IRKA is a Riemannian Gradient Descent Method

Petar Mlinarić [iD][1], **Christopher A. Beattie** [iD][2], **Zlatko Drmač** [iD][3], **and Serkan Gugercin** [iD][4]

[1] Department of Mathematics, Virginia Tech, Blacksburg, VA 24061 (`mlinaric@vt.edu`)
[2] Department of Mathematics, Virginia Tech, Blacksburg, VA 24061 (`beattie@vt.edu`)
[3] Faculty of Natural Sciences, University of Zagreb, 10000 Zagreb, Croatia (`drmac@math.hr`)
[4] Department of Mathematics and Division of Computational Modeling and Data Analytics, Academy of Data Science, Virginia Tech, Blacksburg, VA 24061 (`gugercin@vt.edu`)

Abstract

The iterative rational Krylov algorithm (IRKA) is a commonly used fixed point iteration developed to minimize the $\mathcal{H}_2$ model order reduction error. In this work, IRKA is recast as a Riemannian gradient descent method with a fixed step size over the manifold of rational functions having fixed degree. This interpretation motivates the development of a Riemannian gradient descent method utilizing as a natural extension variable step size and line search. Comparisons made between IRKA and this extension on a few examples demonstrate significant benefits.

Keywords linear systems, reduced order modeling, optimization, computational methods

1 Introduction

The iterative rational Krylov algorithm (IRKA) is a commonly used method for $\mathcal{H}_2$-optimal model order reduction (MOR) of linear time-invariant systems [GAB06, GAB08, ABG10]. As usually conceived, it is presented as a fixed point iteration based on interpolatory $\mathcal{H}_2$-optimality conditions; we describe it in Section 2.

Notably, there have also been approaches for $\mathcal{H}_2$ MOR via Riemannian optimization over various matrix manifolds (see, e.g., [YL99, XZ13, XJY15, SS15, BCS19]). In this work, we focus on the interpolatory IRKA framework and recast IRKA as a Riemannian optimization method over the manifold of stable rational functions of fixed McMillan degree $\Sigma_{r,m,p}^-(\mathbb{C})$, embedded in the Hardy space $\mathcal{H}_2$. This facilitates development of a geometric framework for IRKA and the interpretation of it as a Riemannian gradient descent method with fixed step size. We build on this observation, introduce a variable step size extension, and explore the numerical performance.

These key results depend on the observation that the set of stable rational functions of fixed McMillan degree $\Sigma_{r,m,p}^-(\mathbb{C})$ is an embedded submanifold in the Hardy space $\mathcal{H}_2$. We prove this while also drawing out issues related to connectedness of the manifold and the Cauchy index. An accessible reference for these facts in the continuous-time case was not known to us. This framework allows a new geometric interpretation of IRKA as a Riemannian optimization method with fixed step size. This in turn leads to the introduction of an enhanced approach utilizing a backtracking line search that preserves stability, and guarantees a decrease in the $\mathcal{H}_2$ error at each step. Significantly, the utilization of different step sizes can be interpreted as a fixed point iteration with an equivalent set of interpolatory conditions.

We begin with a background on $\mathcal{H}_2$-optimal MOR and Riemann optimization in Sections 2 and 3, respectively. We then discuss known results on the manifold structure of $\Sigma_{r,m,p}^-(\mathbb{C})$ and introduce the new result for continuous-time systems in Section 4. This allows us to interpret the $\mathcal{H}_2$-optimal MOR prob-

lem as a Riemannian optimization problem and derive the Riemannian gradient of the $\mathcal{H}_2$ error in Section 5. Then we give a geometric interpretation of necessary optimality conditions and IRKA in Section 6, leading to the interpretation of IRKA as a Riemannian gradient descent method with a fixed step size in Section 7. In Section 8, we introduce a variant of IRKA, called Riemannian Gradient Descent IRKA (RGD-IRKA), using a variable step size that enforces stability preservation and decreases the $\mathcal{H}_2$ error at every step. We compare the two methods in Section 9 and conclude with Section 10.

2 $\mathcal{H}_2$-optimal Model Order Reduction

We briefly describe IRKA and summarize relevant background and context for later discussion.

2.1 *Model Order Reduction Problem*

Given a full-order model (FOM) of order n

$$E\dot{x}(t) = Ax(t) + Bu(t), \ x(0) = 0,$$
$$y(t) = Cx(t),$$

the goal is to find a reduced-order model (ROM) of order r

$$\widehat{E}\dot{\widehat{x}}(t) = \widehat{A}\widehat{x}(t) + \widehat{B}u(t), \ \widehat{x}(0) = 0,$$
$$\widehat{y}(t) = \widehat{C}\widehat{x}(t),$$

such that $r \ll n$ and the outputs $y(t), \widehat{y}(t) \in \mathbb{C}^p$ are close to one another when the systems are presented with the same input $u(t) \in \mathbb{C}^m$. Here, $x(t) \in \mathbb{C}^n$ is the full-order state and $\widehat{x}(t) \in \mathbb{C}^r$ is the reduced-order state; $E, A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$, $C \in \mathbb{C}^{p \times n}$, $\widehat{E}, \widehat{A} \in \mathbb{C}^{r \times r}$, $\widehat{B} \in \mathbb{C}^{r \times m}$, and $\widehat{C} \in \mathbb{C}^{p \times r}$. We assume E and $\widehat{E}$ are invertible and that $E^{-1}A$ and $\widehat{E}^{-1}\widehat{A}$ are Hurwitz (their eigenvalues are in the open left half-plane). We develop and present theory for general complex systems. The case of real systems follow directly from our discussion and we add remarks when necessary to highlight the real case.

To measure the distance between y and $\widehat{y}$, we use the transfer functions H and $\widehat{H}$ of the FOM and ROM, respectively:

$$H(s) = C(sE - A)^{-1}B \quad \text{and} \quad \widehat{H}(s) = \widehat{C}\left(s\widehat{E} - \widehat{A}\right)^{-1}\widehat{B}.$$

The $\mathcal{H}_2$ error is defined as

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

A uniform bound on the output error can be formulated:

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

This motivates reducing the $\mathcal{H}_2$ error, thus assuring the output error $\| y - \widehat{y} \|_{\mathcal{L}_\infty}$ to be small for bounded input energy. This leads one to the $\mathcal{H}_2$-optimal MOR problem

$$\underset{\substack{\widehat{H} \text{ of order } r \\ \widehat{H} \text{ stable}}}{\text{minimize}} \quad \left\| H - \widehat{H} \right\|_{\mathcal{H}_2}, \tag{2.1}$$

where stability of $\widehat{H}$ means that its poles are in the open left half-plane. Our focus here is on interpolatory and $\mathcal{H}_2$-optimal MOR. For further approaches, see, e.g., [ZDG96, Ant05, BOCW17, ABG20, BGTQ+21a, BGTQ+21b].

2.2 Hardy Space

Under the assumptions made for the FOM and ROM, both H and $\widehat{H}$ are elements of the Hardy space $\mathcal{H}_2 = \mathcal{H}_2^{p\times m}(\mathbb{C}^+)$ (see, e.g., [Ant05, Section 5.1.3]), where $\mathbb{C}^+$ denotes the open right half-plane. $\mathcal{H}_2$ is a Hilbert space of analytic functions $F \colon \mathbb{C}^+ \to \mathbb{C}^{p\times m}$ satisfying $\sup_{\xi>0} \int_{-\infty}^{\infty} \| F(\xi + \iota\omega) \|_{\mathrm{F}}^2 \, \mathrm{d}\omega < \infty$ with the inner product

$$\langle F, G \rangle_{\mathcal{H}_2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathsf{trace}(F(\iota\omega)^* G(\iota\omega)) \, \mathrm{d}\omega,$$

using extensions of F and G from $\mathbb{C}^+$ to $\overline{\mathbb{C}^+}$. We have particular interest in the real Hilbert space of real $p\times m$ transfer functions, $\mathfrak{R}\mathcal{H}_2 \subset \mathcal{H}_2$,

$$\mathfrak{R}\mathcal{H}_2 = \left\{ H \in \mathcal{H}_2 : \overline{H(\overline{s})} = H(\overline{s}), \forall s \in \mathbb{C}^+ \right\}.$$

2.3 Interpolatory Necessary Conditions for Optimality

The following theorem gives necessary $\mathcal{H}_2$-optimality conditions when the ROM has simple poles (see, e.g., [ML67, ABG10]).

Theorem 2.1. *Let $H, \widehat{H} \in \mathfrak{R}\mathcal{H}_2$ be such that $\widehat{H}$ has the pole-residue form*

$$\widehat{H}(s) = \sum_{i=1}^{r} \frac{c_i b_i^*}{s - \lambda_i},$$

where λ_i are pairwise distinct. Let $\widehat{H}$ be an $\mathcal{H}_2$-optimal approximation for H as in (2.1). Then for $i = 1, 2, \ldots, r$

$$H\left(-\overline{\lambda_i}\right) b_i = \widehat{H}\left(-\overline{\lambda_i}\right) b_i, \tag{2.2a}$$

$$c_i^* H\left(-\overline{\lambda_i}\right) = c_i^* \widehat{H}\left(-\overline{\lambda_i}\right), \tag{2.2b}$$

$$c_i^* H'\left(-\overline{\lambda_i}\right) b_i = c_i^* \widehat{H}'\left(-\overline{\lambda_i}\right) b_i. \tag{2.2c}$$

That is, the $\mathcal{H}_2$-optimal ROMs are bitangential Hermite interpolants at reflected poles of the ROM in tangent directions given by the (vector) residues of the ROM. This theorem directly extends from $\mathfrak{R}\mathcal{H}_2$ to $\mathcal{H}_2$ as well. Extension to higher-order poles is possible [VGA10], but we focus on simple poles in the theoretical analysis for simplicity.

2.4 Bitangential Hermite Interpolation

Given interpolation points $\sigma_i \in \mathbb{C}$ and tangential directions $b_i \in \mathbb{C}^m$, $c_i \in \mathbb{C}^p$, for $i = 1, 2, \ldots, r$, bitangential Hermite interpolants (as required for $\mathcal{H}_2$-optimality) can be constructed via Petrov-Galerkin projection based on rational Krylov subspaces [GAB06]: For $H(s) = C(sE - A)^{-1}B$, define projection matrices, $V, W \in \mathbb{C}^{n\times r}$, such that

$$\mathsf{Im}(V) = \mathsf{span}\left\{(\sigma_i E - A)^{-1} B b_i\right\}_{i=1}^{r}, \tag{2.3a}$$

$$\mathsf{Im}(W) = \mathsf{span}\{(\sigma_i E - A)^{-*} C^* c_i\}_{i=1}^{r} \tag{2.3b}$$

($\mathsf{Im}(M)$ denotes the *image* or *range* of M). Then a ROM is constructed as

$$\widehat{E} = W^* E V, \ \widehat{A} = W^* A V, \ \widehat{B} = W^* B, \ \widehat{C} = C V. \tag{2.4}$$

For $\widehat{H}(s) = \widehat{C}(s\widehat{E} - \widehat{A})^{-1}\widehat{B}$, assuming invertibility where needed, we have for $i = 1, 2, \ldots, r$ [ABG10]:

$$H(\sigma_i) b_i = \widehat{H}(\sigma_i) b_i,$$

$$c_i^* H(\sigma_i) = c_i^* \widehat{H}(\sigma_i),$$

$$c_i^* H'(\sigma_i) b_i = c_i^* \widehat{H}'(\sigma_i) b_i.$$

Note that V and W can be computed from solutions to Sylvester equations as we discuss next. Given interpolation data $\{\sigma_i, b_i, c_i\}$, let the ith column of V be $(\sigma_i E - A)^{-1} B b_i$. This construction satisfies (2.3a). Then it directly follows that V satisfies the Sylvester equation

$$AV - EV\Sigma + B\widetilde{R}^* = 0,$$

where $\Sigma = \mathsf{diag}(\sigma_i)$ and the ith row of $\widetilde{R}$ is b_i^*. Then, for any invertible $T, \widetilde{E} \in \mathbb{C}^{r\times r}$, we have

$$\left(AV - EV\Sigma + B\widetilde{R}^*\right) T \widetilde{E}^* = 0.$$

Therefore,

$$A(VT)\widetilde{E}^* + E(VT)\left(-T^{-1}\Sigma T \widetilde{E}^*\right) + B\left(\widetilde{R}^* T \widetilde{E}^*\right) = 0.$$

With $\widetilde{V} = VT$, $\widetilde{A} = -\widetilde{E}T^*\Sigma^* T^{-*}$, and $\widetilde{B} = \widetilde{E}T^*\widetilde{R}$, we find

$$A\widetilde{V}\widetilde{E}^* + E\widetilde{V}\widetilde{A}^* + B\widetilde{B}^* = 0. \tag{2.5}$$

Note that $\mathsf{Im}(\widetilde{V}) = \mathsf{Im}(V)$ and $\widetilde{E}^{-1}\widetilde{A}$ has $-\overline{\sigma_i}$ as eigenvalues. Analogous constructions hold for W, thus, as claimed, both V and W can be constructed from solving Sylvester equations of the form (2.5). This fact will be employed in the algorithmic development in Section 8. Note that solving (2.5) does not require that $\widetilde{E}^{-1}\widetilde{A}$ has simple eigenvalues [BKS11].

