On the Tightness of the Balanced Truncation Error Bound with an Application to Arrowhead Systems

Sean Joseph Reiter

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Masters of Science

in

Mathematics

Mark Embree, Chair

Serkan Gugercin

Christopher Beattie

Vasileios Kekatos

December 16th, 2021

Blacksburg, Virginia

Keywords: Model Reduction, Balanced Truncation, Error Bound, Arrowhead Systems, Power Systems

On the Tightness of the Balanced Truncation Error Bound with an Application to Arrowhead Systems

Sean Joseph Reiter

(ABSTRACT)

Balanced truncation model reduction for linear systems yields reduced-order models that satisfy a well-known error bound in terms of a system's Hankel singular values. This bound is known to hold with equality under certain conditions, such as when the full-order system is state-space symmetric. In this work, we derive more general conditions in which the balanced truncation error bound holds with equality. We show that this holds for single-input, single-output systems that exhibit a generalized type of state-space symmetry based on the sign parameters corresponding to a system's Hankel singular values. We prove an additional result that shows how to determine this state-space symmetry from the arrowhead realization of a system, if available. In particular, we provide a formula for the sign parameters of an arrowhead system in terms of the off-diagonal entries of its arrowhead realization. We then illustrate these results with an example of an arrowhead system arising naturally in power systems modeling that motivated our study.

On the Tightness of the Balanced Truncation Error Bound with an Application to Arrowhead Systems

Sean Joseph Reiter

(GENERAL AUDIENCE ABSTRACT)

Mathematical modeling of dynamical systems provides a powerful means for studying physical phenomena. Due the complexities of real-world problems, many mathematical models face computational difficulties due to the costs of accurate modeling. Model-order reduction of large-scale dynamical systems circumvents this by approximating the large-scale model with a "smaller" one that still accurately describes the problem of interest. Balanced truncation model reduction for linear systems is one such example, yielding reduced-order models that satisfy a tractable upper bound on the approximation error. This work investigates conditions in which this bound is known to hold with equality, becoming an exact formula for the error in reduction. We additionally show how to determine these conditions for a special class of linear dynamical systems known as arrowhead systems, which arise in special applications of network modeling. We provide an example of one such system from power systems modeling that motivated our study.

Dedication

To my mom, thank you for always encouraging me to do what I love.

Acknowledgments

To my advisors, Dr. Mark Embree and Dr. Serkan Gugercin. Thank you both for your guidance, mentorship, and collaboration (especially through the many ups and downs that came with solving these problems). I'm incredibly grateful to have advisors who are as invested in both my mathematical and personal growth as the two of you are. To the other graduate students in the mathematics department, thank you for the helpful conversations and distractions, throughout this past semester in particular. Finally, I'd like to thank the many friends and family members who have provided emotional support throughout this process.

Contents

List of Figures

List of Tables

List of Abbreviations

LTI Linear time-invariant

MOR Model-order reduction

ODE Ordinary differential equation

PDE Partial differential equation

SISO Single-Input Single-Output

Chapter 1

Introduction

The simulation of mathematical models provides a powerful tool for studying the behavior of complex physical phenomena. These models are *dynamical systems*, collections of ordinary differential equations (ODEs) describing the evolution of the state of the problem in time. Most generally, the systems are written as a set of equations

$$x'(t) = f(x(t), u(t)) \in \mathbb{R}^n, \quad t \geq 0, \tag{1.1}$$

$$y(t) = g(x(t), u(t)) \in \mathbb{R}^p, \tag{1.2}$$

where $x(t) \in \mathbb{R}^n$ is the internal state of the system at time t, $u(t) \in \mathbb{R}^m$ is some external forcing function influencing the system over time, $y(t) \in \mathbb{R}^p$ is some chosen observable of interest, and $f(x(t), u(t)) \in \mathbb{R}^n$, $g(x(t), u(t)) \in \mathbb{R}^p$ are vector-valued maps describing the time evolution of the system. The complexity of the system is represented by the number of internal state equations n, which we call the *order* of the system.

In many applications, these models are of an unfeasibly high complexity, rendering numerical simulation infeasible due to physical limitations on storage, computational power, or time. Primary examples include systems of ODEs obtained from spatial discretizations of a partial differential equation (PDE) governing some physical process of interest. The need for increasingly accurate numerical predictions necessitates finer detail in the discretization stage of the modeling process, leading to increasingly larger-scale systems with orders as high as $n = 10^6$ in some applications. *Model-order reduction* (MOR) is a process that seeks to temper these high computational costs while still meeting the demands for highly accurate numerical predictions. This efficiency is achieved by computing approximate low-order systems that mimic the input-output characteristics of the original while demanding significantly fewer computational resources. Once computed, the reduced-order model (ROM) can be used as a surrogate for the original in the application of interest. For more details on model reduction, see [1, 2, 3, 4, 5] and the references therein.

Balanced truncation model reduction [6, 7] is the gold standard for MOR of linear time-invariant (LTI) dynamical systems (those for which f and g in (1.1) are linear functions of the state and input at time t) as given in (1.3) below. The method enjoys many desirable properties, such as preservation of stability and a well-known upper bound on the reduction error in terms of a system's Hankel singular values. This bound is known to hold with equality under certain conditions, such as for single-input, single-output (SISO) systems that are *state space symmetric* [8], or when only reducing the model order by a single dimension. In such instances, the upper bound instead serves as an exact formula for the error in the reduction.

This thesis makes two primary contributions. First, we show that the balanced truncation error bound holds for a wider class of systems than the state-space symmetric case; namely, those exhibiting a generalized type of state-space symmetry based on the canonical balanced form of a dynamical system. We additionally prove this result for singular perturbation approximations of balanced systems. For the second contribution, we show how to derive this sign symmetry from the signs of the off-diagonal entries of a system's *arrowhead realization.* Linear systems having arrowhead realizations arise in applications of network modeling. In such cases, one can determine *a priori* from a system realization whether or not the balanced truncation error bound holds with equality. We prove this result by combining ideas from system theory and eigenvalue perturbation theory. We first encountered these phenomena while studying a model from power system dynamics. We observed numerically that the balanced truncation error bound for these models was tight, and then noted that their canonical balanced realization always exhibited a particular sign symmetry in which the truncated part of the model was state-space symmetric. We then noted that this sign symmetry could be derived from a general state-space realization in which the model takes an arrowhead form.

The rest of the thesis is organized as follows: In the remainder of Chapter 1, we formalize the notion of LTI systems, as well as the relevant concepts and motivations behind the model reduction problem. In Chapter 2, we briefly introduce and review the relevant system theoretic ideas underlying balanced truncation model reduction. In Chapter 3, we restate previous results regarding the equality of the balanced truncation error bound. We also introduce the canonical form of a balanced system, and describe the generalized state-space symmetry as previously mentioned, as well as its connection to the *sign parameters* of a linear system. We then show that the balanced truncation error bound holds with equality when the truncated part of the model in the canonical balanced realization is state-space symmetric. In Chapter 4, we show how to derive these sign parameters (and thus the sign symmetry of the canonical balanced form) from a system's arrowhead realization using ideas from system theory and eigenvalue perturbation theory. We illustrate these results with an example from power systems modeling that motivated our study.

