilde{\mathsf{Z}}^T.\end{aligned}$$

Therefore $\tilde{\mathsf{c}}_i^T\mathbf{H}(\mu_i) = \tilde{\mathsf{z}}_i^T$, which completes the proof. □

Proof (Alternate version). For this proof we assume that the data have been obtained by sampling a known high order system in state space form $(\mathbf{A},\mathbf{E},\mathbf{B},\mathbf{C})$. Direct calculation verifies that

$$\mathbf{H}(\mu_i)-\mathbf{H}(\sigma_j) = (\sigma_j-\mu_i)\,\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}.$$

Pre-multiplication of this expression by $\tilde{\mathsf{c}}_i^T$ and post-multiplication by $\tilde{\mathsf{b}}_j$ immediately yields

$$\mathbb{L} = -\mathbf{W}_q^T\mathbf{E}\mathbf{V}_r.$$

One may directly calculate first,

$$\begin{aligned}\sigma_j(\sigma_j\mathbf{E}-\mathbf{A})^{-1} &= \sigma_j(\mu_i\mathbf{E}-\mathbf{A})^{-1}(\mu_i\mathbf{E}-\mathbf{A})(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\\ &= \mu_i\sigma_j(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\mu_i\mathbf{E}-\mathbf{A})^{-1} - \sigma_j(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\end{aligned}$$

and

$$\begin{aligned}\mu_i(\mu_i\mathbf{E}-\mathbf{A})^{-1} &= \mu_i(\mu_i\mathbf{E}-\mathbf{A})^{-1}(\sigma_j\mathbf{E}-\mathbf{A})(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\\ &= \mu_i\sigma_j(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\mu_i\mathbf{E}-\mathbf{A})^{-1} - \mu_i(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\end{aligned}$$

$$\begin{aligned}\sigma_j\mathbf{H}(\sigma_j) &= \mu_i\sigma_j\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} - \sigma_j\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\\ \mu_i\mathbf{H}(\mu_i) &= \mu_i\sigma_j\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{E}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B} - \mu_i\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}\end{aligned}$$

which in turn implies that

$$\mu_i\,\mathbf{H}(\mu_i)-\sigma_j\,\mathbf{H}(\sigma_j) = (\sigma_j-\mu_i)\,\mathbf{C}(\mu_i\mathbf{E}-\mathbf{A})^{-1}\mathbf{A}(\sigma_j\mathbf{E}-\mathbf{A})^{-1}\mathbf{B}.$$

Pre-multiplication of this expression by $\tilde{\mathsf{c}}_i^T$ and post-multiplication by $\tilde{\mathsf{b}}_j$ immediately yields

$$\mathbb{M} = -\mathbf{W}_q^T\mathbf{A}\mathbf{V}_r.$$

Finally, note that $\widetilde{\mathsf{Z}}^T = \mathbf{W}_q^T\mathbf{B}$ and $\widetilde{\mathsf{Y}} = \mathbf{C}\mathbf{V}_r$. Thus $\mathbf{H}_r(s) = \widetilde{\mathsf{Y}}(\mathbb{M}-s\mathbb{L})^{-1}\widetilde{\mathsf{Z}}^T$ is a tangential interpolant to $\mathbf{H}(s)$. □

Remark 1.4. The Loewner matrix was introduced in [5]. As shown therein, its usefulness derives from the fact that its rank is equal to the McMillan degree of the

underlying interpolant. It also turns out that Loewner matrices for confluent interpolation problems (i.e., problems where the values of a certain number of derivatives are provided as well) have Hankel structure. Thus the Loewner matrix provides a direct link with classical realization theory. For an overview of these results we refer to Sects. 4.4 and 4.5 in [2]. In [1], one way of constructing state space realizations based on the Loewner matrix was presented. But the shifted Loewner matrix, first introduced in [66], was missing and consequently the resulting procedure is only applicable to proper rational interpolants.

7.2.1 The General Case

If the assumption of the above theorem is not satisfied, one needs to project onto the column span and onto the row span of a linear combination of the two Loewner matrices. More precisely, let the following assumption be satisfied:

$$\operatorname{rank}(s\mathbb{L}-\mathbb{M}) = \operatorname{rank}\begin{pmatrix}\mathbb{L} & \mathbb{M}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\mathbb{L}\\ \mathbb{M}\end{pmatrix} \geq \rho, \text{ for all } s \in \{\sigma_i\} \cup \{\mu_j\}.$$

ρ is a truncation index that is assumed to be no larger than $\operatorname{rank} s\mathbb{L}-\mathbb{M}$. $\rho \leq q, r$ and choosing ρ to be the *numerical* rank of $s\mathbb{L}-\mathbb{M}$ is convenient. The best tool for determining the numerical rank of $s\mathbb{L}-\mathbb{M}$, is the SVD (Singular Value Decomposition). To that end, suppose

$$s\mathbb{L}-\mathbb{M} = \mathbf{Y}\Theta\mathbf{X}^*,$$

is the SVD of $s\mathbb{L}-\mathbb{M}$ for some choice of $s \in \{\sigma_i\} \cup \{\mu_j\}$ and consider a truncated SVD as $\mathbf{Y}_\rho \in \mathbb{C}^{q\times\rho}$, $\mathbf{X}_\rho \in \mathbb{C}^{r\times\rho}$.