Assuming that the interpolation data is closed under conjugation, V and W can be chosen to be real so that the ROM is also real. See [ABG20] for further details.

2.5 The IRKA Iteration

Construction of V and W above assumed the knowledge of interpolation data. However, as shown in (2.2), the interpolation data for $\mathcal{H}_2$ optimality depend on the optimal ROM to be constructed. Thus, satisfying (2.2) requires an iterative process. This observation, together with the interpolatory projection framework, led to the development of the *iterative rational Krylov algorithm (IRKA)* [GAB06, GAB08, ABG10], a

fixed point iteration algorithm that, given an iterate $\widehat{H}_k(s) = \sum_{i=1}^{r} \frac{c_i^{(k)}(b_i^{(k)})^*}{s-\lambda_i^{(k)}}$, determines the next iterate $\widehat{H}_{k+1}$ from the interpolation conditions, for $i = 1, 2, \ldots, r$:

$$H\left(-\overline{\lambda_i^{(k)}}\right)b_i^{(k)} = \widehat{H}_{k+1}\left(-\overline{\lambda_i^{(k)}}\right)b_i^{(k)}, \tag{2.6a}$$

$$\left(c_i^{(k)}\right)^* H\left(-\overline{\lambda_i^{(k)}}\right) = \left(c_i^{(k)}\right)^* \widehat{H}_{k+1}\left(-\overline{\lambda_i^{(k)}}\right), \tag{2.6b}$$

$$\left(c_i^{(k)}\right)^* H'\left(-\overline{\lambda_i^{(k)}}\right)b_i^{(k)} = \left(c_i^{(k)}\right)^* \widehat{H}'_{k+1}\left(-\overline{\lambda_i^{(k)}}\right)b_i^{(k)}. \tag{2.6c}$$

Using Petrov-Galerkin projection as described in 2.4, these conditions uniquely determine a rational transfer function of order r. Then, this fixed point iteration, i.e., IRKA, is run until convergence upon which the ROM satisfies the interpolatory conditions (2.2). For details of IRKA, we refer the reader to [GAB08, ABG20]. Despite its success in practice and its usage across numerous applications, since IRKA is a fixed point iteration, the convergence and stability of the ROM are not theoretically guaranteed. In this work, by recasting IRKA in a Riemannian optimization setting, we develop efficient variants that aim to resolve these issues.

3 Riemannian Optimization Background

We summarize here a few topics from Riemannian optimization (see also [AMS08, Bou23]), focusing, in particular, on Riemannian submanifolds (see, e.g., [Lan95, Kli11, Lee12]), that will be used heavily in the rest of the paper.

3.1 General Manifolds

A *(smooth) manifold* $\mathcal{M} = (M, \mathcal{A})$, modeled on a Banach space $\mathbb{E}$, is a set M together with an *atlas* $\mathcal{A}$, which is a collection of *charts* (U_i, φ_i) (i is an element of an index set) such that

1. each U_i is a subset of M and the U_i cover M,
2. each φ_i is a bijection from U_i to an open subset $\varphi(U_i)$ of $\mathbb{E}$ and $\varphi_i(U_i \cap U_j)$ is open in $\mathbb{E}$ for any i, j,
3. the map $\varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is smooth (infinitely differentiable) for each pair of indices i, j.

We additionally assume that the atlas topology defined by the maximal atlas $\mathcal{A}^+$ (the collection of all charts compatible with elements of $\mathcal{A}$) is Hausdorff and second-countable. As is common in the literature, we conflate $\mathcal{M}$ with M and (U_i, φ_i) with φ_i in the following.

The *dimension* of a manifold, denoted by $\dim(\mathcal{M})$, is defined to be $\dim(\mathbb{E})$. We are interested in both finite and infinite-dimensional manifolds. In particular, note that the Hardy space $\mathcal{H}_2$ is a manifold with $\mathbb{E} = \mathcal{H}_2$ and the identity map $\mathrm{id}\colon \mathcal{H}_2 \to \mathcal{H}_2$ as the chart. Additionally, we are interested in embedded submanifolds (which we discuss in Section 3.3). But, first, we need a general concept of a smooth function and its differential.

3.2 General Tangent Space and Smooth Functions

A function $F\colon \mathcal{M} \to \mathcal{N}$ from a manifold $\mathcal{M}$ to a manifold $\mathcal{N}$ is *smooth at p* if $\psi \circ F \circ \varphi^{-1}\colon \varphi(U) \to \psi(V)$ is smooth for a chart (U, φ) of $\mathcal{M}$ and a chart (V, ψ) of $\mathcal{N}$ such that $p \in U$ and $F(p) \in V$. A function is *smooth* if it is smooth at every point.

For a manifold $\mathcal{M}$ and point $p \in \mathcal{M}$, let C be the set of smooth curves $c\colon I \to \mathcal{M}$ such that $c(0) = p$ where $I \subseteq \mathbb{R}$ is an interval around zero. Define an equivalence relation $\sim$ between two curves c_1, c_2 by

$$c_1 \sim c_2 \quad \Leftrightarrow \quad (\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0)$$

for some chart φ of $\mathcal{M}$ around p. The *tangent vectors* of $\mathcal{M}$ at p are then equivalence classes $[c]$ under the equivalence relation $\sim$. The *tangent space* of $\mathcal{M}$ at p, denoted $\mathrm{T}_p\mathcal{M}$, is the set of all tangent vectors at p, i.e., $\mathrm{T}_p\mathcal{M} = C/\sim$. A tangent space is a vector space of the same dimension as the manifold.

The *differential* of a smooth function $F\colon \mathcal{M} \to \mathcal{N}$ at $p \in \mathcal{M}$ is the linear operator $DF(p)\colon \mathrm{T}_p\mathcal{M} \to \mathrm{T}_{F(p)}\mathcal{N}$, defined by $DF(p)[v] = [F \circ c]$, where c is a representative curve, i.e., $v = [c] \in \mathrm{T}_p\mathcal{M}$. Note that the brackets in $[v]$ are only used to denote that $DF(p)$ is applied to v.

3.3 Embedded Submanifolds

A *smooth immersion* is a smooth mapping such that its differential is injective at each point and the image of the differential is a closed subspace. A *smooth embedding* is a smooth immersion that is also a homeomorphism onto its image, i.e., its inverse is continuous.

An *embedded submanifold* of $\overline{\mathcal{M}}$ is a subset $\mathcal{M} \subseteq \overline{\mathcal{M}}$ that is a manifold in the subspace topology, endowed with a smooth structure with respect to which the inclusion map $\mathcal{M} \hookrightarrow \overline{\mathcal{M}}$ is a smooth embedding. $\overline{\mathcal{M}}$ is called an *ambient space*.

A function $F\colon \mathcal{M} \to \mathcal{N}$ between embedded submanifolds is smooth if and only if F can be extended to a smooth function over an open neighborhood of $\mathcal{M}$ in its ambient space.

3.4 Tangent Vectors, Tangent Spaces, Tangent Bundle, and Normal Spaces of Embedded Submanifolds

Let $\mathcal{M}$ be an embedded submanifold of a Banach space $\mathbb{E}$. The tangent space of $\mathcal{M}$ at $p \in \mathcal{M}$ can be identified with the set of velocities at p of all smooth curves in $\mathcal{M}$ going through p, i.e., $\mathrm{T}_p\mathcal{M} \simeq \{c'(0) \mid c\colon I \to \mathcal{M}, \ c(0) = p\}$, where $I \subseteq \mathbb{R}$ is any interval containing zero. In the following, we always interpret the tangent space of an embedded submanifold of a Banach space as a subset of the ambient space. Then, for every $p \in \mathcal{M}$, $\mathrm{T}_p\mathcal{M}$ forms a subspace of $\mathbb{E}$ of the same dimension as the dimension of the manifold.

The *tangent bundle*, denoted by $\mathrm{T}\mathcal{M}$, is a disjoint union of all tangent spaces of $\mathcal{M}$, in particular, $\mathrm{T}\mathcal{M} = \{(p, v) \mid p \in \mathcal{M} \text{ and } v \in \mathrm{T}_p\mathcal{M}\}$. The tangent bundle is itself a manifold of dimension $2\dim(\mathcal{M})$.

If $\mathbb{E}$ is additionally a Hilbert space, the *normal space* of $\mathcal{M}$ at $p \in \mathcal{M}$, denoted $\mathrm{N}_p\mathcal{M}$, is the orthogonal complement of the tangent space at p in $\mathbb{E}$, i.e., $\mathrm{N}_p\mathcal{M} = (\mathrm{T}_p\mathcal{M})^\perp$.

A *scalar field* over $\mathcal{M}$ is a function $f\colon \mathcal{M} \to \mathbb{R}$. A *vector field* over $\mathcal{M}$ is a function $X\colon \mathcal{M} \to \mathrm{T}\mathcal{M}$ such that $X(p) \in \mathrm{T}_p\mathcal{M}$ for all $p \in \mathcal{M}$.

3.5 Riemannian Manifolds and Submanifolds

A *metric* over $\mathcal{M}$ is a collection of inner products $\langle \cdot, \cdot \rangle_p$ over $T_p\mathcal{M}$ for all $p \in \mathcal{M}$. A *Riemannian metric* is a metric $\langle \cdot, \cdot \rangle_p$ such that $p \mapsto \langle X(p), Y(p) \rangle_p$ is a smooth scalar field for all smooth vector fields X, Y. A *Riemannian manifold* is a manifold with a Riemannian metric.

A *Riemannian submanifold* is an embedded submanifold that inherits the Riemannian metric from the ambient manifold. Note that it is always possible for an embedded submanifold of a Riemannian manifold to inherit the metric because, if $\mathcal{M} \subseteq \overline{\mathcal{M}}$, then $T_p\mathcal{M} \subseteq T_p\overline{\mathcal{M}}$ for all $p \in \mathcal{M}$.

3.6 Riemannian Optimization

Now we have all the necessary background to summarize the basics of Riemannian optimization. A *Riemannian optimization problem* is an optimization problem of the form

$$\underset{p \in \mathcal{M}}{\text{minimize}}\ f(p),$$

where $\mathcal{M}$ is a Riemannian manifold and f is a smooth scalar field. The *gradient* at $p \in \mathcal{M}$ of a smooth scalar field f on a Riemannian manifold $\mathcal{M}$ is the unique tangent vector $\operatorname{grad} f(p)$ in $T_p\mathcal{M}$ such that

$$Df(p)[v] = \langle \operatorname{grad} f(p), v \rangle_p$$

for all $v \in T_p\mathcal{M}$. If $\mathcal{M}$ is a Riemannian submanifold of a Hilbert space $\mathbb{E}$, the gradient can be obtained via

$$\operatorname{grad} f(p) = \operatorname{Proj}_p\big(\nabla \overline{f}(p)\big), \tag{3.1}$$

where $\operatorname{Proj}_p \colon \mathbb{E} \to T_p\mathcal{M}$ is the orthogonal projector onto $T_p\mathcal{M}$, $\overline{f}$ is a smooth extension of f to an open neighborhood of $\mathcal{M}$ in $\mathbb{E}$, and $\nabla \overline{f}(p)$ is the Euclidean gradient of $\overline{f}$ at p.