1.1 Linear Systems and the Model Reduction Problem

We consider SISO LTI dynamical systems given by the collection of ODEs

$$\mathcal{G}: \begin{cases} x'(t) = Ax(t) + bu(t) \\ y(t) = cx(t) + du(t), \end{cases} \tag{1.3}$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^{n\times 1}$, $c \in \mathbb{R}^{1\times n}$, and $d \in \mathbb{R}$. We call $x(t) \in \mathbb{R}^n$ the state of the system at time t, and $u(t)$, $y(t) \in \mathcal{L}_2([0,\infty))$ the input and output of the system, respectively. (The restriction of these functions to $\mathcal{L}_2$ can be interpreted as only permitting finite-energy inputs/outputs.) We refer to the *state-space* of a dynamical system $\mathbb{R}^n$ as the vector space containing all possible states of the system at time t. The matrices and vectors (A, b, c, d) describing the system is a *state-space realization* of $\mathcal{G}$. We refer to the first equation describing the evolution of $x(t)$ in (1.3) as the *state equation*, and the second as the *output equation* of $\mathcal{G}$. By taking a Laplace transformation of the state equation and some rearranging, one can derive the *transfer function* of a LTI system, given by

$$G(s) = c(sI - A)^{-1}b + d, \tag{1.4}$$

a scalar-valued rational function of a complex variable, s. This function describes the input-output mapping of $\mathcal{G}$ in the frequency domain, that is, $Y(s) = G(s)U(s)$, where $U(s)$ and $Y(s)$ are Laplace transformations of the system input and output, respectively. We say a realization of $\mathcal{G}$ is *minimal* if the dimension n of the realization is the smallest possible. The *poles* of a dynamical system are the poles of the function $G(s)$. For a minimal realization of $\mathcal{G}$, these poles are the points $s \in \mathbb{C}$ at which $sI - A$ fails to be invertible. (Equivalently, these poles are given by the eigenvalues of A.) Throughout our discussion, we consider systems that are *asymptotically stable*, that is, we assume that all eigenvalues of A lie in the open left half-plane. We are interested in reduced-order models that approximate $\mathcal{G}$ in the sense that this input-output mapping given by $G(s)$ is well-approximated for a wide variety of inputs. We now state the model reduction problem formally as follows. Given an order-n SISO LTI system $\mathcal{G}$ as in (1.3), we seek a reduced-order system

$$\mathcal{G}_r : \begin{cases} x_r'(t) = A_r x_r(t) + b_r u(t) \\ y_r(t) = c_r x_r(t) + d_r u(t), \end{cases} \tag{1.5}$$

having the transfer function

$$G_r(s) = c_r(sI - A_r)^{-1} b_r + d_r,$$

where $A_r \in \mathbb{R}^{r\times r}$, $b_r \in \mathbb{R}^{r\times 1}$, $c_r \in \mathbb{R}^{1\times r}$, $x_r(t) \in \mathbb{R}^r$, $d_r \in \mathbb{R}$, with $r \ll n$, such that $\mathcal{G}_r$ satisfies the following conditions:

(i) $\mathcal{G}_r$ approximates $\mathcal{G}$ in the sense that the reduced-order output $y_r(t)$ sufficiently mimics the full output $y(t)$ for a wide variety of inputs $u(t)$. Posed more formally, $\mathcal{G}_r$ approximates $\mathcal{G}$ if the output error $\|y - y_r\|$ is small in an appropriate norm.

(ii) Key structural and theoretical properties of the original full-order system $\mathcal{G}$, such as stability, passivity, or in more general cases, structure, are preserved in the reduced-order model.

(iii) The algorithm used in computing of the reduced-order model $\mathcal{G}_r$ is numerically feasible.

We next delve into the ideas that lead to a good ROM $\mathcal{G}_r$ of $\mathcal{G}$ that satisfies these conditions.

We say two systems $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ are *equivalent* if they have the same input-output mapping, $G(s) = \widetilde{G}(s)$. It is well-known that equivalent systems differ by an invertible transformation of the state-space $\mathbb{R}^n$. Formally, consider a nonsingular coordinate transformation $T \in \mathbb{R}^{n\times n}$

applied to the state vector $x(t)$. Then, the resulting system $\widetilde{\mathcal{G}}$ as in (1.3) given by the state space realization $\widetilde{A} = T^{-1}AT$, $\widetilde{b} = T^{-1}b$, $\widetilde{c} = cT$, $\widetilde{d} = d$, and having the state vector $\widetilde{x}(t) = T^{-1}x(t)$, is equivalent to $\mathcal{G}$. This can be verified directly by comparing the transfer functions of the two systems:

$$\begin{aligned}
\widetilde{G}(s) &= \widetilde{c}(sI - \widetilde{A})^{-1}\widetilde{b} + \widetilde{d} \\
&= cT^{-1}(sI - TAT^{-1})^{-1}Tb + d \\
&= cT^{-1}T(sI - A)^{-1}T^{-1}Tb + d \\
&= c(sI - A)^{-1}b + d \\
&= G(s).
\end{aligned}$$

This observation illustrates a key concept for LTI systems: the choice of internal state variables $x(t)$ used in describing the system is not significant to the input-output map. Instead, we view $\mathcal{G}$ as a "black-box" process mapping the input function $u(t)$ to the associated output $y(t)$. It is not uncommon that the trajectory of the system's state $x(t)$ can be well approximated by vectors in some lower order r-dimensional subspace of $\mathbb{R}^n$ with $r \ll n$, suggesting that one can accurately reduce the system order by removing states that are in some sense "insignificant" to the underlying dynamics. Further, because the choice of coordinate system in which the system evolves does not change the input-output mapping, one can perform reduction by simply removing unimportant states from the state space in a favorable representation of the system. This is the key idea behind *projection-based model reduction*, which we introduce next.

1.1.1 Model Reduction via Projection

Assume that the system dynamics of $\mathcal{G}$ are well-approximated by some r-dimensional subspace $\mathcal{V}_r$ of $\mathbb{R}^n$ with $r \ll n$; that is, there exists $\mathcal{V}_r = \text{range}(V_r)$, with $V_r \in \mathbb{R}^{n\times r}$, such that $x(t) \approx V_r x_r(t)$. Using this approximation would lead to inexact dynamics, and so, we orthogonalize the resulting error against a second subspace of interest $\mathcal{W}_r = \text{range}(W_r)$, $W_r \in \mathbb{R}^{n\times r}$, such that $W_r^\top V_r = I_r$, the $r \times r$ identity matrix. This orthogonality condition is known as the *Petrov-Galerkin condition*:

$$W_r^\top (V_r x_r'(t) - AV_r x_r(t) - bu(t)) = 0. \tag{1.6}$$

Rearranging (1.6) produces an order-r reduced order model $\mathcal{G}_r$ as in (2.14) obtained *via projection*, where

$$A_r = W_r^\top A V_r, \quad b_r = W_r^\top b, \quad c_r = cV_r, \quad d_r = d.$$

This idea is central to many popular model reduction algorithms.