Theorem 1.10. *A realization* $[\mathbf{E}_\rho, \mathbf{A}_\rho, \mathbf{B}_\rho, \mathbf{C}_\rho]$*, of a minimal solution is given as follows:*

$$\mathbf{E}_\rho = -\mathbf{Y}_\rho^*\mathbb{L}\mathbf{X}_\rho,\ \ \mathbf{A}_\rho = -\mathbf{Y}_\rho^*\mathbb{M}\mathbf{X}_\rho,\ \ \mathbf{B}_\rho = \mathbf{Y}_\rho^*\widetilde{\mathsf{Y}},\ \ \mathbf{C}_\rho = \widetilde{\mathsf{Z}}^T\mathbf{X}_\rho.$$

Depending on whether ρ *is the exact or approximate rank, we obtain either an interpolant or an approximate interpolant of the data, respectively.*

7.3 Loewner and Pick Matrices

The positive real interpolation problem can be formulated as follows. Given triples $(\sigma_i, \tilde{\mathsf{b}}_i, \tilde{\mathsf{y}}_i)$, $i = 1,\ldots,q$, where σ_i are distinct complex numbers in the right-half of the complex plane, $\tilde{\mathsf{b}}_i$ and $\tilde{\mathsf{y}}_i$ are in $\mathbb{C}^r$, we seek a rational matrix function $\mathbf{H}(s)$ of size $r \times r$, such that $\mathbf{H}(\sigma_i)\tilde{\mathsf{b}}_i = \tilde{\mathsf{y}}_i$, $i = 1,\ldots,q$, and in addition $\mathbf{H}$ is positive real. This problem does not always have a solution. It is well known that the necessary and sufficient condition for its solution is that the associated *Pick* matrix

$$\Pi = \begin{bmatrix} \dfrac{\tilde{y}_1^*\tilde{b}_1 + \tilde{b}_1^*\tilde{y}_1}{\overline{\sigma_1}+\sigma_1} & \cdots & \dfrac{\tilde{y}_1^*\tilde{b}_q + \tilde{b}_1^*\tilde{y}_q}{\overline{\sigma_1}+\sigma_q} \\ \vdots & \ddots & \vdots \\ \dfrac{\tilde{y}_q^*\tilde{b}_1 + \tilde{b}_q^*\tilde{y}_1}{\overline{\sigma_q}+\sigma_1} & \cdots & \dfrac{\tilde{y}_q^*\tilde{b}_q + \tilde{b}_q^*\tilde{y}_q}{\overline{\sigma_q}+\sigma_q} \end{bmatrix} \in \mathbb{C}^{q\times q},$$

be positive semi-definite, that is $\Pi = \Pi^* \geq \mathbf{0}$. By comparing Π with the Loewner matrix $\mathbb{L}$ defined in Sect. 7.2, we conclude that if the right (column) array for the former is taken as $(\sigma_i, \tilde{b}_i, \tilde{y}_i)$, $i = 1, \ldots, q$, and the left (row) array as $(-\overline{\sigma_i}, \tilde{b}_i^*, -\tilde{y}_i^*)$, $i = 1, \ldots, q$, then

$$\Pi = \mathbb{L}.$$

The left array is then called the *mirror-image array*. Thus for this choice of interpolation data the Pick matrix is the same as the Loewner matrix. This shows the importance of the Loewner matrix as a tool for studying rational interpolation.

Remark 1.5. **(a)** The above considerations provide an algebraization of the positive real interpolation problem. If namely, $\Pi \geq \mathbf{0}$, the minimal-degree rational functions which interpolate *simultaneously* the original array *and* its mirror image array, are automatically *positive real* and hence *stable* as well. The data in the model reduction problem that we studied in Sect. 4, *automatically* satisfy this positive definiteness constraint, and therefore the reduced system is positive real.

(b) It readily follows that interpolants of the original and the mirror-image arrays constructed by means of the Loewner matrix, satisfy

$$\left[\mathbf{H}(\sigma_i) + \mathbf{H}^T(-\sigma_i)\right] \tilde{b}_i = \mathbf{0}.$$

In general the zeros σ_i of $\mathbf{H}(s) + \mathbf{H}^T(-s)$ are called *spectral zeros*, and $\tilde{b}_i$ are the corresponding (right) zero directions. Thus the construction of positive real interpolants by means of the Loewner (Pick) matrix, forces these interpolants to have the *given* interpolation points as *spectral zeros*.

(c) The observation that the Pick matrix is a special case of the Loewner matrix, the algebraization discussed above, and passivity preservation by means of spectral zero interpolation, first appeared in [3]. See also [4].

7.4 Examples

7.4.1 A Simple Low-order Example

First we will illustrate the above results by means of a simple example. Consider a 2×2 rational function with minimal realization:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Thus the transfer function is

$$\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} = \begin{bmatrix} s & 1 \\ 1 & \frac{1}{s} \end{bmatrix}.$$

Since $\operatorname{rank}\mathbf{E} = 2$, the McMillan degree of $\mathbf{H}$ is 2. Our goal is to recover this function through interpolation. The data will be chosen in two different ways.

First, we will choose *matrix data*, that is the values of the whole matrix are available at each interpolation point:

$$\sigma_1 = 1, \qquad \sigma_2 = 1, \qquad \sigma_3 = 2, \qquad \sigma_4 = 2,$$

$$\tilde{\mathsf{b}}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{\mathsf{b}}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tilde{\mathsf{b}}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{\mathsf{b}}_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\tilde{\mathsf{y}}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{\mathsf{y}}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{\mathsf{y}}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \tilde{\mathsf{y}}_4 = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$$

$$\mu_1 = 1, \qquad \mu_2 = 1, \qquad \mu_3 = 2, \qquad \mu_4 = 2,$$

$$\tilde{\mathsf{c}}_1^T = \begin{pmatrix} 1, 0 \end{pmatrix}, \quad \tilde{\mathsf{c}}_2^T = \begin{pmatrix} 0, 1 \end{pmatrix}, \quad \tilde{\mathsf{c}}_3^T = \begin{pmatrix} 1, 0 \end{pmatrix}, \quad \tilde{\mathsf{c}}_4^T = \begin{pmatrix} 0, 1 \end{pmatrix}$$

$$\tilde{\mathsf{z}}_1^T = \begin{pmatrix} -1, 1 \end{pmatrix}, \tilde{\mathsf{z}}_2^T = \begin{pmatrix} 1, -1 \end{pmatrix}, \tilde{\mathsf{z}}_3^T = \begin{pmatrix} -2, 1 \end{pmatrix}, \tilde{\mathsf{z}}_4^T = \begin{pmatrix} 1, -\frac{1}{2} \end{pmatrix}$$