As for Euclidean optimization, the first-order necessary optimality condition for p to be a local minimum is

$$\operatorname{grad} f(p) = 0, \tag{3.2}$$

If $\mathcal{M}$ is an Riemannian submanifold, using (3.1), we have that the necessary optimality condition (3.2) is equivalent to

$$\nabla \overline{f}(p) \perp T_p\mathcal{M}. \tag{3.3}$$

A step of Riemannian gradient descent takes the form

$$p_{k+1} = R_{p_k}(-\alpha_k \operatorname{grad} f(p_k)), \tag{3.4}$$

where $R_{p_k} \colon T_{p_k}\mathcal{M} \to \mathcal{M}$ is a *retraction* and α_k are positive. In simple terms, a step of Riemannian gradient descent starts from the current iterate p_k, moves in the direction of the negative gradient at p_k, and then returns to the manifold $\mathcal{M}$ via retraction.

Generally, a retraction R is a mapping from the tangent bundle $T\mathcal{M}$ to the manifold $\mathcal{M}$ such that R_p, the restriction of R to $T_p\mathcal{M}$, is "locally rigid" around zero, i.e.,

$$R_p(0) = p \ \text{ and } \ DR_p(0) = \operatorname{id},$$

where $\operatorname{id} \colon T_p\mathcal{M} \to T_p\mathcal{M}$ is the identity map. The definition of a retraction can be relaxed so as to be defined only over an open subset of the tangent bundle containing all of the tangent zero vectors i.e., containing $(p, 0)$ for all $p \in \mathcal{M}$. We employ these Riemannian optimization concepts, especially (3.1)–(3.4), in the theoretical and algorithmic developments presented in the following sections.

4 RATIONAL FUNCTIONS FORM A MANIFOLD

Let $\Sigma^-_{r,m,p}(\mathbb{C})$ be the set of stable rational functions:

$$\Sigma^-_{r,m,p}(\mathbb{C}) = \left\{ \widehat{C}\big(sI - \widehat{A}\big)^{-1}\widehat{B} : \begin{matrix} \widehat{A} \in \mathbb{C}^{r \times r} \text{ is Hurwitz,} \\ \widehat{B} \in \mathbb{C}^{r \times m},\ \widehat{C} \in \mathbb{C}^{p \times r}, \\ \big(\widehat{A}, \widehat{B}, \widehat{C}\big) \text{ is minimal} \end{matrix} \right\},$$

where minimality means that the realization $(\widehat{A}, \widehat{B}, \widehat{C})$ is of the least possible order. We know that $\Sigma^-_{r,m,p}(\mathbb{C})$ is a subset of $\mathcal{H}_2$. In this section, we prove that $\Sigma^-_{r,m,p}(\mathbb{C})$ is an embedded submanifold of $\mathcal{H}_2$, an important result employed in later sections. Additionally, we show that $\Sigma^-_{r,m,p}(\mathbb{R})$, the set of real stable rational functions, is an embedded submanifold of $\mathfrak{R}\mathcal{H}_2$.

4.1 Previous Work on Manifolds of Rational Functions

Let $S^-_{r,m,p}(\mathbb{R})$ be the set of stable, minimal state-space realizations, i.e., $S^-_{r,m,p}(\mathbb{R}) \subset \mathbb{R}^{r \times r} \times \mathbb{R}^{r \times m} \times \mathbb{R}^{p \times r}$ and $(\widehat{A}, \widehat{B}, \widehat{C}) \in S^-_{r,m,p}(\mathbb{R})$ where $\widehat{A}$ is Hurwitz and $(\widehat{A}, \widehat{B}, \widehat{C})$ is minimal. Similarly, we use $\Sigma_{r,m,p}(\mathbb{R})$ and $S_{r,m,p}(\mathbb{R})$ to denote the sets corresponding to, respectively, $\Sigma^-_{r,m,p}(\mathbb{R})$ and $S^-_{r,m,p}(\mathbb{R})$ but without the stability assumptions.

For the single-input single-output (SISO) case ($m = p = 1$), Brockett [Bro76] identified the elements of $\Sigma_{r,m,p}(\mathbb{R})$ by the $2r$ coefficients in the "polynomial over monic polynomial" form and showed that this set forms an embedded submanifold of $\mathbb{R}^{2r}$ consisting of $r + 1$ connected components. As a corollary to [Bro76, Theorem 1], one obtains the same result for stable rational functions. Additionally, we find that $\Sigma^-_{r,m,p}(\mathbb{C})$ is connected for $m = p = 1$.

Hazewinkel [Haz77, Theorem 2.5.17] constructed a manifold structure over the quotient set $S_{r,m,p}(\mathbb{R})/\sim$ (based on Hazewinkel and Kalman [HK76]) where $\sim$ is the state-space equivalence relation, i.e., $(\widehat{A}_1, \widehat{B}_1, \widehat{C}_1) \sim (\widehat{A}_2, \widehat{B}_2, \widehat{C}_2)$ if and only if there is an invertible matrix $T \in \mathbb{R}^{r \times r}$ such that

$$\widehat{A}_1 = T\widehat{A}_2 T^{-1}, \ \ \widehat{B}_1 = T\widehat{B}_2, \ \ \widehat{C}_1 = \widehat{C}_2 T^{-1}.$$

Delchamps asserts $\Sigma_{r,m,p}(\mathbb{R})$ (and $\Sigma_{r,m,p}(\mathbb{C})$) are analytic manifolds of (complex) dimension $r(m + p)$ ([Del85, Theorem 2.1]; Delchamps attributes the result to Clark [Cla66]). Byrnes and Duncan [BD82] show in the proof of Proposition A.7 that $\Sigma_{r,m,p}(\mathbb{R})$ is connected when $\max(m, p) > 1$. The proof can be extended to stable rational functions in $\Sigma^-_{r,m,p}(\mathbb{R})$ and to complex rational functions in $\Sigma^-_{r,m,p}(\mathbb{C})$.

4.2 Stable Systems Form an Embedded Submanifold

Our first main result is to establish the manifold structure of $\Sigma^-_{r,m,p}(\mathbb{C})$ as an embedded submanifold of the Hardy space $\mathcal{H}_2$ (in contrast to earlier results which establish a manifold structure over different spaces).

Theorem 4.1. *The manifold $\Sigma^-_{r,m,p}(\mathbb{C})$ ($\Sigma^-_{r,m,p}(\mathbb{R})$) is an embedded submanifold of $\mathcal{H}_2$ ($\mathfrak{R}\mathcal{H}_2$).*

Proof. Exploiting the ideas from the proof of Theorem 2.1 in [ABG94] for discrete-time dynamical systems, we will prove the result by showing that the natural inclusion $\mathbb{I} \colon \Sigma^-_{r,m,p}(\mathbb{C}) \to$

$\mathcal{H}_2$ is a smooth embedding where the manifold structure of $\Sigma_{r,m,p}^-(\mathbb{C})$ is from Hazewinkel and Kalman [HK76]. We show that $\mathbb{I}$ is an embedding, i.e.,

1. $\mathbb{I}$ is differentiable,

2. the differential $\mathrm{D}\mathbb{I}(\widehat{H})$ is injective and its image $\mathsf{Im}(\mathrm{D}\mathbb{I}(\widehat{H}))$ is closed for all $\widehat{H} \in \Sigma_{r,m,p}^-(\mathbb{C})$, and

3. $\mathbb{I}^{-1}$ is continuous.

Define $\Pi\colon S_{r,m,p}^-(\mathbb{C}) \rightarrow \Sigma_{r,m,p}^-(\mathbb{C})$ by $\Pi(\widehat{A},\widehat{B},\widehat{C})[s] = \widehat{C}(sI-\widehat{A})^{-1}\widehat{B}$. Using the charts $\varphi_i\colon U_i \rightarrow S_{r,m,p}^-(\mathbb{C})$ of $\Sigma_{r,m,p}^-(\mathbb{C})$ from [HK76], $\mathbb{I}$ can be locally expressed as $\mathbb{I}\circ\Pi\circ\varphi_i$. Since φ_i is smooth by definition, we need only show that $\mathbb{I}\circ\Pi\colon S_{r,m,p}^-(\mathbb{C}) \rightarrow \mathcal{H}_2$ is smooth. We see that

$$\mathrm{D}_{\widehat{A}}(\mathbb{I}\circ\Pi)(\widehat{A},\widehat{B},\widehat{C})[X] = -\widehat{C}(sI-\widehat{A})^{-1}X(sI-\widehat{A})^{-1}\widehat{B},$$
$$\mathrm{D}_{\widehat{B}}(\mathbb{I}\circ\Pi)(\widehat{A},\widehat{B},\widehat{C})[W] = \widehat{C}(sI-\widehat{A})^{-1}W,$$
$$\mathrm{D}_{\widehat{C}}(\mathbb{I}\circ\Pi)(\widehat{A},\widehat{B},\widehat{C})[V] = V(sI-\widehat{A})^{-1}\widehat{B},$$

therefore, $\mathbb{I}$ is smooth, proving the first assertion.

For the second assertion, note that $\mathsf{Im}(\mathrm{D}\mathbb{I}(\widehat{H}))$ consists of rational matrix-valued functions having the form

$$V(sI-\widehat{A})^{-1}\widehat{B} + \widehat{C}(sI-\widehat{A})^{-1}W$$
$$- \widehat{C}(sI-\widehat{A})^{-1}X(sI-\widehat{A})^{-1}\widehat{B}$$

for arbitrary $X \in \mathbb{C}^{n\times n}$, $W \in \mathbb{C}^{n\times m}$, and $V \in \mathbb{C}^{p\times n}$. Since the domain of $\mathrm{D}\mathbb{I}(\widehat{H})$ is $\mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C})$, it follows that $\dim\mathsf{Im}(\mathrm{D}\mathbb{I}(\widehat{H})) \leqslant \dim\Sigma_{r,m,p}^-(\mathbb{C}) = r(m+p)$. Using the same procedure as in [ABG94], we can find a linearly independent set with $r(m+p)$ elements, from which we conclude that $\mathrm{D}\mathbb{I}(\widehat{H})$ is injective. The image of $\mathrm{D}\mathbb{I}(\widehat{H})$ is closed since it is finite-dimensional, which proves the second assertion on $\mathbb{I}$.

For the third assertion, we show that $\mathbb{I}^{-1}$ is continuous by showing the convergence of Markov parameters induced by convergence of $(\widehat{H}_k)$ in $\mathcal{H}_2$ for continuous-time dynamical system. Let $(\widehat{H}_k)$ be a convergent sequence in the image of $\mathbb{I}$ with $\widehat{H}$ as the limit. Let $M_i^{(k)}$ and M_i denote the ith Markov parameter of $\widehat{H}_k$ and $\widehat{H}$, respectively. The Markov parameters $M_i^{(k)}$ and M_i can be determined, via contour integrals, as

$$M_i^{(k)} = \frac{1}{2\pi\iota}\oint_{\Gamma_R} s^i \widehat{H}_k(s)\,\mathrm{d}s, \quad M_i = \frac{1}{2\pi\iota}\oint_{\Gamma_R} s^i \widehat{H}(s)\,\mathrm{d}s,$$

where $\Gamma_R = [-\iota R, \iota R] \cup \{Re^{\iota\omega} : \omega \in (\frac{\pi}{2}, \frac{3\pi}{2})\}$ and $R > 0$ is big enough such that Γ_R encloses all of the poles of $\widehat{H}_k$ and $\widehat{H}$, for all $k \geqslant 1$. Then we have that

$$\left\| M_i^{(k)} - M_i \right\|_2$$
$$\leqslant \frac{1}{2\pi}\int_a^b |\gamma(t)|^i \left\| \widehat{H}_k(\gamma(t)) - \widehat{H}(\gamma(t)) \right\|_2 |\gamma'(t)|\,\mathrm{d}t,$$

where $\gamma\colon [a,b] \rightarrow \mathbb{C}$ is a parametrization of Γ_R. The Cauchy-Schwarz inequality gives

$$\left\| M_i^{(k)} - M_i \right\|_2 \leqslant \frac{1}{2\pi}\left(\int_a^b |\gamma(t)|^{2i}|\gamma'(t)|^2\,\mathrm{d}t\right)^{1/2}$$
$$\times \left(\int_a^b \left\| \widehat{H}_k(\gamma(t)) - \widehat{H}(\gamma(t)) \right\|_2^2\,\mathrm{d}t\right)^{1/2}.$$

Since convergence of the sequence of rational functions $\widehat{H}_k$ of bounded degree in $\mathcal{H}_2$ implies $\mathcal{L}_2$ convergence on Γ_R, we conclude $M_i^{(k)}$ converges to M_i.