We wish to make precise the idea that a reduced-order system $\mathcal{G}_r$ should approximate $\mathcal{G}$ in the sense that $y_r \approx y$. For a consistent input $u(t)$, we seek a reduced order system $\mathcal{G}_r$ having transfer function $G_r(s)$ approximating $\mathcal{G}$, such that the resulting output error $y - y_r$ is small in an appropriate norm. If this holds for all inputs, this error can be bounded in a "worst-case" sense. Different norms can be used to measure performance; of particular interest to us is the $\mathcal{H}_\infty$ norm of an asymptotically stable system, defined as

$$\|\mathcal{G}\|_{\mathcal{H}_\infty} = \sup_{\omega\in\mathbb{R}} |G(i\omega)|, \tag{1.7}$$

where i is the imaginary unit such that $i^2 = -1$. Suppose we want to ensure a small output error $y_r - y$ in the $\mathcal{L}_2$ sense. First note that for an arbitrary bounded input $u \in \mathcal{L}_2([0,\infty))$, in the frequency domain we have that

$$Y(s) - Y_r(s) = (G(s) - G_r(s))U(s).$$

According to the Plancherel's theorem, the norm of a matrix-valued function in the frequency domain and in time domain differ up to a constant multiple. So, the $\mathcal{L}_2$ error of the output can be written as

$$\begin{aligned}
\|y - y_r\|_{\mathcal{L}_2}^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |Y(i\omega) - Y_r(i\omega)|^2 d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} |(G(i\omega) - G_r(i\omega))U(i\omega)|^2 d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(i\omega) - G_r(i\omega)|^2 |U(i\omega)|^2 d\omega \\
&\leq \|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty}^2 \frac{1}{2\pi}\int_{-\infty}^{\infty} |U(i\omega)|^2 d\omega \\
&= \|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty}^2 \|u\|_{\mathcal{L}_2}^2.
\end{aligned}$$

A similar bound can be derived for the $\mathcal{L}_\infty$ error of the approximation $y_r \approx y$ using a different measure of system error. Note the important observation that by making the $\mathcal{H}_\infty$ error small, i.e., by finding a rational function G_r that is a good approximation to G, one can make the resulting output error small. This sufficiently accurate low-order approximation $\mathcal{G}_r$ to $\mathcal{G}$ can be obtained by the above projection-based framework, so long as the result yields a system with a transfer function $G_r(s)$ sufficiently approximating $G(s)$.

Chapter 2

Balanced Truncation Model Reduction

In order to establish notation, we review the necessary ideas and theory behind balanced truncation model reduction. We begin by introducing the state-theoertic concepts of reachability and observability, and their connection to the system Gramians. Afterwards, we establish the idea of balancing a linear system, and its natural connection to model-order reduction. We then discuss the useful theoretical properties of the method.

2.1 Reachability and Observability of a State

We begin by discussing the ideas of the reachability and observability of a state, and establish the use of these ideas in MOR.

Given an LTI system $\mathcal{G}$, we say a state $\widehat{x} \in \mathbb{R}^n$ is *reachable* if there exists an input $u \in \mathcal{L}_2([0,\infty))$ and finite time $T < \infty$ such that $\widehat{x}$ is the solution to the state equation as in (1.3) at time T subject to the initial condition $x_0 = 0$ and input $u(t)$. We say $\mathcal{G}$ if *fully reachable* if every state in the state-space has this property. Roughly speaking, we think of reachable states as those to which we can drive the system dynamics in finite time with a finite energy input.

Similarly, we say a state $x_0 \in \mathbb{R}^n$ is *unobservable* if the output $y(t) = cx(t) = 0$ for all $t \geq 0$, where $x(t)$ is the solution to the state equation as in (1.3) subject to the initial condition x_0 and zero forcing $u(t) = 0$. We say a $\mathcal{G}$ is *fully observable* if only the zero state is unobservable. Intuitively, a state is *unobservable* if it is indistinguishable from the zero state by observing the output of the system.

The significance of these ideas for our purposes lies in their application to MOR, as we can apply the ideas in Chapter 1 and categorize the significance of a state to the underlying system dynamics by its degree of reachability and observability. To investigate this further, we introduce the *system Gramians*.

2.1.1 The System Gramians

The *reachability Gramian* $\mathcal{P} \in \mathbb{R}^{n \times n}$ and *observability Gramain* $\mathcal{Q} \in \mathbb{R}^{n \times n}$ of an asymptotically stable LTI system $\mathcal{G}$ as in (1.3) are defined as the unique solutions to the Lyapunov equations

$$A\mathcal{P} + \mathcal{P}A^\top + bb^\top = 0 \quad \text{and} \quad A^\top\mathcal{Q} + \mathcal{Q}A + c^\top c = 0. \tag{2.1}$$

Given the asymptotic stability assumption, these objects have the useful integral representations

$$\mathcal{P} = \int_0^\infty e^{At} bb^\top e^{A^\top t} \, dt \tag{2.2}$$

and

$$\mathcal{Q} = \int_0^\infty e^{A^\top t} c^\top c e^{At} \, dt. \tag{2.3}$$

It is readily verifiable that $\mathcal{P}$ and $\mathcal{Q}$ as defined in (2.2) and (2.3) solve the corresponding Lyapunov equations (2.1). The Gramians $\mathcal{P}$ and $\mathcal{Q}$ can also be formulated in the frequency domain by taking Laplace transformations of (2.2) and (2.3) and applying the Plancherel's theorem. This yields

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(A - i\omega I)^{-1} b b^{\top}(A^{\top} + i\omega I)^{-1}\, d\omega \tag{2.4}$$

and

$$\mathcal{Q} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(A^{\top} - i\omega I)^{-1} c^{\top} c(A + i\omega I)^{-1}\, d\omega. \tag{2.5}$$

We now state the following fact; highlighting the connection of these system Gramians to the ideas of reachability and observability.

Theorem 2.1. *Let $\mathcal{G}$ be an asymptotically stable LTI system as in (1.3). Then $\mathcal{G}$ is fully reachable (observable) if and only if $\mathcal{P}$ ($\mathcal{Q}$) is symmetric positive definite.*

This also implies a system $\mathcal{G}$ is fully reachable and observable if and only if the system Gramians are full rank. We additionally note the following useful characterization of the minimality of a system $\mathcal{G}$ in terms of reachability and observability.

Definition 2.2. An LTI system $\mathcal{G}$ is minimal if it is both reachable and observable.

2.1.2 Gramian-based MOR

Part of the significance of these Gramians lies in their use in determining the *degree* to which a state is reachable or observable. We define the *reachability energy* $\mathcal{E}_{reach}$ of a state $\widehat{x} \in \mathbb{R}^n$ as the minimal energy required to drive the dynamics from $x_0 = 0$ to $\widehat{x}$ in some finite time. Similarly, we define the *observability energy* $\mathcal{E}_{obsv}$ of a state $x_0 \in \mathbb{R}^n$ as the energy obtained by observing the system having initial condition x_0 under zero forcing. Formally, these quantities are defined as

$$\mathcal{E}_{reach}(\widehat{x}) = \min\{\|u\|_{\mathcal{L}_2} : \widehat{x} \text{ is reachable by } u\}$$

and

$$\mathcal{E}_{obsv}(x_0) = \|y\|_{\mathcal{L}_2} \text{ where } y(t) = cx(t).$$

We have the more approachable characterizations of the reachability energy and observability energy of a state in terms of the system Gramians. These results are from [9, Lemma 4.29].