The associated (block) Loewner and shifted Loewner matrices turn out to be:

$$\mathbb{L} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ 1 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{4} \end{pmatrix}, \quad \mathbb{M} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Notice that the rank of both Loewner matrices is 2 while the rank of $x_i\mathbb{L} - \mathbb{M}$ is 3, for all x equal to a σ_i or μ_i. It can be readily verified that the column span of $\sigma_1\mathbb{L} - \mathbb{M} = \mathbb{L} - \mathbb{M}$ is the same as that of Π, where

$$\Pi = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Furthermore the row span of $\mathbb{L} - \mathbb{M}$ is the same as that of Π^* Thus

$$\hat{\mathbf{A}} = -\Pi^*\mathbb{M}\Pi = \begin{pmatrix} -2 & -3 & 1 \\ -1 & 0 & -4 \\ 1 & 0 & 4 \end{pmatrix}, \quad \hat{\mathbf{E}} = -\Pi^*\mathbb{L}\Pi = \begin{pmatrix} -2 & -2 & \frac{3}{2} \\ -2 & -4 & 4 \\ \frac{3}{2} & 4 & -\frac{17}{4} \end{pmatrix},$$

$$\hat{\mathbf{B}} = \Pi^* \widetilde{\mathsf{Z}}^T = \begin{pmatrix} 0 & 0 \\ -3 & 2 \\ 3 & -\frac{5}{2} \end{pmatrix}, \quad \hat{\mathbf{C}} = \widetilde{\mathsf{Y}} \Pi = \begin{pmatrix} 2 & 3 & -1 \\ 2 & 2 & -\frac{3}{2} \end{pmatrix},$$

satisfy $\mathbf{H}(s) = \hat{\mathbf{C}}(s\hat{\mathbf{E}} - \hat{\mathbf{A}})\hat{\mathbf{B}}$, which shows that a (second) minimal realization of $\mathbf{H}$ has been obtained.

The *second* experiment involves tangential data, that is, at each interpolation point only values along certain directions are available.

$$\sigma_1 = 1, \qquad \sigma_2 = 2, \qquad \sigma_3 = 3,$$

$$\tilde{\mathsf{b}}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{\mathsf{b}}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tilde{\mathsf{b}}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\tilde{\mathsf{y}}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{\mathsf{y}}_2 = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \tilde{\mathsf{y}}_3 = \begin{pmatrix} 4 \\ \frac{4}{3} \end{pmatrix}$$

$$\mu_1 = -1, \qquad \mu_2 = -2, \qquad \mu_3 = -3,$$

$$\tilde{\mathsf{c}}_1^T = \begin{pmatrix} 1, 0 \end{pmatrix}, \quad \tilde{\mathsf{c}}_2^T = \begin{pmatrix} 0, 1 \end{pmatrix}, \quad \tilde{\mathsf{c}}_3^T = \begin{pmatrix} 1, 1 \end{pmatrix}$$

$$\tilde{\mathsf{z}}_1^T = \begin{pmatrix} -1, 1 \end{pmatrix}, \tilde{\mathsf{z}}_2^T = \begin{pmatrix} 1, -\frac{1}{2} \end{pmatrix}, \tilde{\mathsf{z}}_3^T = \begin{pmatrix} -2, \frac{2}{3} \end{pmatrix}.$$

Thus the associated Loewner and shifted Loewner matrices are:

$$\mathbb{L} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & \frac{1}{4} & \frac{1}{6} \\ 1 & \frac{1}{6} & \frac{10}{9} \end{bmatrix}, \quad \mathbb{M} = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

It readily follows that the conditions of theorem 1.9 are satisfied and hence the quadruple $(-\mathbb{M}, -\mathbb{L}, \widetilde{\mathsf{Z}}^T, \widetilde{\mathsf{Y}})$, provides a (third) minimal realization of the original rational function: $\mathbf{H}(s) = -\widetilde{\mathsf{Y}} \left[s\mathbb{L} - \mathbb{M} \right]^{-1} \widetilde{\mathsf{Z}}^T$.

7.4.2 Coupled Mechanical System

Figure 6 depicts a constrained mechanical system described in [67]. There are g masses in total; the ith mass of weight m_i, is connected to the $(i+1)$st mass by a spring and a damper with constants k_i and d_i, respectively, and also to the ground by a spring and a damper with constants κ_i and δ_i, respectively. Additionally, the first mass is connected to the last one by a rigid bar (holonomic constraint) and it is influenced by the control $\mathbf{u}(t)$. The vibration of this constrained system is described in generalized state space form as:

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t),$$

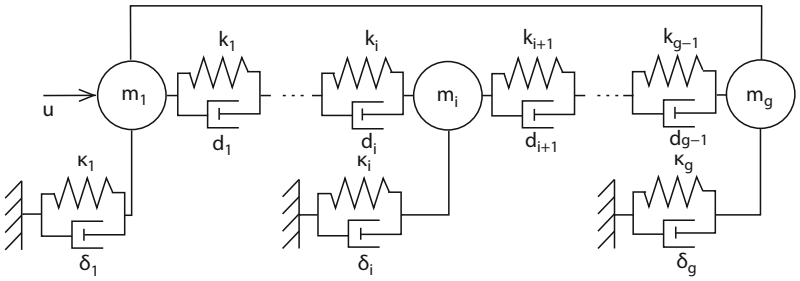


Fig. 6 Constrained mechanical system

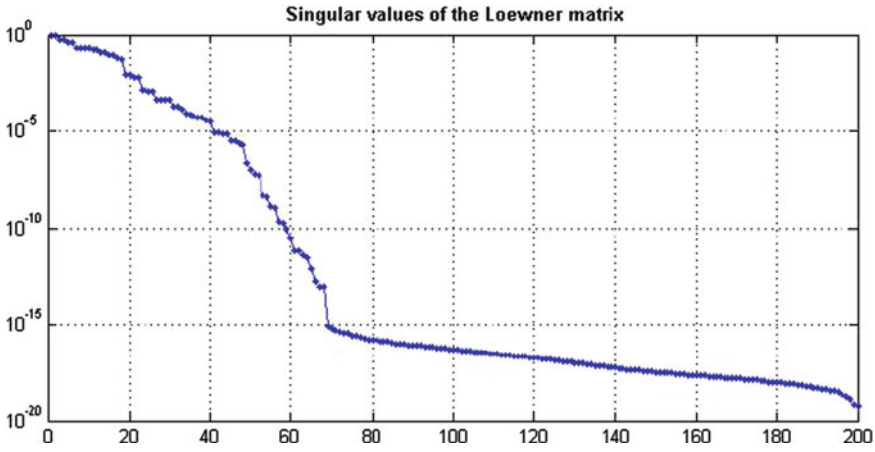


Fig. 7 The singular values of the Loewner matrix

where $\mathbf{x}$ contains the positions and velocities of the masses,

$$\mathbf{E} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{K} & \mathbf{D} & -\mathbf{G}^* \\ \mathbf{G} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_2 \\ \mathbf{0} \end{bmatrix}, \ \mathbf{C} = [\mathbf{C}_1, \ \mathbf{C}_2, \ \mathbf{C}_3];$$

furthermore $\mathbf{M}$ is the mass matrix ($g \times g$, diagonal, positive definite), $\mathbf{K}$ is the stiffness matrix ($g \times g$, tri-diagonal), $\mathbf{D}$ is the damping matrix ($g \times g$, tri-diagonal), $\mathbf{G} = [1, \ 0, \ \cdots, \ 0, \ -1]$, is the $1 \times g$ constraint matrix.

In [67], balanced truncation methods for descriptor systems are used to reduce this system. Here we will reduce this system by means of the Loewner framework. Towards this goal, we compute 200 frequency response data, that is $\mathbf{H}(i\omega_i)$, where $\omega_i \in [-2, +2]$. Figure 7 shows the singular values of the Loewner matrix pair, which indicate that a system of order 20 will have an approximate error 10^{-3} (-60 dB).

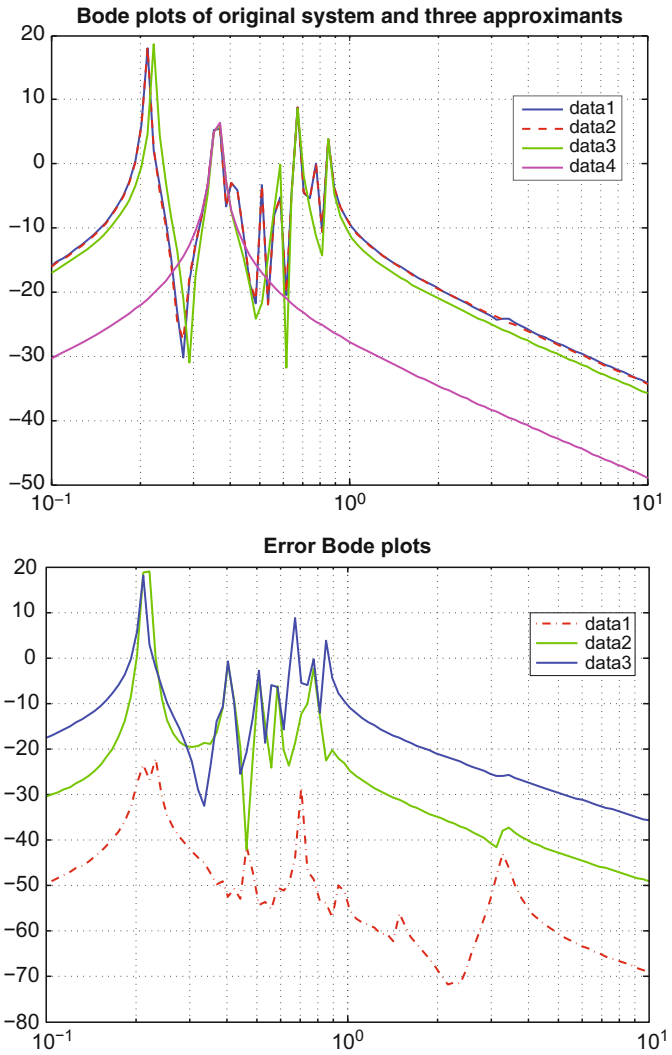


Fig. 8 *Upper pane*: Frequency responses of original system and approximants (orders 2, 10, 18). *Lower pane*: Frequency responses of error systems (orders 2,10,18)

Figure 8 shows that (for the chosen values of the parameters) the frequency response has about seven peaks. A second order approximant reproduces (approximately) the highest peak, a tenth order system reproduces (approximately) five peaks, while a system of order 18 provides a good approximation of the whole frequency response (see in particular the error plots – lower pane of Figure 8).

7.4.3 Four-pole Band-pass Filter

In this case 1,000 frequency response measurements are given, of the 2×2 S-parameters of a semi-conductor device which is meant to be a band-pass filter. There is no a priori model available. The range of frequencies is between 40 and 120 GHz; We will use the Loewner matrix procedure applied to the S-parameters. This yields $\mathbb{L}, \mathbb{M} \in \mathbb{C}^{2,000 \times 2,000}$.