Finally, it follows similarly that $\Sigma_{r,m,p}^-(\mathbb{R})$ is an embedded submanifold of $\mathfrak{R}\mathcal{H}_2$. $\qquad\square$

5 $\mathcal{H}_2$ Minimization as a Riemannian Optimization

Theorem 4.1 shows that $\Sigma_{r,m,p}^-(\mathbb{C})$ is an embedded submanifold of $\mathcal{H}_2$. Building on that result, we now give an interpretation of the $\mathcal{H}_2$-optimal MOR problem (2.1) as a Riemannian optimization problem. We then compute the Riemannian gradient in this setting, which will be used to derive necessary optimality conditions and facilitate the interpretation of IRKA as a Riemannian gradient descent method.

First, we establish $\Sigma_{r,m,p}^-(\mathbb{C})$ as a Riemannian manifold. Clearly, it is possible to define a Riemannian metric over $\Sigma_{r,m,p}^-(\mathbb{C})$ by inheriting the inner product from the ambient $\mathcal{H}_2$ space. In this way, $\Sigma_{r,m,p}^-(\mathbb{C})$ forms a Riemannian submanifold and the $\mathcal{H}_2$-optimal MOR problem (2.1) can be written as a Riemannian optimization problem

$$\underset{\widehat{H}\in\Sigma_{r,m,p}^-(\mathbb{C})}{\text{minimize}} f(\widehat{H}) = \left\| H - \widehat{H} \right\|_{\mathcal{H}_2}^2. \tag{5.1}$$

To find $\operatorname{grad} f(\widehat{H})$ using (3.1), we first need to find a smooth extension of f to an open neighborhood of $\Sigma_{r,m,p}^-(\mathbb{C})$ in $\mathcal{H}_2$. To do so, we define $\overline{f}\colon \mathcal{H}_2 \rightarrow \mathbb{R}$ by $\overline{f}(G) = \|H - G\|_{\mathcal{H}_2}^2$, which satisfies $\overline{f}|_{\Sigma_{r,m,p}^-(\mathbb{C})} = f$. To find the Euclidean gradient of the extension $\overline{f}$, note that

$$\overline{f}(G + \Delta G) = \overline{f}(G) + 2\operatorname{Re}(\langle G - H, \Delta G\rangle_{\mathcal{H}_2}) + \|\Delta G\|_{\mathcal{H}_2}^2.$$

If we restrict to real rational functions, then we directly find that $\nabla\overline{f}|_{\mathfrak{R}\mathcal{H}_2}(G) = 2(G - H)$ for all $G \in \mathfrak{R}\mathcal{H}_2$. Since $\overline{f}\colon \mathcal{H}_2 \rightarrow \mathbb{R}$ is not complex differentiable (the only real-valued analytic functions are constant functions), we need a different concept of a complex gradient. Following [NAP08] (based on Wirtinger calculus), we use $2\nabla_{\overline{G}}\overline{f} = \nabla_{G^{\mathfrak{R}}}\overline{f} + \iota\nabla_{G^{\mathfrak{I}}}\overline{f}$, where $G = G^{\mathfrak{R}} + \iota G^{\mathfrak{I}}$ and $G^{\mathfrak{R}}, G^{\mathfrak{I}}$ are real. We find that

$$\operatorname{Re}(\langle G - H, \Delta G\rangle_{\mathcal{H}_2}) = \left\langle G^{\mathfrak{R}} - H^{\mathfrak{R}}, \Delta G^{\mathfrak{R}} \right\rangle_{\mathcal{H}_2}$$
$$+ \left\langle G^{\mathfrak{I}} - H^{\mathfrak{I}}, \Delta G^{\mathfrak{I}} \right\rangle_{\mathcal{H}_2}.$$

Thus, $\nabla_{G^{\mathfrak{R}}}\overline{f}(G^{\mathfrak{R}}) = 2(G^{\mathfrak{R}} - H^{\mathfrak{R}})$ and $\nabla_{G^{\mathfrak{I}}}\overline{f}(G^{\mathfrak{I}}) = 2(G^{\mathfrak{I}} - H^{\mathfrak{I}})$, and

$$\nabla_{\overline{G}}\overline{f}(G) = G - H. \tag{5.2}$$

Then, from (3.1), it follows that the Riemannian gradient of the $\mathcal{H}_2$-minimization problem is given by

$$\operatorname{grad} f(\widehat{H}) = \operatorname{Proj}_{\widehat{H}}\left(\nabla_{\overline{\widehat{H}}}\overline{f}(\widehat{H})\right) = \operatorname{Proj}_{\widehat{H}}\left(\widehat{H} - H\right), \tag{5.3}$$
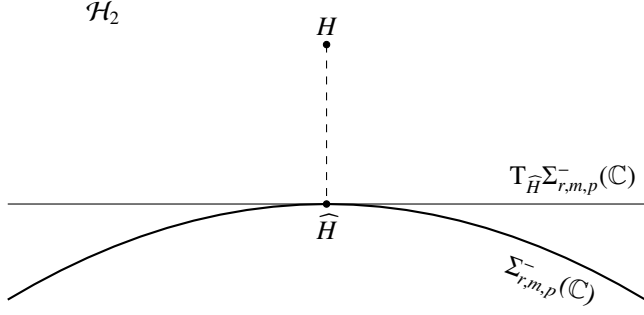
Figure 1: Necessary $\mathcal{H}_2$-optimality conditions in terms of orthogonality. H is the full-order transfer function, $\Sigma_{r,m,p}^-(\mathbb{C})$ is the manifold of stable rational functions of McMillan degree r, $\widehat{H}$ is the $\mathcal{H}_2$-optimal reduced-order transfer function, $\mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C})$ is the tangent space of $\Sigma_{r,m,p}^-(\mathbb{C})$ at $\widehat{H}$.

where $\mathrm{Proj}_{\widehat{H}}\colon \mathcal{H}_2 \to \mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C})$ is the orthogonal projector onto $\mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C})$.

6 GEOMETRIC INTERPRETATION OF NECESSARY OPTIMALITY CONDITIONS AND IRKA

Meier and Luenberger [ML66] derived the interpolatory $\mathcal{H}_2$-optimality conditions (2.2) and their geometric interpretation for the SISO case, i.e., $m = p = 1$. Using our analysis from Section 5, we generalize these results to multiple-input multiple-output (MIMO) systems using Riemannian optimization theory. In particular, in Section 6.1, we show that any (locally) $\mathcal{H}_2$-optimal ROM $\widehat{H}$ is such that the error system $H - \widehat{H}$ is orthogonal to the manifold $\Sigma_{r,m,p}^-(\mathbb{C})$. Next, we show that the orthogonality conditions are equivalent to the known interpolatory necessary optimality conditions (2.2). Then, in Section 6.2, these results lead to a geometric interpretation of IRKA. In turn, this leads to the interpretation of IRKA as a Riemannian gradient descent method.

6.1 Geometry of Interpolatory Conditions

Having interpreted the $\mathcal{H}_2$-minimization as a Riemannian optimization problem, it follows that the necessary optimality condition (3.3), in view of (5.2), becomes

$$H - \widehat{H} \perp \mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C}). \tag{6.1}$$

This orthogonality result is illustrated in Figure 1. Therefore, $\mathcal{H}_2$-optimal ROMs are necessarily such that the error system is orthogonal to the manifold $\Sigma_{r,m,p}^-(\mathbb{C})$.

To derive the interpolatory necessary conditions from (6.1), we first need to determine the tangent vectors of $\Sigma_{r,m,p}^-(\mathbb{C})$ at $\widehat{H}$. Our next result achieves this result for MIMO systems with simple poles.

Lemma 6.1. *Let $\widehat{H} \in \Sigma_{r,m,p}^-(\mathbb{C})$, with a pole-residue form $\widehat{H}(s) = \sum_{i=1}^r \frac{c_i b_i^*}{s - \lambda_i}$, have pairwise distinct poles. Then the tangent space $\mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C})$ is spanned by*

$$\frac{e_j b_i^*}{s - \lambda_i}, \quad \frac{c_i e_\ell^*}{s - \lambda_i}, \quad \frac{c_i b_i^*}{(s - \lambda_i)^2}, \tag{6.2}$$

for $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, p$, $\ell = 1, 2, \ldots, m$.

Proof. Let $\gamma\colon I \to \Sigma_{r,m,p}^-(\mathbb{C})$ be a smooth curve where $0 \in I \subset \mathbb{R}$ and $\gamma(0) = \widehat{H}$. Since rational functions with simple poles form a relatively open subset of $\Sigma_{r,m,p}^-(\mathbb{C})$, there is a neighborhood of 0 such that $\gamma(t)$ has simple poles for all t in that neighborhood. Let $\gamma(t)(s) = \sum_{i=1}^r \frac{c_i(t)b_i(t)^*}{s - \lambda_i(t)}$ with $\lambda_i(0) = \lambda_i$, $b_i(0) = b_i$, and $c_i(0) = c_i$. Then

$$\gamma'(t)(s) = \sum_{i=1}^r \left(\frac{c_i'(t)b_i(t)^*}{s - \lambda_i(t)} + \frac{c_i(t)b_i'(t)^*}{s - \lambda_i(t)} + \frac{c_i(t)b_i(t)^*\lambda_i'(t)}{(s - \lambda_i(t))^2} \right).$$

Therefore,

$$\gamma'(0)(s) = \sum_{i=1}^r \left(\frac{c_i'(0)b_i^*}{s - \lambda_i} + \frac{c_ib_i'(0)^*}{s - \lambda_i} + \frac{c_ib_i^*\lambda_i'(0)}{(s - \lambda_i)^2} \right),$$

which is a linear combination of (6.2), proving that all tangent vectors are in the span of (6.2). Conversely, since $\lambda_i'(0) \in \mathbb{C}$, $b_i'(0) \in \mathbb{C}^m$, $c_i'(0) \in \mathbb{C}^p$ are arbitrary, it follows that all elements of the span of (6.2) are tangent vectors. $\square$

Now with the full characterization of the tangent vectors, we can give an interpolatory interpretation of orthogonality.

Theorem 6.2. *Let $\widehat{H} \in \Sigma_{r,m,p}^-(\mathbb{C})$, with a pole-residue form $\widehat{H}(s) = \sum_{i=1}^r \frac{c_ib_i^*}{s - \lambda_i}$, have pairwise distinct poles. Furthermore, let $F \in \mathcal{H}_2$ be arbitrary. Then*

$$F \perp \mathrm{T}_{\widehat{H}}\Sigma_{r,m,p}^-(\mathbb{C}) \tag{6.3}$$

if and only if

$$F\left(-\overline{\lambda_i}\right)b_i = 0, \quad c_i^* F\left(-\overline{\lambda_i}\right) = 0, \quad c_i^* F'\left(-\overline{\lambda_i}\right)b_i = 0 \tag{6.4}$$

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

Proof. Using Lemma 6.1, it follows that (6.3) is equivalent to

$$\left\langle \frac{e_j b_i^*}{s - \lambda_i}, F \right\rangle_{\mathcal{H}_2} = 0, \quad j = 1, 2, \ldots, p,$$

$$\left\langle \frac{c_i e_\ell^*}{s - \lambda_i}, F \right\rangle_{\mathcal{H}_2} = 0, \quad \ell = 1, 2, \ldots, m, \tag{6.5}$$

$$\left\langle \frac{c_i b_i^*}{(s - \lambda_i)^2}, F \right\rangle_{\mathcal{H}_2} = 0,$$

for $i = 1, 2, \ldots, r$. Using Cauchy's integral formula gives that, for any $G \in \mathcal{H}_2$, $\lambda \in \mathbb{C}^-$, $b \in \mathbb{C}^m$, and $c \in \mathbb{C}^p$ (see, e.g., [ABG20, Lemma 2.1.4]),

$$\left\langle \frac{cb^*}{s - \lambda}, G \right\rangle_{\mathcal{H}_2} = c^* G\left(-\overline{\lambda}\right)b \quad \text{and}$$

$$\left\langle \frac{cb^*}{(s - \lambda)^2}, G \right\rangle_{\mathcal{H}_2} = -c^* G'\left(-\overline{\lambda}\right)b.$$

Therefore, it follows that (6.5) is equivalent to (6.4). $\square$

Applying Theorem 6.2 to $F = H - \widehat{H}$ directly yields the desired result.