Theorem 2.3. *Let $\mathcal{G}$ be an LTI system as in (1.3) having the reachability Gramian $\mathcal{P}$ as in (2.4) and observability Gramian $\mathcal{Q}$ as in (2.5). Suppose $\widehat{x} \in \mathbb{R}^n$ is reachable. Then, the reachability energy of $\widehat{x}$ is given by*

$$\mathcal{E}_{reach}(\widehat{x}) = \widehat{x}^\top \mathcal{P}^{-1} \widehat{x}. \tag{2.6}$$

Suppose $x_0 \in \mathbb{R}^n$ is observable. Then, the observability energy of x_0 is given by

$$\mathcal{E}_{obsv}(x_0) = x_0^\top \mathcal{Q} x_0. \tag{2.7}$$

We think of states that are "hard to reach" as those having a high reachability energy and states that are "hard to observe" as those having a low observability energy. Note that this characterization makes intuitive sense: states with high reachability energies (2.6) are those that can only be reached with high energy inputs. The system is unlikely to arrive at such states naturally and so they likely do not contribute much to the underlying dynamics. Similarly, if a state x_0 has a low observability energy (2.7) then it does not contribute much to the corresponding output $y(t)$ for $t \geq 0$, and is insignificant in this sense. This suggests a natural means of reducing the model order. If one wants to remove states that are insignificant to the system dynamics, then one should move states that are hard to reach or hard to observe.

We can formalize this notion as follows. Consider an LTI system $\mathcal{G}$ and a reachable state $\widehat{x} \in \mathbb{R}^n$. Let $\mathcal{P} = VMV^\top$ be the singular value decomposition of $\mathcal{P}$ (note that this form follows from the fact that $\mathcal{P}$ is symmetric). Plugging this into (2.6) with the change of variable $z = V^\top \widehat{x}$ gives

$$\mathcal{E}_{reach}(\widehat{x}) = \widehat{x}^\top \mathcal{P}^{-1} \widehat{x} = \widehat{x}^\top V M^{-1} V^\top \widehat{x} = \sum_{i=1}^{n} \mu_i^{-1} |z_i|^2,$$

where $M = \text{diag}(\mu_1, \ldots, \mu_n)$ are the singular values of $\mathcal{P}$ with $\mu_1 \geq \cdots \geq \mu_n > 0$, and z_i is the component of $\widehat{x}$ in the span of $\mathcal{P}$'s ith singular vector. Note the reachability energy of $\widehat{x}$ is *maximized* when $\widehat{x}$ has a large component in the span of the singular vectors of $\mathcal{P}$ corresponding to small singular values. In other words, we can see states that are hard to reach are *precisely* those corresponding to the trailing singular values of $\mathcal{P}$ (in this basis of $\mathcal{P}$'s singular vectors). One can apply the same analysis to the observability energy of a state x_0 to observe that $\mathcal{E}_{obsv}(x_0)$ is *minimized* when x_0 has a large component in the span of the singular vectors of $\mathcal{Q}$ corresponding to small singular values. And so similarly, states that are hard to observe are precisely those corresponding to the trailing singular values of $\mathcal{Q}$ (in the basis of $\mathcal{Q}$'s singular vectors). Thus, if one wishes to obtain an accurate reduced-order model to $\mathcal{G}$, then one should project the system dynamics onto the leading singular vectors of either $\mathcal{P}$ or $\mathcal{Q}$.

Unfortunately, this methodology has certain drawbacks. Notably, the reachability and observability of a state are *basis-dependent*. Indeed, if one performs a change of coordinate transformation $\widetilde{x} = Tx$ on the state-space, then the Gramians $\widetilde{\mathcal{P}}$ and $\widetilde{\mathcal{Q}}$ under this transformation are given respectively by

$$\widetilde{\mathcal{P}} = T\mathcal{P}T^{\top}, \quad \text{and} \quad \widetilde{\mathcal{Q}} = T^{-\top}\mathcal{Q}T^{-1}, \tag{2.8}$$

which can be verified by making the corresponding change of variable in the Lyapunov equations (2.1). So in any given basis, states that are hard to reach may not be hard to observe (and vice versa). This leads to the natural question of whether or not there exists a basis more suitable for reduction; namely, can one find a coordinate system in which states that are hard to reach are simultaneously hard to observe? The answer is yes, and such a basis is called a *balanced basis* for the system.

2.2 Balanced Truncation Model Reduction

2.2.1 Balancing Transformations

In a balanced basis those states that are hard to reach are simultaneously hard to observe. Finding such a basis corresponds to simultaneous diagonalization of the system Gramians $\mathcal{P}$ and $\mathcal{Q}$, leading us to the following definition. We say a system $\mathcal{G}$ is *principal-axis balanced* if the system Gramians satisfy $\mathcal{P} = \mathcal{Q} = \Sigma$, where

$$\Sigma = \operatorname{diag}(\sigma_1 I_{m_1}, \ldots, \sigma_q I_{m_q}), \tag{2.9}$$

with $\sigma_1 > \sigma_2 > \cdots > \sigma_q > 0$ and $m_1 + \cdots + m_q = n$. Here $\sigma_1, \ldots, \sigma_q$ are the *Hankel singular values* of $\mathcal{G}$. For brevity, we will use the term *balanced* as shorthand for principal-axis balanced moving forward. The state-space realization in balanced coordinates is called a *balanced realization* of $\mathcal{G}$. In this basis, the singular values of the system Gramians are trivially given by the Hankel singular values Σ; hence the alignment of hard to reach states with hard to observe ones. Further, such states are those associated with the trailing Hankel singular values of $\mathcal{G}$. This is the main idea behind *balanced truncation* model reduction. It proceeds as follows: First, one computes a *balancing transformation* T of a system $\mathcal{G}$ under which the transformed Gramians (2.8) satisfy $\widetilde{\mathcal{P}} = \widetilde{\mathcal{Q}} = \Sigma$ (i.e., one *balances* $\mathcal{G}$). Secondly, one reduces the model order by truncating states corresponding to small Hankel singular values from the state-space. Note that this is the idea behind balanced truncation theoretically, not numerically. In practice, the system of interest is simultaneously balanced and truncated; the full balanced system is never formed (see Remark 2.7).

Remark 2.4. Note that the Hankel singular values (2.9) are *system invariants*, that is, they are a property of the system $\mathcal{G}$ itself and are invariant under coordinate transformations of the state-space. This calls back to the general model reduction framework as described in Chapter 1. In the balanced basis, states are easily characterizable in terms of their significance. So, we can transform the state-space into the balanced basis and truncate these insignificant states.

First, we note that for every minimal and stable LTI system $\mathcal{G}$ can be balanced. This fact is from [9, Lemma 7.3].

Theorem 2.5. *Let $\mathcal{G}$ be a minimal and asymptotically stable LTI system with system Gramians $\mathcal{P}$ and $\mathcal{Q}$. Let the Cholesky factorization of $\mathcal{P}$ be given by*

$$\mathcal{P} = UU^\top$$

and the eigenvalue decomposition of $U^\top \mathcal{Q} U$ be given by

$$U^\top \mathcal{Q} U = K\Sigma^2 K^\top.$$

Then, a balancing transformation of $\mathcal{G}$ is given by

$$T = \Sigma^{1/2} K^\top U^{-1} \quad \textit{and} \quad T^{-1} = UK\Sigma^{1/2}. \tag{2.10}$$

It is easily verifiable that this transformation balances $\mathcal{G}$ by applying T and T^{-1} to (2.8) and noting that $\widetilde{\mathcal{P}} = \widetilde{\mathcal{Q}} = \Sigma$ under this transformation.