In the upper left-hand plot of Fig. 9, the singular values of the Loewner matrix corresponding to the two-port system (upper curve) is compared with the singular values of two one-port subsystems (lower curves). As the decay of all curves is fast, an approximant of order around 20, is expected to provide a good fit. Indeed, as the upper right-hand plot shows, a 21st order approximant provides fits with error less than $-60\,dB$. For comparison the fit of a 15th order model is shown in the lower left-hand plot. Sometimes in practical applications, the entries of the two-port S-parameters are modeled separately. In our case 14th order models are sufficient, but the McMillan degree of the two-port is 28 or higher (depending on the symmetries involved, e.g. $S_{11} = S_{22}$, $S_{12} = S_{21}$).

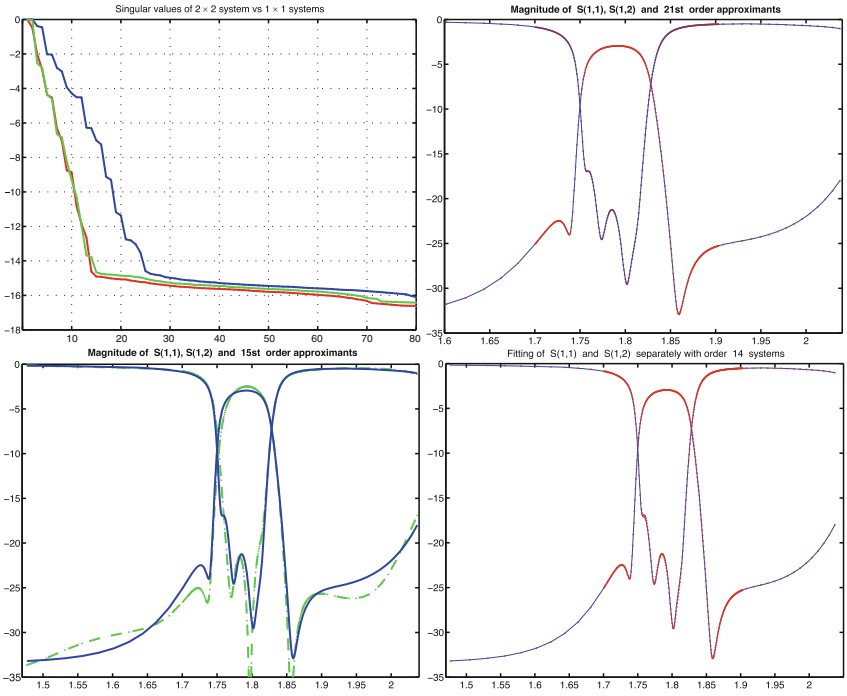


Fig. 9 *Upper row*, *left pane*: The singular values of $x\mathbb{L} - \mathbb{M}$, for the two-port and for two one-ports. *Upper row*, *right pane*: The $S(1,1)$ and $S(1,2)$ parameter data for a 21st order model. *Lower row*, *left pane*: Fitting $S(1,1)$, $S(1,2)$ jointly with a 15th order approximant. *Lower row*, *right pane*: Fitting $S(1,1)$, $S(1,2)$ separately with 14th order approximants

8 Conclusions

In this chapter, we have surveyed two projection-based model reduction frameworks both of which make fundamental use of tangential rational interpolation. This approach is extremely flexible in applications; scales well to handle extremely large problems; and is capable of producing very high fidelity reduced order models very efficiently. Throughout, examples are given that illustrate the theoretical concepts discussed.

The first framework we consider assumes the availability of a high-order dynamical system model in state space form (often these are acquired and assembled through the independent discretization and coupling of high resolution distributed parameter models). The retention of high model fidelity is directly recast as the problem of choosing appropriate interpolation points and associated tangent directions. We address this with a detailed discussion of how interpolation points and tangent directions can be be chosen to obtain reduced order models which are *optimal* with respect to $\mathcal{H}_2$ error measures (Sect. 3), or so that reduced order models are obtained which retain the property of passivity (Sect. 4).

The flexibility of our approach is demonstrated in Sects. 5 and 6 where we explore significant generalizations to the basic problem setting. We consider how to preserve second-order system structure; reduction of systems that involve delays or memory terms; and systems having a structured dependence on parameters that must be retained in the reduced systems.

The second major framework we explore allows for the original (high-order) dynamical system model to be *inaccessible*, and assumes that only system response data (e.g. frequency response measurements) are available. We describe an approach using the Loewner matrix pencil and place it in the context of interpolatory model reduction methods. We show that the Loewner matrix pencil constitutes an effective tool in obtaining minimal state-space realizations of reduced order models directly from measured data.

Acknowledgments The work of the A.C. Antoulas was supported in part by the NSF through Grant CCF-0634902. The work of C. Beattie and S. Gugercin has been supported in part by NSF Grants DMS-0505971 and DMS-0645347.

References

1. Anderson, B., Antoulas, A.: Rational interpolation and state variable realizations. In: Decision and Control, 1990. Proceedings of the 29th IEEE Conference on pp. 1865–1870 (1990)
2. Antoulas, A.: Approximation of Large-Scale Dynamical Systems (Advances in Design and Control). Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2005)
3. Antoulas, A.: A new result on passivity preserving model reduction. Systems and Control Letters **54**, 361–374 (2005)
4. Antoulas, A.: On the construction of passive models from frequency response data. Automatisierungstechnik **56**, 447–452 (2008)