Corollary 6.3. *Let* $\widehat{H} \in \Sigma_{r,m,p}^{-}(\mathbb{C})$*, with a pole-residue form* $\widehat{H}(s) = \sum_{i=1}^{r} \frac{c_i b_i^*}{s - \lambda_i}$*, have pairwise distinct poles. Then the geometric necessary optimality condition* (6.1) *is equivalent to the interpolatory necessary conditions* (2.2).

Note that the above results extend to real systems. In particular, in Lemma 6.1, the spanning set (6.2) of $\mathrm{T}_{\widehat{H}} \Sigma_{r,m,p}^{-}(\mathbb{R})$ becomes of twice the size by splitting the rational functions in (6.2) into real and imaginary parts. Then Theorem 6.2 remains the same after replacing $\Sigma_{r,m,p}^{-}(\mathbb{C})$ and $\mathcal{H}_2$ by $\Sigma_{r,m,p}^{-}(\mathbb{R})$ and $\mathfrak{R}\mathcal{H}_2$, respectively. Therefore, also Corollary 6.3 stays the same.

6.2 Geometric Interpretation of IRKA

IRKA has been initially introduced and studied as a fixed point iteration [GAB08]. In this section, our goal is to analyze the geometry of the IRKA iteration for the MIMO case using the tools from Section 6.1.

First, Theorem 6.2 allows us to represent each IRKA step as an orthogonal projection as we prove next.

Corollary 6.4. *Let* $H \in \mathcal{H}_2$ *be given. Furthermore, let* $\widehat{H}_k, \widehat{H}_{k+1} \in \Sigma_{r,m,p}^{-}(\mathbb{C})$ *be two consecutive iterates of IRKA such that* $\widehat{H}_k$ *has pairwise distinct poles. Then*

$$H - \widehat{H}_{k+1} \perp \mathrm{T}_{\widehat{H}_k} \Sigma_{r,m,p}^{-}(\mathbb{C}), \qquad (6.6)$$

i.e., $\widehat{H}_{k+1}$ *is an orthogonal projection of H onto $\Sigma_{r,m,p}^{-}(\mathbb{C})$ along the normal space of $\Sigma_{r,m,p}^{-}(\mathbb{C})$ at $\widehat{H}_k$.*

Proof. Applying Theorem 6.2 with $F = H - \widehat{H}_{k+1}$ shows that (6.6) is equivalent to (2.6). $\square$

Thus, Corollary 6.4 gives an interpretation of IRKA as an iterative orthogonal projection method. This is illustrated in Figure 2a. In the next section, this will help us in interpreting IRKA as a Riemannian optimization method.

7 IRKA IS A RIEMANNIAN OPTIMIZATION METHOD

So far, we have interpreted $\mathcal{H}_2$-optimal MOR as a Riemannian optimization problem and gave a geometric interpretation of IRKA based on the embedded manifold structure of $\Sigma_{r,m,p}^{-}(\mathbb{C})$. With this, it becomes clear that IRKA can be interpreted as a Riemannian gradient descent method applied to (5.1) for a particular choice of step size and retraction.

Indeed, Figure 2b suggests that IRKA is a Riemannian gradient descent method (3.4) with $\alpha_k = 1$ and the orthographic retraction, i.e., a projection to the manifold that is orthogonal to the tangent space. In Section 7.1, we extend a previous result on orthographic retractions from finite-dimensional Euclidean spaces to infinite-dimensional spaces. Then using this result in Section 7.2, we show that IRKA does behave as a Riemannian gradient descent method when well-defined.

7.1 Orthographic Retraction

The orthographic retraction R is such that

$$\widehat{H} + G - \mathrm{R}_{\widehat{H}}(G) \perp \mathrm{T}_{\widehat{H}} \Sigma_{r,m,p}^{-}(\mathbb{C}),$$

for all $\widehat{H} \in \Sigma_{r,m,p}^{-}(\mathbb{C})$ and all G in some open subset of $\mathrm{T}_{\widehat{H}} \Sigma_{r,m,p}^{-}(\mathbb{C})$ containing zero. There is work on projection-like retractions, including orthographic retractions, in [AM12, Lemma 20] for embedded submanifolds of finite-dimensional Euclidean spaces. The following lemma generalizes it to infinite-dimensional spaces.

Lemma 7.1. *Let* $\mathcal{M}$ *be an embedded submanifold of a Hilbert space* $\mathbb{E}$*. For all* $\widehat{p} \in \mathcal{M}$*, there exists a neighborhood* $\mathcal{U}_{\mathrm{T}\mathcal{M}}$ *of* $(\widehat{p}, 0)$ *in* $\mathrm{T}\mathcal{M}$ *such that for all* $(p, v) \in \mathcal{U}_{\mathrm{T}\mathcal{M}}$*, there is one and only one smallest* $w \in \mathrm{N}_p \mathcal{M}$ *such that* $p + v + w \in \mathcal{M}$*. Call it* $w(p, v)$ *and define* $\mathrm{R}(p, v) = p + v + w(p, v)$*. We have* $\mathrm{D}_v w(p, 0) = 0$ *and thus* R *defines a retraction around* $\widehat{p}$*. Since the expression of* R *does not depend on* $\widehat{p}$ *or* $\mathcal{U}_{\mathrm{T}\mathcal{M}}$*,* R *defines a retraction on* $\mathcal{M}$*.*

Proof. The proof of [AM12, Lemma 20] works exactly the same in this setting, except that $\mathbb{R}^d$ and $\mathbb{R}^{n-d}$ need to be replaced by $\mathbb{E}_1$ and $\mathbb{E}_2$, respectively, where $\mathbb{E} \simeq \mathbb{E}_1 \times \mathbb{E}_2$ such that for any $\widehat{p} \in \mathcal{M}$ there exists a chart $\varphi \colon U \to \mathbb{E}_1 \times \mathbb{E}_2$ of $\mathbb{E}$ around $\widehat{p}$ such that $\varphi(p) \in \mathbb{E}_1 \times \{0\}$ if and only if $p \in \mathcal{M} \cap U$ (see [Kli11, Theorem 1.3.5]). $\square$

7.2 IRKA as Riemannian Gradient Descent

We can now show that IRKA is a Riemannian gradient descent method, assuming IRKA iterates are well-defined.

Theorem 7.2. *Let* $\widehat{H}_k, \widehat{H}_{k+1} \in \Sigma_{r,m,p}^{-}(\mathbb{C})$ *be two consecutive iterates of IRKA,* f *as in* (5.1)*, and* $\mathrm{R} \colon \mathrm{T}\Sigma_{r,m,p}^{-}(\mathbb{C}) \to \Sigma_{r,m,p}^{-}(\mathbb{C})$ *the orthographic retraction. Furthermore, let* $\widehat{H}_{k+1}$ *be the unique element of* $\Sigma_{r,m,p}^{-}(\mathbb{C})$ *satisfying* (6.6) *and let* $\mathrm{R}_{\widehat{H}_k}(-\operatorname{grad} f(\widehat{H}_k))$ *exist. Then*

$$\widehat{H}_{k+1} = \mathrm{R}_{\widehat{H}_k}\big(-\operatorname{grad} f(\widehat{H}_k)\big). \qquad (7.1)$$

Proof. Since R is the orthographic retraction, we have

$$\widehat{H}_k - \operatorname{grad} f(\widehat{H}_k) - \mathrm{R}_{\widehat{H}_k}\big(-\operatorname{grad} f(\widehat{H}_k)\big) \perp \mathrm{T}_{\widehat{H}_k} \Sigma_{r,m,p}^{-}(\mathbb{C}).$$

Next, the relation (5.3) can be written as

$$\operatorname{grad} f(\widehat{H}_k) + H - \widehat{H}_k \perp \mathrm{T}_{\widehat{H}_k} \Sigma_{r,m,p}^{-}(\mathbb{C}).$$

Adding the above two relations, we find

$$H - \mathrm{R}_{\widehat{H}_k}\big(-\operatorname{grad} f(\widehat{H}_k)\big) \perp \mathrm{T}_{\widehat{H}_k} \Sigma_{r,m,p}^{-}(\mathbb{C}).$$

Since $\widehat{H}_{k+1}$ satisfying (6.6) was assumed to be unique, it follows that $\widehat{H}_{k+1}$ and $\mathrm{R}_{\widehat{H}_k}(-\operatorname{grad} f(\widehat{H}_k))$ are equal. $\square$

This shows that IRKA is a Riemannian gradient descent method with a fixed step size. The qualifier "well-defined" is needed as some iterates of IRKA may not lie on the manifold $\Sigma_{r,m,p}^{-}(\mathbb{C})$ (they may be unstable or of lower order) or there may not be a unique interpolant (2.6).

8 ALGORITHMIC DEVELOPMENTS

Now that we have established the understanding of IRKA as a Riemannian gradient descent method with a fixed step size, we are in a position to improve its performance and develop
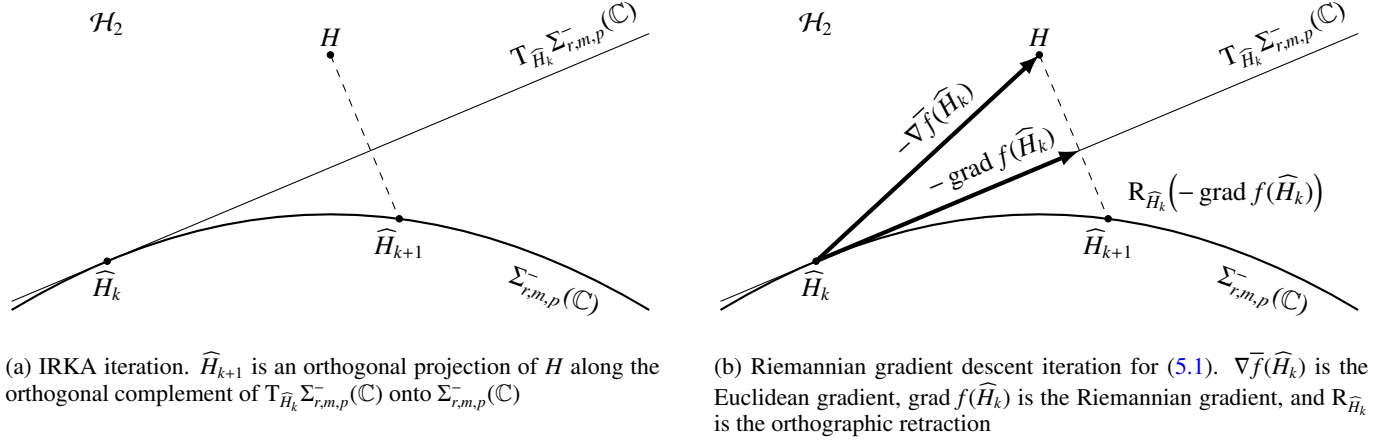
(a) IRKA iteration. $\widehat{H}_{k+1}$ is an orthogonal projection of H along the orthogonal complement of $T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C})$ onto $\Sigma^-_{r,m,p}(\mathbb{C})$

(b) Riemannian gradient descent iteration for (5.1). $\nabla \overline{f}(\widehat{H}_k)$ is the Euclidean gradient, grad $f(\widehat{H}_k)$ is the Riemannian gradient, and $R_{\widehat{H}_k}$ is the orthographic retraction

Figure 2: Two methods for $\mathcal{H}_2$-optimal model order reduction. H is the full-order transfer function, $\Sigma^-_{r,m,p}(\mathbb{C})$ is the manifold of stable rational functions of degree r, $\widehat{H}_k$ is the current reduced-order transfer function, $T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C})$ is the tangent space of $\Sigma^-_{r,m,p}(\mathbb{C})$ at $\widehat{H}_k$, and $\widehat{H}_{k+1}$ is the next iterate

new algorithms by considering various Riemannian optimization techniques.