2.2.2 Balanced Truncation in Theory

Once balanced, one reduces the order of a system by removing components of the state-space corresponding to small Hankel singular values, which amounts truncating the trailing components of the state vector $x \in \mathbb{R}^n$. This can be seen by formulating balanced truncation in the projection framework outlined in Subsection 1.1.1. Assuming we have a balanced system $\mathcal{G}$, an equivalent way to express the reduction is by projecting the system dynamics onto the dominant singular vectors of the system Gramians. Because $\mathcal{G}$ is balanced, $\mathcal{P} = \mathcal{Q} = \Sigma$, and so their singular vectors are trivially the standard basis vectors of $\mathbb{R}^n$. So, we take $W_r = V_r = \begin{bmatrix} I_r & 0 \end{bmatrix}^\top \in \mathbb{R}^{n \times r}$ as our projection bases, where I_r is the $r \times r$ identity matrix, and the order-r reduced model via balanced truncation is given by the leading r-dimensional submatrices of $\mathcal{G}$'s balanced realization. This is stated formally in Theorem 2.6.

In addition to its elegance, balanced truncation enjoys two main properties: preservation of stability in the reduced-order model, and a simple *a priori* error bound in terms of the neglected Hankel singular values. This is given in Theorem 2.6 below. The asymptotic stability result is due to Pernebo and Silverman [10] and the error bound (2.15) is due to Enns [11].

Theorem 2.6. *Let $\mathcal{G}$ be an order-n minimal and asymptotically stable dynamical system having the balanced realization*

$$\left[\begin{array}{c|c} A & b \\ \hline c & d \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & A_{12} & b_1 \\ A_{21} & A_{22} & b_2 \\ \hline c_1 & c_2 & d \end{array}\right], \tag{2.11}$$

where the state space matrices are partitioned according to the system Gramian $\Sigma = diag(\Sigma_1, \Sigma_2)$, with

$$\Sigma_1 = diag(\sigma_1 I_{m_1}, \ldots, \sigma_k I_{m_k}) \tag{2.12}$$

$$\Sigma_2 = diag(\sigma_{k+1} I_{m_{k+1}}, \ldots, \sigma_q I_{m_q}), \tag{2.13}$$

where $r := m_1 + \cdots + m_k$. Then the rth order reduced model obtained via balanced truncation

$$\mathcal{G}_r : \begin{cases} x_r'(t) = A_{11} x_r(t) + b_1 u(t) \\ y_r(t) = c_1 x_r(t) + du(t) \end{cases} \tag{2.14}$$

having the transfer function

$$G_r(s) = c_1(sI - A_{11})^{-1} b_1 + d$$

is asymptotically stable and satisfies the error bound

$$\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty} \leq 2(\sigma_{k+1} + \cdots + \sigma_q). \tag{2.15}$$

Remark 2.7. As previously mentioned, in practice one does not proceed with balanced truncation as outlined above. The most common implementation is via the so-called *square-root algorithm*, which is more numerically stable. In this implementation, balancing transformation and truncation are performed simultaneously. For further details, see [9, Sec. 7.4].

The rest of this work is related to the balanced truncation error bound (2.15). In Chapter 3, we explore conditions in which the error bound (2.15) is known to hold with equality. In particular, we describe new conditions for when this holds based on the canonical balanced form and the sign parameters of a system $\mathcal{G}$.

Chapter 3

On the Balanced Truncation Error Bound

The balanced truncation error bound (2.15) is known to hold with equality under certain conditions. Most notably, the bound is tight when only one Hankel singular value is truncated; that is $\Sigma_2 = \sigma_q I_{m_q}$. We next explore other situations in which the balanced truncation error bound (2.15) holds with equality, thus providing an exact formula for the error in the reduction. We begin by describing a crucial sign symmetry property of balanced systems, as characterized in [12, 13], that will be central to our results. Note that this chapter is a lightly edited version of the work [14].

3.1 The Canonical Form of a Balanced System

Theorem 3.1. *(Wilson and Kumar [12], Ober [13]) Let $\mathcal{G}$ be an asymptotically stable and minimal SISO LTI system as in (1.3). Then $\mathcal{G}$ has a balanced realization satisfying*

$$A = SA^{\top}S \quad \textit{and} \quad b = (cS)^{\top}, \tag{3.1}$$

where $S = diag(s_1, s_2, \ldots, s_n)$ and $s_i \in \{\pm 1\}$ for $i = 1, \ldots, n$.

We refer to these $s_1, \ldots, s_n$ as the *sign parameters* of a system $\mathcal{G}$. Ober [13, 15] shows that every asymptotically stable and minimal SISO LTI system $\mathcal{G}$ is equivalent to a balanced system having a realization satisfying (3.1), which we refer to as the *canonical form* of a balanced system. We also occasionally refer to this as the *canonical balanced realization* of a system $\mathcal{G}$. Each sign parameter is naturally associated with one of the system's Hankel singular values; that is, each s_i corresponds to σ_j for some $j = 1, \dots, q$. For a system with distinct Hankel singular values, the ith sign parameter s_i corresponds directly to the ith Hankel singular value, with $\sigma_1 > \cdots > \sigma_n$. For a system with repeated Hankel singular values, multiple sign parameters can correspond to the same Hankel singular value. In particular, if we consider a system having a Hankel singular value σ_k of multiplicity $m_k > 1$, then the m_k sign parameters corresponding to σ_k are given by the diagonal entries of $s_k \widehat{I}_{m_k} \in \mathbb{R}^{m_k \times m_k}$ where $\widehat{I}_{m_k} = \mathrm{diag}(1, -1, 1, \dots) \in \mathbb{R}^{m_k \times m_k}$ and $s_k = \pm 1$ (for further details, see [16, Section 2.7]). In Chapter 4, we will discuss these sign parameters further, and show how to determine them from a system's arrowhead realization, when available.

Remark 3.2. We note that any balanced realization of a system $\mathcal{G}$ is unique up to orthogonal transformations of the state-space [7], permitting multiple balanced realizations of a system $\mathcal{G}$ that obey same sign symmetry (3.1) with different permutations of the signs on the diagonal of S. In these realizations, the associated balanced coordinates are *not* ordered in decreasing significance (in terms of the Hankel singular values), in contrast to the canonical form.

For an order-n system $\mathcal{G}$ satisfying the hypotheses of Theorem 3.1 and having distinct Hankel singular values $\sigma_1 > \cdots > \sigma_n$, the canonical balanced realization of $\mathcal{G}$ satisfying (3.1) can be written explicitly as in [9, Eq. 7.24]:

$$a_{ij} = \frac{-\gamma_i \gamma_j}{s_i s_j \sigma_i + \sigma_j}, \quad b_i = \gamma_i, \quad c_i = s_i \gamma_i, \qquad i, j = 1, \dots, n, \tag{3.2}$$

where $\gamma_i > 0$ for all $i = 1, \dots, n$. Note that this realization of the system is completely determined by the distinct Hankel singular values $\sigma_1 > \cdots > \sigma_n > 0$, the signs $s_1, \dots, s_n \in \{\pm 1\}$, and the entries of the input vector $b \in \mathbb{R}^n$. Conversely, any SISO LTI system having a realization defined by (3.2) is guaranteed to be balanced, minimal, and asymptotically stable [13].

We emphasize the connection that the sign parameters of a system $\mathcal{G}$ explicitly determine the sign symmetry of its balanced realization given in (3.1). Conversely, the sign parameters of $\mathcal{G}$ can be deduced from this canonical form. Further, they are a *system invariant*, just like the Hankel singular values (see Remark 2.4). While they can be derived from the canonical form of $\mathcal{G}$, the parameters are not basis dependent, but rather a property of the system itself.