5. Antoulas, A., Anderson, B.: On the scalar rational interpolation problem. IMA Journal of Mathematics Control and Information **3**, 61–68 (1986)
6. Bagley, R., Torvik, P.: A theoretical basis for the application of fractional calculus to viscoelasticity. Journal of Rheology **27**, 201–210 (1983)
7. Bai, Z.: Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. Applied Numerical Mathematics **43**(1-2), 9–44 (2002)
8. Bai, Z., Bindel, D., Clark, J., Demmel, J., Pister, K., Zhou, N.: New numerical techniques and tools in SUGAR for 3D MEMS simulation. In: Technical Proceedings of the Fourth International Conference on Modeling and Simulation of Microsystems, pp. 31–34 (2000)
9. Bai, Z., Feldmann, P., Freund, R.: Stable and passive reduced-order models based on partial Pade approximation via the Lanczos process. Numerical Analysis Manuscript **97**, 3–10 (1997)
10. Bai, Z., Feldmann, P., Freund, R.: How to make theoretically passive reduced-order models passive inpractice. In: Custom Integrated Circuits Conference, 1998. Proceedings of the IEEE 1998, pp. 207–210 (1998)
11. Bai, Z., Freund, R.: A partial Pade-via-Lanczos method for reduced-order modeling. Linear Algebra and its Applications **332**(334), 141–166 (2001)
12. Bai, Z., Skoogh, D.: A projection method for model reduction of bilinear dynamical systems. Linear Algebra and Its Applications **415**, 406–425 (2006)
13. Bai, Z., Su, Y.: Dimension reduction of second order dynamical systems via a second-order Arnoldi method. SIAM Journal on Scientific Computing **5**, 1692–1709 (2005)
14. Baur, U., Beattie, C., Benner, P., Gugercin, S.: Interpolatory projection methods for parameterized model reduction. Chemnitz Scientific Computing Preprints 09-08, TU Chemntiz, (ISSN 1864-0087) November 2009
15. Beattie, C., Gugercin, S.: Krylov-based minimization for optimal $\mathcal{H}_2$ model reduction. In: Proceedings of the 46th IEEE Conference on Decision and Control pp. 4385–4390 (2007)
16. Beattie, C., Gugercin, S.: Interpolatory projection methods for structure-preserving model reduction. Systems and Control Letters **58**(3), 225–232 (2009)
17. Beattie, C., Gugercin, S.: A trust region method for optimal $\mathcal{H}_2$ model reduction. In: Proceedings of the 48th IEEE Conference on Decision and Control (2009)
18. Benner, P.: Solving large-scale control problems. IEEE Control Systems Magazine **24**(1), 44–59 (2004)
19. Benner, P., Feng, L.: A robust algorithm for parametric model order reduction based on implicit moment matching. In B. Lohmann and A. Kugi (eds.) Tagungsband GMA-FA 1.30 "Modellbildung, Identifizierung und Simulation in der Automatisierungstechnik", Workshop in Anif, 26.–28.9.2007, pp. 34–47 (2007)
20. Benner, P., Saak, J.: Efficient numerical solution of the LQR-problem for the heat equation. Proceedings in Applied Mathematics and Mechanics **4**(1), 648–649 (2004)
21. Benner, P., Sokolov, V.: Partial realization of descriptor systems. Systems and Control Letters **55**(11), 929–938 (2006)
22. Bond, B., Daniel, L.: Parameterized model order reduction of nonlinear dynamical systems. In: IEEE/ACM International Conference on Computer-Aided Design, 2005. ICCAD-2005, pp. 487–494 (2005)
23. Bui-Thanh, T., Willcox, K., Ghattas, O.: Model reduction for large-scale systems with high-dimensional parametric input space. SIAM Journal on Scientific Computing **30**(6), 3270–3288 (2008)
24. Bungartz, H.: Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung. Dissertation, Institut fur Informatik, TU München (1992)
25. Bunse-Gerstner, A., Kubalinska, D., Vossen, G., Wilczek, D.: $\mathcal{H}_2$-optimal model reduction for large scale discrete dynamical MIMO systems. Journal of Computational and Applied Mathematics (2009). Doi:10.1016/j.cam.2008.12.029
26. Chahlaoui, V., Gallivan, K.A., Vandendorpe, A., Van Dooren, P.: Model reduction of second-order system. In: P. Benner, V. Mehrmann, D. Sorensen (eds.) Dimension Reduction of Large-Scale Systems, *Lecture Notes in Computational Science and Engineering*, vol. 45, pp. 149–172. Springer, Berlin/Heidelberg, Germany (2005)