The first such approach would be to implement a variation of IRKA using Riemannian gradient descent with variable step size. In particular, this is relevant because of two issues:

1. the orthographic retraction may only be defined locally,
2. $\Sigma^-_{r,m,p}(\mathbb{R})$ is disconnected for $m = p = 1$.

Regarding the first point, it is known that IRKA may generate intermediate iterates with unstable poles. For the second point, we find that IRKA can jump between different connected components, but the theory of Riemannian optimization assumes that the manifold is connected. We resolve both of these issues using backtracking line search, as discussed in this section.

We discuss connectedness and how to detect connected components in Section 8.1. We emphasize that connectedness is a potential concern only in the SISO case. In Section 8.2, we employ backtracking line search to enforce stability, a decrease in $\mathcal{H}_2$ error, and that the iterates remain on the same connected component.

8.1 Connectedness

As mentioned in Section 4.1, $\Sigma^-_{r,m,p}(\mathbb{R})$ has $r+1$ connected components when $m = p = 1$. Furthermore, the connected components can be identified by the Cauchy index.

For a real proper rational function with partial fraction expansion

$$F(s) = \sum_{i=1}^{k} \sum_{j=1}^{\mu_i} \frac{\varphi_i^{(j)}}{(s - \lambda_i)^j} + G(s),$$

where $\lambda_1, \ldots, \lambda_k$ are all real distinct poles of F (G has no real poles), the Cauchy index of F, colloquially speaking, is the number of real poles at which the function jumps from $-\infty$ to ∞ minus the number of real poles with jumps from ∞ to $-\infty$.

Formally, the Cauchy index of F is defined as

$$CI(F) = \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j \text{ odd}}}^{\mu_i} \text{sgn}\left(\varphi_i^{(j)}\right).$$

Assuming moderate r, one can compute the pole-residue form of $\widehat{H}_k$ in a numerically efficient way and compute $CI(\widehat{H}_k)$ at every iteration in IRKA. Doing precisely that, we show via various numerical experiments in Section 9 that, for SISO systems, during the IRKA iterations, the Cauchy index can change. Using backtracking line search as we explain in the next sections, we ensure that the Cauchy index remains the same throughout the iteration for SISO systems. We re-emphasize that connectedness is not a concern when $m > 1$ and/or $p > 1$.

8.2 Line Search

Here we derive a Riemannian gradient descent method with variable step size and propose a backtracking line search for $\mathcal{H}_2$-optimal MOR. This will resolve both potential issues listed at the beginning of this section.

Recall that Theorem 7.2 states that IRKA is a Riemannian gradient descent method with a fixed step size $\alpha_k = 1$, i.e., $\widehat{H}_{k+1} = R_{\widehat{H}_k}(-\text{grad } f(\widehat{H}_k))$ as shown in (7.1) where R denotes the orthographic retraction. Here, exploiting this new Riemannian optimization perspective of IRKA, we develop a new formulation where the classical IRKA step is replaced with $\widehat{H}_{k+1} = R_{\widehat{H}_k}(-\alpha_k \text{ grad } f(\widehat{H}_k))$ for a positive step size α_k. With an appropriate choice of α_k with backtracking, this modification guarantees stability of $\widehat{H}_k$ at every step and also makes sure that the new modified Riemannian-based IRKA iterate do not jump between different connected components.

Two issues remain. Recall that in the classical IRKA case ($\widehat{H}_{k+1} = R_{\widehat{H}_k}(-\text{grad } f(\widehat{H}_k))$), every iterate is obtained via bi-tangential Hermite interpolation. How can we compute the iterates with a variable step size? First, we answer this question and show that the new updates can still be achieved using interpolation. Also recall that at every IRKA step, one performs a

Petrov-Galerkin projection as in (2.4). However, a line search with backtracking will require re-computing $\widehat{H}_{k+1}$ for all the α_k's tested. A naive implementation of this will then require repeated Petrov-Galerkin projections during line search, which is computationally very expensive for the large-scale problems of interest to MOR. Thus, as a second contribution in this section, we develop an efficient implementation for constructing $\widehat{H}_{k+1}$ that avoids repeated Petrov-Galerkin projection during line search.

The next result shows that $\widehat{H}_{k+1}$ obtained via a step size $\alpha_k \neq 1$ is still an interpolant, more precisely it is a bitangential Hermite interpolant of $H_{k,\alpha_k} = (1-\alpha_k)\widehat{H}_k + \alpha_k H$ (Figure 2b helps to show why this is the case).

Theorem 8.1. *Let $H \in \mathcal{H}_2$, $\widehat{H}_k \in \Sigma^-_{r,m,p}(\mathbb{C})$, $\alpha_k > 0$. Let*

$$\widehat{H}_{k+1} = R_{\widehat{H}_k}\left(-\alpha_k \operatorname{grad} f(\widehat{H}_k)\right)$$

be obtained via Riemannian gradient descent with step size α_k. Define

$$H_{k,\alpha_k} = (1 - \alpha_k)\widehat{H}_k + \alpha_k H.$$

Then

$$H_{k,\alpha_k} - \widehat{H}_{k+1} \perp T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C}),$$

thus $\widehat{H}_{k+1}$ is a bitangential Hermite interpolant to H_{k,α_k}.

Proof. From the property of $\operatorname{Proj}_{\widehat{H}_k}$, we have that

$$H_{k,\alpha_k} - \widehat{H}_k - \operatorname{Proj}_{\widehat{H}_k}\left(H_{k,\alpha_k} - \widehat{H}_k\right) \perp T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C}).$$

Using that R is the orthographic retraction, we have

$$\widehat{H}_k - \alpha_k \operatorname{grad} f(\widehat{H}_k) - \widehat{H}_{k+1} \perp T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C}).$$

Next, we observe that

$$\operatorname{Proj}_{\widehat{H}_k}\left(H_{k,\alpha_k} - \widehat{H}_k\right) = \operatorname{Proj}_{\widehat{H}_k}\left(\alpha_k\left(H - \widehat{H}_k\right)\right)$$
$$= \alpha_k \operatorname{Proj}_{\widehat{H}_k}\left(H - \widehat{H}_k\right) = -\alpha_k \operatorname{grad} f(\widehat{H}_k).$$

Therefore,

$$H_{k,\alpha_k} - \widehat{H}_{k+1} = \left(H_{k,\alpha_k} - \widehat{H}_k - \operatorname{Proj}_{\widehat{H}_k}\left(H_{k,\alpha_k} - \widehat{H}_k\right)\right)$$
$$+ \left(\widehat{H}_k - \alpha_k \operatorname{grad} f(\widehat{H}_k) - \widehat{H}_{k+1}\right)$$

is orthogonal to $T_{\widehat{H}_k}\Sigma^-_{r,m,p}(\mathbb{C})$. $\quad\square$

This result shows that $\widehat{H}_{k+1}$ is a bitangential Hermite interpolant to H_{k,α_k} and thus can be constructed, e.g., using Petrov-Galerkin projection as discussed in Section 2.4 but now applied to the affine combination $H_{k,\alpha_k} = (1 - \alpha_k)\widehat{H}_k + \alpha_k H$ (as opposed to H as done in classical IRKA). A naive implementation of line search would then, for every new α_k, build the state-space form of H_{k,α_k}, compute the projection matrices by solving $2r$ linear systems in (2.3), and then project to find the new ROM, which would be costly. The following theorem show how to avoid these repeated projections during line search.

Theorem 8.2. *Let $H(s) = C(sE - A)^{-1}B$ and $\widehat{H}_k(s) = \widehat{C}_k(s\widehat{E}_k - \widehat{A}_k)^{-1}\widehat{B}_k$ be real systems. Then a state-space realization of $\widehat{H}_{k+1} = R_{\widehat{H}_k}\left(-\alpha_k \operatorname{grad} f(\widehat{H}_k)\right)$ is*

$$\widehat{E}_{k+1} = \widehat{E}_k - \alpha_k\left(\widehat{E}_k - \widetilde{Q}_k^{-1}\widetilde{Q}_k^{\mathrm{T}}E\widetilde{P}_k\widehat{P}_k^{-1}\right), \tag{8.1a}$$

$$\widehat{A}_{k+1} = \widehat{A}_k - \alpha_k\left(\widehat{A}_k - \widetilde{Q}_k^{-1}\widetilde{Q}_k^{\mathrm{T}}A\widetilde{P}_k\widehat{P}_k^{-1}\right), \tag{8.1b}$$

$$\widehat{B}_{k+1} = \widehat{B}_k - \alpha_k\left(\widehat{B}_k - \widetilde{Q}_k^{-1}\widetilde{Q}_k^{\mathrm{T}}B\right), \tag{8.1c}$$

$$\widehat{C}_{k+1} = \widehat{C}_k - \alpha_k\left(\widehat{C}_k - C\widetilde{P}_k\widehat{P}_k^{-1}\right), \tag{8.1d}$$

for any $\alpha_k > 0$ where $\widetilde{P}_k$, $\widehat{P}_k$, $\widetilde{Q}_k$, $\widehat{Q}_k$ are solutions to Sylvester and Lyapunov equations

$$A\widetilde{P}_k\widehat{E}_k^{\mathrm{T}} + E\widetilde{P}_k\widehat{A}_k^{\mathrm{T}} + B\widehat{B}_k^{\mathrm{T}} = 0, \tag{8.2a}$$

$$\widehat{A}_k\widehat{P}_k\widehat{E}_k^{\mathrm{T}} + \widehat{E}_k\widehat{P}_k\widehat{A}_k^{\mathrm{T}} + \widehat{B}_k\widehat{B}_k^{\mathrm{T}} = 0, \tag{8.2b}$$

$$A^{\mathrm{T}}\widetilde{Q}_k\widehat{E}_k + E^{\mathrm{T}}\widetilde{Q}_k\widehat{A}_k + C^{\mathrm{T}}\widehat{C}_k = 0, \tag{8.2c}$$

$$\widehat{A}_k^{\mathrm{T}}\widehat{Q}_k\widehat{E}_k + \widehat{E}_k^{\mathrm{T}}\widehat{Q}_k\widehat{A}_k + \widehat{C}_k^{\mathrm{T}}\widehat{C}_k = 0. \tag{8.2d}$$

Proof. A state-space realization of H_{k,α_k} for $\alpha_k \neq 1$ is

$$E_{k,\alpha_k} = \begin{bmatrix} \alpha_k E & 0 \\ 0 & (1 - \alpha_k)\widehat{E}_k \end{bmatrix}, \tag{8.3a}$$

$$A_{k,\alpha_k} = \begin{bmatrix} \alpha_k A & 0 \\ 0 & (1 - \alpha_k)\widehat{A}_k \end{bmatrix}, \tag{8.3b}$$

$$B_{k,\alpha_k} = \begin{bmatrix} \alpha_k B \\ (1 - \alpha_k)\widehat{B}_k \end{bmatrix}, \tag{8.3c}$$

$$C_{k,\alpha_k} = \begin{bmatrix} \alpha_k C & (1 - \alpha_k)\widehat{C}_k \end{bmatrix}. \tag{8.3d}$$

Then $\widehat{H}_{k+1}$ can be realized by

$$\begin{aligned} \widehat{E}_{k+1} &= W^{\mathrm{T}}E_{k,\alpha_k}V, & \widehat{A}_{k+1} &= W^{\mathrm{T}}A_{k,\alpha_k}V, \\ \widehat{B}_{k+1} &= W^{\mathrm{T}}B_{k,\alpha_k}, & \widehat{C}_{k+1} &= C_{k,\alpha_k}V, \end{aligned} \tag{8.4}$$

where V and W are solutions to Sylvester equations (see Section 2.4)

$$A_{k,\alpha_k}V\widehat{E}_k^{\mathrm{T}} + E_{k,\alpha_k}V\widehat{A}_k^{\mathrm{T}} + B_{k,\alpha_k}\widehat{B}_k^{\mathrm{T}} = 0, \tag{8.5a}$$

$$A_{k,\alpha_k}^{\mathrm{T}}W\widehat{E}_k + E_{k,\alpha_k}^{\mathrm{T}}W\widehat{A}_k + C_{k,\alpha_k}^{\mathrm{T}}\widehat{C}_k = 0. \tag{8.5b}$$