Now suppose $\mathcal{G}$ is balanced with the canonical form (3.1). Define $r := m_1 + \cdots + m_k$ for some k with $1 \leq k < q$, and partition the sign matrix given in (3.1) as

$$S = \operatorname{diag}(S_1, S_2),$$

where

$$S_1 = \operatorname{diag}(s_1, \dots, s_k) \quad \text{and} \quad S_2 = \operatorname{diag}(s_{k+1}, \dots, s_n). \tag{3.3}$$

Partition A, b, and c as in (2.11). Then direct multiplication of (3.1) gives the sign symmetries

$$A_{11} = S_1 A_{11}^\top S_1, \quad A_{12} = S_1 A_{21}^\top S_2, \quad A_{21} = S_2 A_{12}^\top S_1, \tag{3.4}$$

$$A_{22} = S_2 A_{22}^\top S_2, \quad b_1 = (c_1 S_1)^\top, \text{ and } b_2 = (c_2 S_2)^\top. \tag{3.5}$$

Because the reduced-order model obtained via balanced truncation is independent of the initial system realization, for the time being we assume without loss of generality that any system $\mathcal{G}$ we work with is already balanced, having the realization given in (3.1) and defined by (3.2). So, the Lyapunov equations in (2.1) become

$$A\Sigma + \Sigma A^{\top} + bb^{\top} = 0 \quad \text{and} \quad A^{\top}\Sigma + \Sigma A + c^{\top}c = 0. \tag{3.6}$$

3.1.1 State-space Symmetric Systems

Definition 3.3. If an LTI SISO system $\mathcal{G}$ has a realization satisfying (3.1) with $S = I$, then $\mathcal{G}$ is called *state-space symmetric*. In this case, $A = A^{\top}$ and $b = c^{\top}$.

Alternatively, we can define any system $\mathcal{G}$ having a realization such that $A = A^{\top}$ and $b = c^{\top}$ to be state-space symmetric, but Definition 3.3 is more applicable in the context of our discussion. As mentioned previously, state-space symmetric systems have the property that the balanced truncation error bound holds with equality for *any* truncation order r. Before extending this result to more general class of systems in Section 3.2, we we prove that the Hankel singular values of systems with a slight generalization of state-space symmetry (allowing $S = \pm I$) have multiplicity one. This result has been already pointed out in [17]; here we provide a different proof based on the balanced canonical form of a system. An analogous result holds for zero interlacing pole (ZIP) systems, which are closely related to state-space symmetric systems in the SISO case [18].

Proposition 3.4. *Let $\mathcal{G}$ be an asymptotically stable, minimal, balanced SISO LTI system as in (1.3), and suppose $\mathcal{G}$ has a balanced realization (3.1) with either $S = I$ or $S = -I$. Then the Hankel singular values of $\mathcal{G}$ must all have multiplicity one, that is, $m_1 = \cdots = m_n = 1$.*

Proof. Suppose $\mathcal{G}$ has some Hankel singular value σ of multiplicity m. The balanced realization of a system is unique up to orthogonal transformations [9, Corollary 7.4]. So by [15, Corollary 2.1], there exists a unitary matrix $U \in \mathbb{R}^{n\times n}$ such that the upper-left $m \times m$ block of $\widehat{A} := U^\top A U$ has the form

$$\widehat{A}(1:m,1:m) = \begin{bmatrix} \frac{-\gamma^2}{2\sigma} & \alpha_1 & & \\ -\alpha_1 & 0 & \ddots & \\ & \ddots & \ddots & \alpha_{m-1} \\ & & -\alpha_{m-1} & 0 \end{bmatrix}$$

for some $\gamma \in \mathbb{R}$ and $\alpha_1, \ldots, \alpha_{m-1} > 0$. However, note that $S = \pm I$ implies $A = A^\top$, and so $\widehat{A}^\top = (U^\top A U)^\top = \widehat{A}$; such symmetry is impossible if $m > 1$, given the form of $\widehat{A}(1:m,1:m)$. Thus, $\mathcal{G}$ cannot have a Hankel singular value of multiplicity greater than one, and so all of $\mathcal{G}$'s Hankel singular values must have multiplicity $m = 1$. □

We now restate the following important result for state-space symmetric systems; namely, they have the property that the balanced truncation error bound holds with equality for any truncation order r. This can be obtained using results from [8].

Theorem 3.5. *Let $\mathcal{G}$ be an asymptotically stable LTI SISO system that is state-space symmetric. Let $\mathcal{G}_r$ denote the order-r approximation to $\mathcal{G}$ obtained via balanced truncation as in* (2.11)*. Then, the Hankel singular values of $\mathcal{G}$ are distinct, $\mathcal{G}_r$ is state-space symmetric, and $\mathcal{G}_r$ achieves the error bound* (2.15)*:*

$$\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty} = 2(\sigma_{r+1} + \cdots + \sigma_n).$$

3.2 A Tight Error Bound when the Truncated Sign Parameters are Consistent

We now extend Theorem 3.5 to a more general class of systems than the state-space symmetric ones previously described. To this end, we define the *truncated system*. In accordance with the partitioning of a system $\mathcal{G}$'s balanced realization (2.11), let the truncated system corresponding to the order-r balanced approximation to $\mathcal{G}$ be given by

$$\mathcal{G}^{tr} : \begin{cases} x_2'(t) = A_{22}x_2(t) + b_2u(t), \\ y_2(t) = c_2\widetilde{x}(t), \end{cases} \tag{3.7}$$

with the transfer function $G^{tr}(s) = c_2(sI - A_{22})^{-1}b_2$ and system Gramian

$$\Sigma_2 = \text{diag}(\sigma_{k+1}I_{m_k+1}, \ldots, \sigma_q I_{m_q}).$$

Note this implies that the sign parameters of $\mathcal{G}^{tr}$ are given by S_2 as in (3.3). We call (3.7) the *truncated system*, since its state $x_2(t) \in \mathbb{R}^{n-r}$ corresponds to the truncated states in the balanced realization of $\mathcal{G}$. Note from (3.5) that $\mathcal{G}^{tr}$ satisfies the sign symmetries $A_{22} = S_2A_{22}^\top S_2$ and $b_2 = (c_2S_2)^\top$. In the subsequent results, we relax the assumption that the full-order system be state-space symmetric, and show that the balanced truncation error bound holds with equality when sign parameters corresponding to the truncated Hankel singular values are the same. That is, we allow the signs in S_1 to vary; we only assume that $S_2 = I_{n-r}$ or $S_2 = -I_{n-r}$. Put alternatively, this assumption can be stated as the sign parameters of the *truncated* system being consistent.