27. Clark, J.V., Zhou, N., Bindel, D., Schenato, L., Wu, W., Demmel, J., Pister, K.S.J.: 3D MEMS simulation using modified nodal analysis. In: Proceedings of Microscale Systems: Mechanics and Measurements Symposium, p. 6875 (2000)
28. Craig Jr., R.R.: Structral Dynamics: An Introduction to Computer Methods. Wiley, New York (1981)
29. Daniel, L., Siong, O., Chay, L., Lee, K., White, J.: A Multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **23**(5), 678–693 (2004)
30. De Villemagne, C., Skelton, R.: Model reductions using a projection formulation. International Journal of Control **46**(6), 2141–2169 (1987)
31. Fanizza, G., Karlsson, J., Lindquist, A., Nagamune, R.: Passivity-preserving model reduction by analytic interpolation. Linear Algebra and Its Applications **425**(2-3), 608–633 (2007)
32. Farle, O., Hill, V., Ingelström, P., Dyczij-Edlinger, R.: Multi-parameter polynomial order reduction of linear finite element models. Mathematical and Computational Modeling of Dynamics Systems **14**(5), 421–434 (2008)
33. Feldmann, P., Freund, R.: Efficient linear circuit analysis by Padé approximation via the Lanczos process. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **14**(5), 639–649 (1995)
34. Feng, L.: Parameter independent model order reduction. Mathematics and Computers in Simulations **68**(3), 221–234 (2005)
35. Feng, L., Benner, P.: A robust algorithm for parametric model order reduction based on implicit moment matching. PAMM Proceedings of the Applied Mathematics and Mechanism **7**(1), 1021,501–1021,502 (2008)
36. Feng, L., Rudnyi, E., Korvink, J.: Preserving the film coefficient as a parameter in the compact thermal model for fast electrothermal simulation. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **24**(12), 1838–1847 (2005)
37. Freund, R., Feldmann, P.: Efficient small-signal circuit analysis and sensitivity computations with the PVL algorithm. In: Proceedings of the 1994 IEEE/ACM International Conference on Computer-aided Design, pp. 404–411. IEEE Computer Society Press Los Alamitos, CA, USA (1994)
38. Freund, R.W.: Padé-type model reduction of second-order and higher-order linear dynamical systems. In: P. Benner, V. Mehrmann, D.C. Sorensen (eds.) Dimension Reduction of Large-Scale Systems, *Lecture Notes in Computational Science and Engineering*, vol. 45, pp. 191–223. Springer, Berlin/Heidelberg (2005)
39. Gallivan, K., Vandendorpe, A., Dooren, P.: Model reduction of MIMO systems via tangential interpolation. SIAM Journal on Matrix Analysis and Applications **26**(2), 328–349 (2005)
40. Glover, K.: All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. International Journal of Control **39**(6), 1115–1193 (1984)
41. Griebel, M.: A parallelizable and vectorizable multi-level algorithm on sparse grids. In: Parallel Algorithms for Partial Differential Equations (Kiel, 1990), *Notes in Numerical Fluid Mechanics*, vol. 31, pp. 94–100. Vieweg, Braunschweig (1991)
42. Grimme, E.: Krylov projection methods for model reduction. Ph.D. thesis, Coordinated-Science Laboratory, University of Illinois at Urbana-Champaign (1997)
43. Grimme, E., Sorensen, D., Dooren, P.: Model reduction of state space systems via an implicitly restarted Lanczos method. Numerical Algorithms **12**(1), 1–31 (1996)
44. Gugercin, S.: Projection methods for model reduction of large-scale dynamical systems. Ph.D. thesis, Ph.D. Dissertation, ECE Department, Rice University, December 2002 (2002)
45. Gugercin, S.: An iterative rational Krylov algorithm (IRKA) for optimal $\mathcal{H}_2$ model reduction. In: Householder Symposium XVI. Seven Springs Mountain Resort, PA, USA (2005)
46. Gugercin, S., Antoulas, A.: An $_2$ error expression for the Lanczos procedure. In: Proceedings of the 42nd IEEE Conference on Decision and Control (2003)
47. Gugercin, S., Antoulas, A.: Model reduction of large-scale systems by least squares. Linear Algebra and its Applications **415**(2-3), 290–321 (2006)

48. Gugercin, S., Antoulas, A., Beattie, C.: A rational Krylov iteration for optimal $\mathcal{H}_2$ model reduction. In: Proceedings of MTNS, vol. 2006 (2006)
49. Gugercin, S., Antoulas, A., Beattie, C.: $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. SIAM Journal on Matrix Analysis and Applications **30**(2), 609–638 (2008)
50. Gugercin, S., Willcox, K.: Krylov projection framework for fourier model reduction. Automatica **44**(1), 209–215 (2008)
51. Gunupudi, P., Khazaka, R., Nakhla, M.: Analysis of transmission line circuits using multidimensional model reduction techniques. IEEE Transactions on Advanced Packaging **25**(2), 174–180 (2002)
52. Gunupudi, P., Khazaka, R., Nakhla, M., Smy, T., Celo, D.: Passive parameterized time-domain macromodels for high-speed transmission-line networks. IEEE Transactions on Microwave Theory and Techniques **51**(12), 2347–2354 (2003)
53. Halevi, Y.: Frequency weighted model reduction via optimal projection. IEEE Transactions on Automatic Control **37**(10), 1537–1542 (1992)
54. Hyland, D., Bernstein, D.: The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore. IEEE Transactions on Automatic Control **30**(12), 1201–1211 (1985)
55. Ionutiu, R., Rommes, J., Antoulas, A.: Passivity preserving model reduction using dominant spectral zero interpolation. IEEE Transactions on CAD (Computer-Aided Design of Integrated Circuits and Systems) **27**, 2250–2263 (2008)
56. Kellems, A., Roos, D., Xiao, N., Cox, S.J.: Low-dimensional morphologically accurate models of subthreshold membrane potential. Journal of Computational Neuroscience, **27**(2), 161–176 (2009)
57. Korvink, J., Rudnyi, E.: Oberwolfach benchmark collection. In: P. Benner, V. Mehrmann, D.C. Sorensen (eds.) Dimension Reduction of Large-Scale Systems, *Lecture Notes in Computational Science and Engineering*, vol. 45, pp. 311–315. Springer, Berlin/Heidelberg, Germany (2005)
58. Krajewski, W., Lepschy, A., Redivo-Zaglia, M., Viaro, U.: A program for solving the L2 reduced-order model problem with fixed denominator degree. Numerical Algorithms **9**(2), 355–377 (1995)
59. Kubalinska, D., Bunse-Gerstner, A., Vossen, G., Wilczek, D.: $\mathcal{H}_2$-optimal interpolation based model reduction for large-scale systems. In: Proceedings of the 16th International Conference on System Science, Poland (2007)
60. Kunkel, P.: Differential-Algebraic Equations: Analysis and Numerical Solution. European Mathematical Society, Amsterdam (2006)
61. Lasance, C.: Two benchmarks to facilitate the study of compact thermal modelingphenomena. IEEE Transactions on Components and Packaging Technologies [see also IEEE Transactions on Packaging and Manufacturing Technology, Part A: Packaging Technologies] **24**(4), 559–565 (2001)
62. Lefteriu, S., Antoulas, A.: A new approach to modeling multi-port systems from frequency domain data. IEEE Transactions on CAD (Computer-Aided Design of Integrated Circuits and Systems) **29**, 14–27 (2010)
63. Leitman, M., Fisher, G.: The linear theory of viscoelasticity. Handbuch der Physik **6**, 10–31 (1973)
64. Leung, A.M., Khazaka, R.: Parametric model order reduction technique for design optimization. In: IEEE International Symposium on Circuits and Systems, 2005. ISCAS 2005, vol. 2, pp. 1290–1293 (2005)
65. Ma, M., Leung, A.M., Khazaka, R.: Sparse Macromodels for Parametric Networks. In: Proceedings of the IEEE International Symposium on Circuits and Systems (2006)
66. Mayo, A., Antoulas, A.: A framework for the solution of the generalized realization problem. Linear Algebra and Its Applications **425**(2-3), 634–662 (2007)
67. Mehrmann, V., Stykel, T.: Balanced truncation model reduction for large-scale systems in descriptor form. In: P. Benner, V. Mehrmann, D. Sorensen (eds.) Dimension Reduction of Large-Scale Systems, pp. 83–115. Springer, Belin (2005)