Exploiting the block structure (8.3) of the coefficient matrices in (8.5) and comparing it to (8.2), we find that $V = \begin{bmatrix} \widetilde{P}_k \\ \widehat{P}_k \end{bmatrix}$ and $W = \begin{bmatrix} \widetilde{Q}_k \\ \widehat{Q}_k \end{bmatrix}$. Therefore, the expressions in (8.4) simplify to

$$\widehat{E}_{k+1} = \alpha_k\widetilde{Q}_k^{\mathrm{T}}E\widetilde{P}_k + (1 - \alpha_k)\widehat{Q}_k\widehat{E}_k\widehat{P}_k,$$

$$\widehat{A}_{k+1} = \alpha_k\widetilde{Q}_k^{\mathrm{T}}A\widetilde{P}_k + (1 - \alpha_k)\widehat{Q}_k\widehat{A}_k\widehat{P}_k,$$

$$\widehat{B}_{k+1} = \alpha_k\widetilde{Q}_k^{\mathrm{T}}B + (1 - \alpha_k)\widehat{Q}_k\widehat{B}_k,$$

$$\widehat{C}_{k+1} = \alpha_k C\widetilde{P}_k + (1 - \alpha_k)\widehat{C}_k\widehat{P}_k,$$

which also holds for $\alpha_k = 1$. Premultiplying by $\widehat{Q}_k^{-1}$ and postmultiplying by $\widehat{P}_k^{-1}$ gives the equivalent realization of $\widehat{H}_{k+1}$ in (8.1). $\quad\square$

We now have all the ingredients to design an effective numerical algorithm for a variation of IRKA using Riemannian gradient descent with backtracking. A pseudo-code is presented in

Algorithm 1. We initialize the method with a stable ROM $\widehat{H}_1$ and construct a new ROM with $\alpha_k = 1$, i.e., a step of IRKA. Then, we do backtracking and reduce the step size by two if $\widehat{H}_{k+1}$ is unstable or the $\mathcal{H}_2$ has increased, or the Cauchy index has changed (in the SISO case only). The algorithm terminates if the step size becomes smaller than a specified minimum step size $\alpha_{\min} > 0$ or if a desired tolerance is reached. Thus, Algorithm 1 guarantees stability throughout, the $\mathcal{H}_2$ error decreases at every step, and upon convergence (interpolatory) $\mathcal{H}_2$-optimality conditions are satisfied.

Algorithm 1 Riemannian Gradient Descent IRKA (RGD-IRKA)

Input: FOM (E, A, B, C), initial ROM $(\widehat{E}_1, \widehat{A}_1, \widehat{B}_1, \widehat{C}_1)$, maximum iteration number `maxit`, minimum step size $\alpha_{\min} > 0$, tolerance `tol` > 0.

Output: ROM $(\widehat{E}, \widehat{A}, \widehat{B}, \widehat{C})$.

1: Set $\widehat{H}_1$ as the transfer function of the initial ROM.
2: **for** k in $1, 2, \ldots,$ `maxit` **do**
3:      $\alpha_k = 1$
4:      **while** ROM in (8.1) is unstable, has a different Cauchy index, or it increases the $\mathcal{H}_2$ error **do**
5:          $\alpha_k = \alpha_k/2$
6:          **if** $\alpha_k < \alpha_{\min}$ **then**
7:              Exit the **for** loop.
8:      Set $(\widehat{E}_{k+1}, \widehat{A}_{k+1}, \widehat{B}_{k+1}, \widehat{C}_{k+1})$ as in (8.1) with transfer function $\widehat{H}_{k+1}$.
9:      **if** $\|\widehat{H}_k - \widehat{H}_{k+1}\|_{\mathcal{H}_2} \leqslant$ `tol` $\cdot \alpha_k \|\widehat{H}_{k+1}\|_{\mathcal{H}_2}$ **then**
10:         Exit the **for** loop.
11: Return the last computed ROM.

A couple of remarks are in order. In checking whether $\mathcal{H}_2$ error has decreased, we do *not* explicitly compute $\|H - \widehat{H}_k\|_{\mathcal{H}_2}$ as this will require solving a large-scale Lyapunov equation at every step. Instead note that

$$\left\| H - \widehat{H}_k \right\|_{\mathcal{H}_2}^2 = \|H\|_{\mathcal{H}_2}^2 + 2\,\mathrm{Re}\left(\left\langle H, \widehat{H}_k \right\rangle_{\mathcal{H}_2}\right) + \left\| \widehat{H}_k \right\|_{\mathcal{H}_2}^2.$$

The first term in this formula is a constant. Thus, all we need to check is the last two terms which can be easily and effectively computed. We chose a stopping criterion based on a relative change in the ROM. Other stopping criteria are also possible.

Remark 8.3. *We note that computations in* (8.1) *involve solving linear systems with reduced Gramians* $\widehat{P}_k$ *and* $\widehat{Q}_k$ *as system matrices. We observed that in some cases these reduced Gramians can be ill-conditioned, which then leads to choosing small step sizes, thus requiring small* $\alpha_{\min}$. *In this paper, our main goal is to establish the main theoretical framework. Developing better ways of handling such numerical issues, e.g. by regularizing the Gramians via truncated SVD, would be a topic of future research. We have also observed that in some cases* $\widehat{E}_k$ *can become ill-conditioned. As a way to circumvent this, we check the condition number of* $\widehat{E}_k$ *and if it becomes large (e.g., larger than* 10^4*, which we use in the numerical examples), we apply a state-space transform to make* $\widehat{E}_k$ *the identity matrix.*

8.3 Interpretation as a Fixed Point Iteration

Note that Theorem 8.1, together with Theorem 6.2, gives that $\widehat{H}_{k+1}$ is a bitangential Hermite interpolant of H_{k,α_k} at the reflected poles and in the directions of $\widehat{H}_k(s) = \sum_{i=1}^{r} \frac{c_i^{(k)}(b_i^{(k)})^*}{s - \lambda_i^{(k)}}$, i.e.,

$$H_{k,\alpha_k}\left(-\overline{\lambda_i^{(k)}}\right) b_i^{(k)} = \widehat{H}_{k+1}\left(-\overline{\lambda_i^{(k)}}\right) b_i^{(k)}, \tag{8.6a}$$

$$\left(c_i^{(k)}\right)^* H_{k,\alpha_k}\left(-\overline{\lambda_i^{(k)}}\right) = \left(c_i^{(k)}\right)^* \widehat{H}_{k+1}\left(-\overline{\lambda_i^{(k)}}\right), \tag{8.6b}$$

$$\left(c_i^{(k)}\right)^* H_{k,\alpha_k}''\left(-\overline{\lambda_i^{(k)}}\right) b_i^{(k)} = \left(c_i^{(k)}\right)^* \widehat{H}_{k+1}'\left(-\overline{\lambda_i^{(k)}}\right) b_i^{(k)}, \tag{8.6c}$$

for $i = 1, 2, \ldots, r$. This can be interpreted as a fixed point iteration applied to

$$\left(\alpha H + (1 - \alpha)\widehat{H}\right)\left(-\overline{\lambda_i}\right) b_i = \widehat{H}\left(-\overline{\lambda_i}\right) b_i,$$

$$c_i^*\left(\alpha H + (1 - \alpha)\widehat{H}\right)\left(-\overline{\lambda_i}\right) = c_i^* \widehat{H}\left(-\overline{\lambda_i}\right),$$

$$c_i^*\left(\alpha H + (1 - \alpha)\widehat{H}\right)'\left(-\overline{\lambda_i}\right) b_i = c_i^* \widehat{H}'\left(-\overline{\lambda_i}\right) b_i,$$

which is clearly equivalent to (2.2) for any $\alpha > 0$. Therefore, the proposed Riemannian gradient descent formulation with line search for $\mathcal{H}_2$ minimization can be considered as a fixed point iteration applied to modified interpolatory necessary optimality conditions; thus more clearly illustrating the distinction from the original IRKA formulation, which corresponds to $\alpha = 1$.

9 NUMERICAL EXAMPLES

The code producing the presented results is available at [Mli23], written in the Python programming language and based on pyMOR [MRS16]. In all examples, we set `maxit` = 100, $\alpha_{\min} = 10^{-20}$, and `tol` $= 10^{-4}$ in Algorithm 1.

9.1 Example 1

We start with the following example from [GAB08, Section 5.4]:

$$H(s) = \frac{-s^2 + (7/4)s + 5/4}{s^3 + 2s^2 + (17/16)s + 15/32}.$$

Notably, [GAB08] shows that for the reduced order $r = 1$, the globally $\mathcal{H}_2$-optimal ROM for H, with the pole at 0.27272, is repellent for IRKA, thus IRKA does not converge. We apply IRKA and the proposed method, i.e. RGD-IRKA, as implemented in Algorithm 1, to this model. Figure 3 shows the results when both method are initialized with $\widehat{E}_1 = 1$, $\widehat{A}_1 = -0.27$, $\widehat{B}_1 = 1$, $\widehat{C}_1 = 1$, which has a Cauchy index of 1 and it is close to the global minimum (as used in [GAB08]). Indeed, we observe that IRKA iterates move away from the optimum as expected since the minimizer is a repellent fixed point. And note that many intermediate IRKA iterates are unstable. Furthermore, the Cauchy index changes many times during the IRKA iterations. On the other hand, RGD-IRKA converges quickly (with stability guarantee at every step) and it often uses $\alpha_k < 1$ as the step size. Moreover, note that the Cauchy index stays constant during RGD-IRKA.
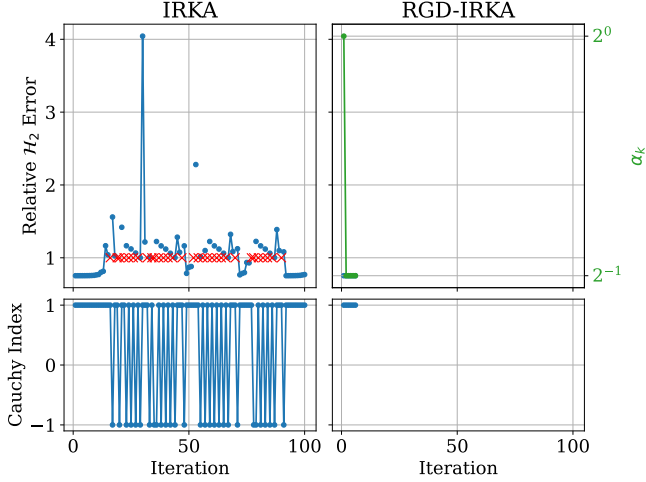
Figure 3: IRKA and RGD-IRKA results for the [GAB08] example and $r = 1$. Red crosses represent unstable models
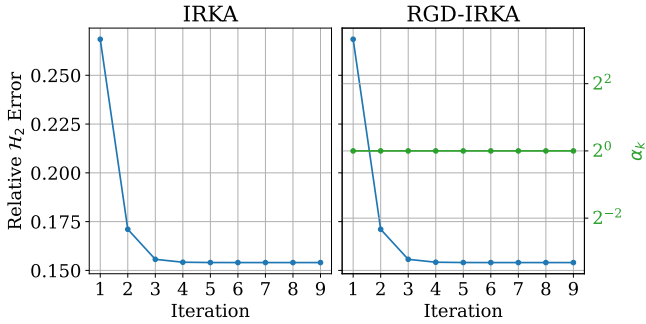

Figure 5: IRKA and RGD-IRKA results for the [GAB08] example, $r = 2$, and initial Cauchy index of 2

Figure 4: IRKA and RGD-IRKA results for the [GAB08] example, $r = 2$, and initial Cauchy index of 0

Figure 6: IRKA and RGD-IRKA results for the CD player example and $r = 6$. Red crosses represent unstable models

Next we compare IRKA and RGD-IRKA on the same example for $r = 2$ in Figure 4, initialized with

$$\widehat{E}_1 = I_2, \ \widehat{A}_1 = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, \ \widehat{B}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \widehat{C}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

Note that its transfer function has Cauchy index of 0. In this case with this specific initialization, IRKA and RGD-IRKA behave exactly the same. So, for these specific choices, the original IRKA steps indeed correspond to descent steps.