Theorem 3.6. *Let $\mathcal{G}$ be an order-n asymptotically stable, minimal, balanced SISO system as in (1.3), with its matrix of Hankel singular values $\Sigma = diag(\Sigma_1, \Sigma_2)$ partitioned as in*

(2.12) *and* (2.13)*; set* $r := m_1 + \cdots + m_k$*. Let* $\mathcal{G}_r$ *denote the order-*r *approximation of* $\mathcal{G}$ *via balanced truncation, as in* (2.14)*. Conformally partition the sign matrix,* $S = diag(S_1, S_2)$*, with* $S_1 \in \mathbb{R}^{r\times r}$ *and* $S_2 \in \mathbb{R}^{(n-r)\times(n-r)}$*. If all the signs in* S_2 *are the same, i.e.,*

$$S_2 = diag\,(+1, \ldots, +1) \;\; or \;\; S_2 = diag\,(-1, \ldots, -1), \tag{3.8}$$

then the truncated Hankel singular values are distinct,

$$\Sigma_2 = diag(\sigma_{k+1}, \ldots, \sigma_q) \tag{3.9}$$

and $\mathcal{G}_r$ *in* (2.14) *achieves the balanced truncation error bound* (2.15)*:*

$$\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty} = 2(\sigma_{k+1} + \cdots + \sigma_q).$$

Proof. First note that the truncated system $\mathcal{G}^{tr}$ satisfies the conditions of Proposition 3.4, so we can conclude that the associated Hankel singular values in (3.9) are distinct. To prove that the bound is tight, we begin by repeating the necessary details of the balanced truncation error bound proof; see [11] and [19, Thm. 7.3]. Recall the definition of the $\mathcal{H}_\infty$-norm in (1.7). Because $\mathcal{G}_r$ is obtained via balanced truncation, the error system satisfies the upper bound (2.15), so it suffices to show that this bound is attained for the particular frequency $\omega = 0$, i.e., $s := \mathrm{i}\omega = 0$. Assume without loss of generality that $\mathcal{G}$ is balanced and in the canonical form (3.1). Then the state-space matrices satisfy the sign symmetries given in (3.4) and (3.5). Following [19, Thm. 7.3], the transfer function for the error system can be written as

$$\begin{aligned} G(s) - G_r(s) &= c(sI - A)^{-1}b - c_1(sI - A_{11})^{-1}b_1 \\ &= \widetilde{c}(s)\psi(s)^{-1}\widetilde{b}(s), \end{aligned}$$

where

$$\begin{aligned} \psi(s) &= sI - A_{22} - A_{21}(sI - A_{11})^{-1}A_{12}, \\ \widetilde{b}(s) &= A_{21}(sI - A_{11})^{-1}b_1 + b_2, \\ \widetilde{c}(s) &= c_1(sI - A_{11})^{-1}A_{12} + c_2. \end{aligned}$$

We claim that the error bound is achieved at $s = 0$. Observe that

$$\psi(0)^\top = -A_{22}^\top + (A_{21}A_{11}^{-1}A_{12})^\top = -A_{22} + A_{12}^\top A_{11}^{-\top} A_{21}^\top,$$

since $A_{22} = S_2 A_{22}^\top S_2$ by (3.5), and $S_2 = \pm I_{n-r}$. Now by (3.4),

$$\begin{aligned} \psi(0)^\top &= -A_{22} + S_2 A_{21} S_1 S_1 A_{11}^{-1} S_1 S_1 A_{12} S_2 \\ &= -A_{22} + A_{21}A_{11}^{-1}A_{12} \\ &= \psi(0), \end{aligned}$$

and hence $\psi(0)$ is a symmetric matrix. (The sign assumption $S_2 = \pm I_{n-r}$ is only needed to establish this symmetry.) Now evaluate the error at $s = 0$:

$$\begin{aligned}|G(0) - G_r(0)|^2 &= \big(\widetilde{c}(0)\psi(0)^{-1}\widetilde{b}(0)\big)^\top\big(\widetilde{c}(0)\psi(0)^{-1}\widetilde{b}(0)\big)\\ &= \widetilde{b}(0)^\top\psi(0)^{-\top}\widetilde{c}(0)^\top\widetilde{c}(0)\psi(0)^{-1}\widetilde{b}(0)\\ &= \operatorname{tr}\big(\widetilde{b}(0)^\top\psi(0)^{-\top}\widetilde{c}(0)^\top\widetilde{c}(0)\psi(0)^{-1}\widetilde{b}(0)\big)\\ &= \operatorname{tr}\big(\widetilde{b}(0)\widetilde{b}(0)^\top\psi(0)^{-\top}\widetilde{c}(0)^\top\widetilde{c}(0)\psi(0)^{-1}\big),\end{aligned}$$

where $\operatorname{tr}(\cdot)$ denotes the trace, and we have used the invariance of this operation under cyclic permutation. Since $\widetilde{b}(0)\widetilde{b}(0)^\top$ and $\widetilde{c}(0)^\top\widetilde{c}(0)$ are both rank-one matrices (a consequence of the system being SISO), the matrix in this last trace expression has at most one nonzero eigenvalue, and so the trace (the sum of the eigenvalues) is simply that eigenvalue. Since $|G(0)-G_r(0)|^2$ is nonnegative, this eigenvalue must be real and nonnegative, and we designate it with $\lambda_{\max}(\cdot)$:

$$|G(0) - G_r(0)|^2 = \lambda_{\max}\big(\widetilde{b}(0)\widetilde{b}(0)^\top\psi(0)^{-\top}\widetilde{c}(0)^\top\widetilde{c}(0)\psi(0)^{-1}\big).$$

One can show that $\widetilde{b}(0)\widetilde{b}(0)^\top$ and $\widetilde{c}(0)^\top\widetilde{c}(0)$ are the right-hand sides of Lyapunov equations that both have Σ_2 as their solution:

$$\Sigma_2\psi(0)^\top + \psi(0)\Sigma_2 = \widetilde{b}(0)\widetilde{b}(0)^\top, \tag{3.10}$$

$$\Sigma_2\psi(0) + \psi(0)^\top\Sigma_2 = \widetilde{c}(0)^\top\widetilde{c}(0). \tag{3.11}$$

(See [19, Sect. 7.1] for further details.) Substituting the left-hand sides into the last expression for $|G(0) - G_r(0)|^2$ gives

$$|G(0) - G_r(0)|^2 = \lambda_{\max}\big((\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-\top})(\Sigma_2 + \psi(0)^\top\Sigma_2\psi(0)^{-1})\big).$$

Since $\psi(0)$ is symmetric,

$$\begin{aligned} |G(0) - G_r(0)|^2 &= \lambda_{\max}\big((\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-1})^2\big) \\ &= \big(\lambda_{\max}(\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-1})\big)^2. \end{aligned}$$

We emphasize that $\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-1}$ must be rank-one by equations (3.10) and (3.11), and thus has one nonzero eigenvalue. Additionally note that Σ_2 and $\psi(0)\Sigma_2\psi(0)^{-1}$ have identical eigenvalues, since the latter is a similarity transformation of the former. This fact and the aforementioned rank-one structure allow us to conclude that

$$\begin{aligned} |G(0) - G_r(0)| &= \lambda_{\max}\left(\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-1}\right) \\ &= \operatorname{tr}(\Sigma_2 + \psi(0)\Sigma_2\psi(0)^{-1}) \\ &= \operatorname{tr}(\Sigma_2) + \operatorname{tr}(\psi(0)\Sigma_2\psi(0)^{-1}) \\ &= 2\operatorname{tr}(\Sigma_2) \\ &= 2(\sigma_{k+1} + \cdots + \sigma_q), \end{aligned}$$

completing the proof. □

Note that this result also neatly generalizes the sharpness of the balanced truncation error bound when only one Hankel singular value is truncated: in this case the hypothesis of Theorem 3.6 is always satisfied, since the truncated system contains only one distinct sign parameter.

Remark 3.7. We note Theorem 3.6 can also be proven in a different context, namely by adapting results from [17], which show an $\mathcal{H}_\infty$ lower bound on the error in balanced truncation model reduction for systems with semi-definite Hankel operators. While [17, Proposition 13] is stated for systems that are semi-definite, its proof only requires that the *truncated system* be semi-definite. With this insight, we can apply [17, Corollary 14] to the systems in Theorem 3.6, showing that the lower bound on the balanced truncation error in [17, Corollary 14] holds with equality. We thank Christian Himpe for bringing this reference [17] to our attention.