68. Meier III, L., Luenberger, D.: Approximation of linear constant systems. IEEE Transactions on Automatic Control **12**(5), 585–588 (1967)
69. Meyer, D., Srinivasan, S.: Balancing and model reduction for second-order form linear systems. IEEE Transactions on Automatic Control **41**(11), 1632–1644 (1996)
70. Moore, B.: Principal component analysis in linear systems: Controllability, observability, and model reduction. IEEE Transactions on Automatic Control **26**(1), 17–32 (1981)
71. Moosmann, K., Korvink, J.: Automatic parametric mor for mems design. In: B. Lohmann, A. Kugi (eds.) Tagungsband GMA-FA 1.30 "Modellbildung, Identifikation und Simulation in der Automatisierungstechnik", Workshop am Bostalsee, 27.–29.9.2006, pp. 89–99 (2006)
72. Mullis, C., Roberts, R.: Synthesis of minimum roundoff noise fixed point digital filters. IEEE Transactions on Circuits and Systems **23**(9), 551–562 (1976)
73. Odabasioglu, A., et al.: PRIMA. IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems **17**(8) (1998)
74. Pillage, L., Rohrer, R.: Asymptotic waveform evaluation for timing analysis. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **9**(4), 352–366 (1990)
75. Preumont, A.: Vibration Control of Active Structures: An Introduction. Springer, Berlin (2002)
76. Raghavan, V., Rohrer, R., Pillage, L., Lee, J., Bracken, J., Alaybeyi, M.: AWE-inspired. In: Custom Integrated Circuits Conference, 1993. Proceedings of the IEEE 1993 (1993)
77. Rommes, J., Martins, M.: Efficient computation of multivariable transfer function dominant poles using subspace acceleration. IEEE Transactions on Power Systems **21**, 1471–1483 (2006)
78. Rudnyi, E., Moosmann, C., Greiner, A., Bechtold, T., Korvink, J.: Parameter Preserving Model Reduction for MEMS System-level Simulation and Design. In: Fifth MATHMOD Proceedings **1** (2006)
79. Ruhe, A.: Rational Krylov algorithms for nonsymmetric eigenvalue problems. II: matrix pair. Linear Algebra and Its Applications Applications, pp. 282–295 (1984)
80. Sorensen, D.: Passivity preserving model reduction via interpolation of spectral zeros. Systems and Control Letters **54**, 354–360 (2005)
81. Spanos, J., Milman, M., Mingori, D.: A new algorithm for L^2 optimal model reduction. Automatica (Journal of IFAC) **28**(5), 897–909 (1992)
82. Su, T.J., Jr., R.C.: Model reduction and control of flexible structures using Krylov vectors. Journal of Guid. Control Dyn. **14**, 260–267 (1991)
83. Van Dooren, P., Gallivan, K., Absil, P.: $\mathcal{H}_2$-optimal model reduction of MIMO systems. Applied Mathematics Letters **21**(12), 1267–1273 (2008)
84. Weaver, W., Johnston, P.: Structural dynamics by finite elements. Prentice-Hall, Upper Saddle River (1987)
85. Weile, D., Michielssen, E., Grimme, E., Gallivan, K.: A method for generating rational interpolant reduced order models of two-parameter linear systems. Applied Mathematics Letters **12**(5), 93–102 (1999)
86. Weile, D.S., Michielssen, E., Grimme, E., Gallivan, K.: A method for generating rational interpolant reduced order models of two-parameter linear systems. Applied Mathematics Letters **12**(5), 93–102 (1999)
87. Willcox, K., Megretski, A.: Fourier series for accurate, stable, reduced-order models in large-scale applications. SIAM Journal of Scientific Computing **26**(3), 944–962 (2005)
88. Wilson, D.: Optimum solution of model-reduction problem. Proceedings of IEE **117**(6), 1161–1165 (1970)
89. Yan, W., Lam, J.: An approximate approach to h^2 optimal model reduction. IEEE Transactions on Automatic Control **44**(7), 1341–1358 (1999)
90. Yousuff, A., Skelton, R.: Covariance equivalent realizations with applications to model reduction of large-scale systems. Control and Dynamic Systems **22**, 273–348 (1985)
91. Yousuff, A., Wagie, D., Skelton, R.: Linear system approximation via covariance equivalent realizations. Journal of Mathematical Analysis and Applications **106**(1), 91–115 (1985)

92. Zenger, C.: Sparse grids. In: Parallel Algorithms for Partial Differential Equations (Kiel, 1990), *Notes Numerical Fluid Mechanics*, vol. 31, pp. 241–251. Vieweg, Braunschweig (1991)
93. Zhaojun, B., Qiang, Y.: Error estimation of the Pade approximation of transfer functions via the Lanczos process. Electronic Transactions on Numerical Analysis **7**, 1–17 (1998)
94. Zhou, K., Doyle, J., Glover, K.: Robust and Optimal Control. Prentice-Hall, Upper Saddle River, NJ (1996)
95. Zigic, D., Watson, L., Beattie, C.: Contragredient transformations applied to the optimal projection equations. Linear Algebra and Its Applications **188**, 665–676 (1993)