Next for $r = 2$, we change the initialization to

$$\widehat{E}_1 = I_2, \ \widehat{A}_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \ \widehat{B}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \widehat{C}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

which has Cauchy index of 2. The results, depicted in Figure 5, show that, in the SISO case, where the manifold $\Sigma_{r,m,p}^-(\mathbb{R})$ is disconnected and RGD-IRKA enforces the iterates to stay on the same connected component, a bad choice of the initial Cauchy index can lead to a bad ROM for RGD-IRKA, but IRKA still converges to a good local minimum as it jumps to a model with a different Cauchy index. Furthermore, we note that the pole-residue form of the ROM from RGD-IRKA is

$$\frac{0.97188}{s + 0.27344} + \frac{8.4933}{s + 6.9933 \times 10^{12}}.$$

It suggests that, since RGD-IRKA is restricted to one connected component due to line search, the iterates converge to
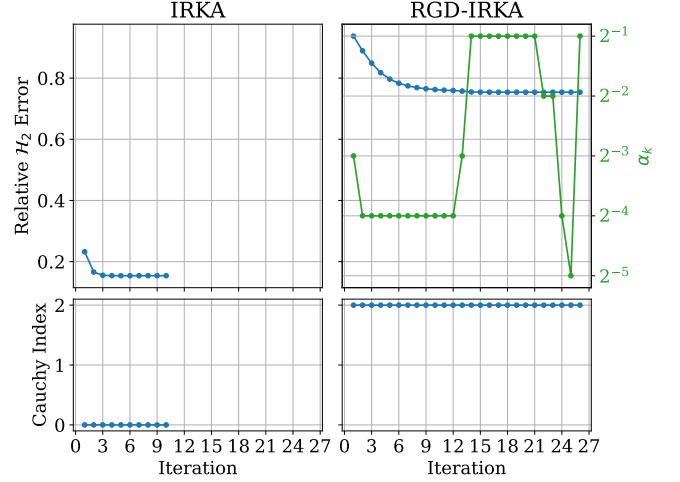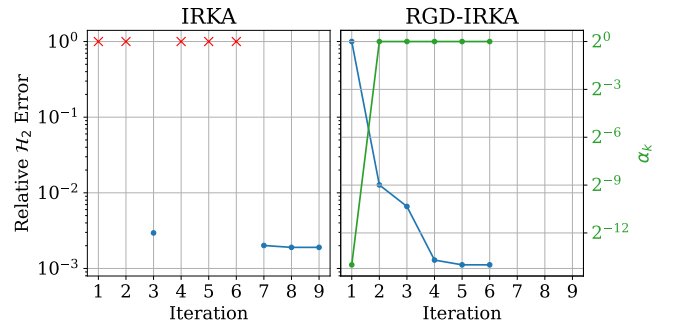
the boundary of the connected component, which are systems of lower order. We emphasize that Cauchy index initialization is only a potential concern in the SISO case.

9.2 Example 2

Here we show results on the CD player example from the NICONET benchmark collection [CVD02], which is a MIMO system with 2 inputs and 2 outputs. Figure 6 shows the result for $r = 6$, initialized with $\widehat{E}_1 = I_6$, $\widehat{A}_1 = \mathsf{diag}(-1, -2, \ldots, -6)$, $\widehat{B}_1 = \mathbb{1}_{6 \times 2}$, $\widehat{C}_1 = \mathbb{1}_{2 \times 6}$, where $\mathbb{1}_{m \times n} \in \mathbb{R}^{m \times n}$ is a matrix of ones. Note that, since the manifold $\Sigma_{r,m,p}^-(\mathbb{R})$ is now connected, Cauchy index is not necessary (and is not even defined). Here we observe convergence of IRKA with most iterates being unstable. On the other hand, RGD-IRKA converges faster and reaches a slightly better ROM (relative $\mathcal{H}_2$ errors are 1.9003×10^{-3} for IRKA and 1.1167×10^{-3} for RGD-IRKA). The first step size in RGD-IRKA is $2^{-14} \approx 6 \times 10^{-5}$ while the later ones are 1.

10 Conclusion

We showed that IRKA can be interpreted as a Riemannian gradient descent method with a fixed step size. This interpretation directly leads to a development of a Riemannian gradient de-

scent formulation of IRKA, called RGD-IRKA, that employs variable step size and line search, with guaranteed convergence and stability preservation. Numerical examples illustrated the benefits of RGD-IRKA.

REFERENCES

[ABG94] D. Alpay, L. Baratchart, and A. Gombani. On the differential structure of matrix-valued rational inner functions. In A. Feintuch and I. Gohberg, editors, *Nonselfadjoint Operators and Related Topics*, pages 30–66, Basel, 1994. Birkhäuser Basel. doi:10.1007/978-3-0348-8522-5_2.

[ABG10] A. C. Antoulas, C. A. Beattie, and S. Gugercin. Interpolatory model reduction of large-scale dynamical systems. In J. Mohammadpour and K. M. Grigoriadis, editors, *Efficient Modeling and Control of Large-Scale Systems*, pages 3–58. Springer US, Boston, MA, 2010. doi:10.1007/978-1-4419-5757-3_1.

[ABG20] A. C. Antoulas, C. A. Beattie, and S. Güğercin. *Interpolatory methods for model reduction.* Computational Science and Engineering 21. SIAM, Philadelphia, PA, 2020. doi:10.1137/1.9781611976083.

[AM12] P.-A. Absil and J. Malick. Projection-like retractions on matrix manifolds. *SIAM J. Optim.*, 22(1):135–158, 2012. doi:10.1137/100802529.

[AMS08] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, Princeton, N.J. Woodstock, 2008. URL: https://press.princeton.edu/absil.

[Ant05] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*, volume 6 of *Adv. Des. Control.* SIAM, Philadelphia, PA, 2005. doi:10.1137/1.9780898718713.

[BCS19] P. Benner, X. Cao, and W. Schilders. A bilinear $\mathcal{H}_2$ model order reduction approach to linear parameter-varying systems. *Advances in Computational Mathematics*, 45(5-6):2241–2271, April 2019. doi:10.1007/s10444-019-09695-9.

[BD82] C. I. Byrnes and T. E. Duncan. On certain topological invariants arising in system theory. In *New Directions in Applied Mathematics*, pages 29–71. Springer New York, 1982. doi:10.1007/978-1-4612-5651-9_3.

[BGTQ⁺21a] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors. *Model Order Reduction. Volume 1: System- and Data-Driven Methods and Algorithms.* De Gruyter, Berlin, 2021. doi:10.1515/9783110498967.

[BGTQ⁺21b] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors. *Model Order Reduction. Volume 2: Snapshot-Based Methods and Algorithms.* De Gruyter, Berlin, 2021. doi:10.1515/9783110671490.

[BKS11] P. Benner, M. Köhler, and J. Saak. Sparse-dense Sylvester equations in $\mathcal{H}_2$-model order reduction. Preprint MPIMD/11-11, Max Planck Institute Magdeburg, December 2011. URL: https://csc.mpi-magdeburg.mpg.de/preprints/2011/11/.

[BOCW17] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, editors. *Model Reduction and Approximation: Theory and Algorithms.* Computational Science & Engineering. SIAM, Philadelphia, PA, 2017. doi:10.1137/1.9781611974829.

[Bou23] N. Boumal. *An introduction to optimization on smooth manifolds.* Cambridge University Press, Cambridge, England, March 2023.

[Bro76] R. Brockett. Some geometric questions in the theory of linear systems. *IEEE Trans. Autom. Control*, 21(4):449–455, 1976. doi:10.1109/TAC.1976.1101301.

[Cla66] J. M. C. Clark. The consistent selection of local coordinates in linear system identification. In *Joint Automatic Control Conf.*, pages 576–580, 1966.

[CVD02] Y. Chahlaoui and P. Van Dooren. A collection of benchmark examples for model reduction of linear time invariant dynamical systems. Technical Report 2002-2, SLICOT Working Note, 2002. Available from www.slicot.org.

[Del85] D. F. Delchamps. Global structure of families of multivariable linear systems with an application to identification. *Mathematical Systems Theory*, 18(1):329–380, December 1985. doi:10.1007/bf01699476.

[GAB06] S. Gugercin, A. C. Antoulas, and C. A. Beattie. A rational Krylov iteration for optimal $\mathcal{H}_2$ model reduction. In *Proc. of the 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 1665–1667, 2006.

[GAB08] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30(2):609–638, 2008. doi:10.1137/060666123.

[Haz77] M. Hazewinkel. Moduli and canonical forms for linear dynamical systems II: The topological case. *Mathematical Systems Theory*, 10(1):363–385, December 1977. doi:10.1007/bf01683285.

[HK76] M. Hazewinkel and R. E. Kalman. On invariants, canonical forms and moduli for linear, constant, finite dimensional, dynamical systems. In G. Marchesini and S. K. Mitter, editors, *Mathematical Systems Theory*, pages 48–60, Berlin,

Heidelberg, 1976. Springer Berlin Heidelberg. doi:10.1007/978-3-642-48895-5_4.

[Kli11]   W. P. A. Klingenberg.   *Riemannian Geometry*.   De Gruyter, 2011.   doi:10.1515/9783110905120.

[Lan95]   S. Lang, editor. *Differential and Riemannian Manifolds*. Springer New York, 1995. doi:10.1007/978-1-4612-4182-9.

[Lee12]   J. M. Lee. *Introduction to Smooth Manifolds*. Springer New York, 2012. doi:10.1007/978-1-4419-9982-5.

[ML66]   L. Meier and D. G. Luenberger. Approximation of linear constant systems. In *Joint Automatic Control Conf.*, pages 585–588, 1966.

[ML67]   L. Meier and D. G. Luenberger. Approximation of linear constant systems. *IEEE Trans. Autom. Control*, 12(5):585–588, 1967. doi:10.1109/TAC.1967.1098680.

[Mli23]   P. Mlinarić. Code for Riemannian gradient descent IRKA numerical experiments, November 2023. Available from https://github.com/pmli/rgd-irka-experiments/tree/v1.

[MRS16]   R. Milk, S. Rave, and F. Schindler. pyMOR – generic algorithms and interfaces for model order reduction. *SIAM J. Sci. Comput.*, 38(5):S194–S216, 2016. doi:10.1137/15M1026614.

[NAP08]   Y. Nishimori, S. Akaho, and M. D. Plumbley. Natural conjugate gradient on complex flag manifolds for complex independent subspace analysis. In *Artificial Neural Networks - ICANN 2008*, pages 165–174. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-87536-9_18.

[SS15]   H. Sato and K. Sato. Riemannian trust-region methods for H^2 optimal model reduction. In *54th IEEE Conference on Decision and Control (CDC)*, pages 4648–4655, 2015. doi:10.1109/CDC.2015.7402944.

[VGA10]   P. Van Dooren, K. A. Gallivan, and P.-A. Absil. $\mathcal{H}_2$-optimal model reduction with higher-order poles. *SIAM J. Matrix Anal. Appl.*, 31(5):2738–2753, 2010. doi:10.1137/080731591.

[XJY15]   K. Xu, Y. Jiang, and Z. Yang. H_2 order-reduction for bilinear systems based on Grassmann manifold. *Journal of the Franklin Institute*, 352(10):4467–4479, 2015. doi:10.1016/j.jfranklin.2015.06.022.

[XZ13]   Y. Xu and T. Zeng. Fast optimal $\mathcal{H}_2$ model reduction algorithms based on Grassmann manifold optimization. *Int. J. Numer. Anal. Model.*, 10(4):972–991, 2013. URL: http://www.math.ualberta.ca/ijnam/Volume-10-2013/No-4-13/2013-04-13.pdf.

[YL99]   W.-Y. Yan and J. Lam. An approximate approach to $\mathcal{H}_2$ optimal model reduction. *IEEE Trans. Autom. Control*, 44(7):1341–1358, 1999. doi:10.1109/9.774107.

[ZDG96]   K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996. doi:10.1007/978-1-4471-6257-5.