3.2.1 An Example with Flipped Signs and a Strict Bound

We illustrate Theorem 3.6 with a synthetic example showing how the balanced truncation error bound holds with equality when the truncated system obeys the sign consistency in (3.8); that is, when the sign parameters corresponding to the trailing Hankel singular values are consistent. We construct a system $\mathcal{G}$ of order $n = 4$ in its canonical balanced form (3.2). Start by specifying the Hankel singular values

$$\Sigma = \mathrm{diag}(10^1, 10^0, 10^{-1}, 10^{-2})$$

and the corresponding sign parameters

$$S = \mathrm{diag}(1, 1, -1, -1).$$

By specifying the entries γ_i of the input vector b and applying the formula (3.2), we construct a balanced realization of a SISO LTI system $\mathcal{G}$ that is guaranteed to be minimal and asymptotically stable, having the canonical form

$$A = \left[\begin{array}{cc|cc} -0.05 & -0.18 & 0.30 & 0.40 \\ -0.18 & -2.00 & 6.67 & 8.08 \\ \hline -0.30 & -6.67 & -45.00 & -109.09 \\ -0.40 & -8.08 & -109.09 & -800.00 \end{array}\right], \quad b = \left[\begin{array}{c} 1 \\ 2 \\ \hline 3 \\ 4 \end{array}\right], \quad c^\top = \left[\begin{array}{c} 1 \\ 2 \\ \hline -3 \\ -4 \end{array}\right], \quad d = 0.$$

We compute reduced order models via balanced truncation of orders $r = 1, 2, 3$. We highlight the partitioning of $\mathcal{G}$ with respect to the truncation order $r = 2$ to expose the sign symmetry of the truncated system. Table 3.1 compares the $\mathcal{H}_\infty$-norm of the error system to the balanced truncation upper bound (2.15). When performing reduction to orders $r = 2$ and $r = 3$, the condition (3.8) is met, and the balanced truncation bound holds with equality, as guaranteed by Theorem 3.6. However when performing reduction to order $r = 1$, the truncated system does not obey the required sign consistency (3.8), and the upper bound (2.15) holds with a *strict inequality*.

Table 3.1: $\mathcal{H}_\infty$ norm of the error system, compared to the balanced truncation upper bound (2.15) for a system where Theorem 3.6 holds for $r = 2$ and 3, but not for $r = 1$.

	$\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty}$	$2(\sigma_{r+1} + \cdots + \sigma_n)$
$r = 1$	1.780×10^0	2.220×10^0
$r = 2$	2.200×10^{-1}	2.200×10^{-1}
$r = 3$	2.000×10^{-2}	2.000×10^{-2}

3.2.2 Extension to Singular Perturbation Balancing

We now show that the result of Theorem 3.6 holds when performing a variant of balanced truncation model reduction. As opposed to truncating the state $x_2(t)$ corresponding to Σ_2 in the balanced form (2.11), one can perform model reduction via *singular perturbation* *balancing* [20] by setting $x_2'(t) = 0$. Starting with the balanced realization of $\mathcal{G}$ in (2.11), the order-r *singular perturbation balancing approximation* of $\mathcal{G}$ is

$$\mathcal{G}_r^{sp} : \begin{cases} x_r'(t) = A_r^{sp} x_r(t) + b_r^{sp} u(t) \\ y_r(t) = c_r^{sp} x_r(t) + d_r^{sp} u(t), \end{cases} \tag{3.12}$$

where

$$A_r^{sp} = A_{11} - A_{21} A_{22}^{-1} A_{12}, \quad b_r^{sp} = b_1 - A_{12} A_{22}^{-1} b_2 \tag{3.13}$$

$$c_r^{sp} = c_1 - c_2 A_{22}^{-1} A_{21}, \qquad d_r^{sp} = d - c_2 A_{22}^{-1} b_2, \tag{3.14}$$

having the transfer function

$$G_r^{sp}(s) = c_r^{sp}(sI - A_r^{sp})^{-1} b_r^{sp} + d_r^{sp}.$$

In addition to the reduced model $\mathcal{G}_r^{sp}$ satisfying the same $\mathcal{H}_\infty$ error bound (2.15) (see [20, Thm. 3.2]), $\mathcal{G}_r^{sp}$ has the property that it matches the DC gain of the full-order system; namely $G(0) = G_r^{sp}(0)$. This interpolation is of importance in certain applications of reduced-order modeling, where the physical properties of the system guarantee a certain value of the transfer function at the frequency $s = 0$.

As a consequence of Theorem 3.6, we will show that the $\mathcal{H}_\infty$ error bound (2.15) also holds with equality when performing singular perturbation balancing, provided the sign parameters of the system $\mathcal{G}$ satisfy (3.8). This is also pointed out in [17, Section V.A], also using the concept of reciprocal system as in the proof below, but in a slightly different manner.

Theorem 3.8. *Let $\mathcal{G}$ be an order-n asymptotically stable, minimal, balanced SISO system as in (1.3) with its Hankel singular values $\Sigma = diag(\Sigma_1, \Sigma_2)$ and sign parameters $S =$*

diag(S_1, S_2) *partitioned as in Theorem 3.6. Let* $\mathcal{G}_r^{sp}$ *be the singular perturbation approximation of* $\mathcal{G}$ *given by* (3.12)*, truncated after the* k*th distinct Hankel singular value and having order* $r := m_1 + \cdots + m_k$*. If all the signs in* S_2 *are the same, as in* (3.8)*, then* $\mathcal{G}_r^{sp}$ *in* (3.12) *achieves the error bound* (2.15)*:*

$$\left\|\mathcal{G} - \mathcal{G}_r^{sp}\right\|_{\mathcal{H}_\infty} = 2(\sigma_{k+1} + \cdots + \sigma_q).$$

Proof. Without loss of generality assume that $\mathcal{G}$ is balanced and given as in the canonical form (3.1). It is shown in [20, Thm. 3.2] that the model reduction error from the rth order singular perturbation approximation to $\mathcal{G}$ can be written as

$$\left\|\mathcal{G} - \mathcal{G}_r^{sp}\right\|_{\mathcal{H}_\infty} = \left\|\widehat{\mathcal{G}} - \widehat{\mathcal{G}}_r\right\|_{\mathcal{H}_\infty},$$

where $\widehat{\mathcal{G}}$ is the *reciprocal system* of $\mathcal{G}$ given by the realization

$$\widehat{A} = A^{-1}, \quad \widehat{b} = A^{-1}b, \quad \widehat{c} = cA^{-1},$$

and $\widehat{\mathcal{G}}_r$ is the rth order balanced truncation reduced model for $\widehat{\mathcal{G}}$. Then [20, Lemma 3.1] states that the given realization of $\widehat{\mathcal{G}}$ is balanced with the Gramian Σ, and so the Hankel singular values of $\widehat{\mathcal{G}}$ are the same as those of the original system $\mathcal{G}$. Notice that the reciprocal system obeys the same sign symmetry as the original system, that is, $\widehat{A} = S\widehat{A}^\top S$ and $\widehat{b} = (\widehat{c}S)^\top$, where S is the sign matrix of $\mathcal{G}$: inverting both sides of $A = SA^\top S$ shows that $\widehat{A} = S\widehat{A}^\top S$; additionally we see

$$\widehat{b} = A^{-1}b = A^{-1}(cS)^\top = A^{-1}Sc^\top = SA^{-\top}c^\top = S\widehat{c}^\top.$$

It follows that the submatrices of the reciprocal system partitioned according to (2.11) satisfy the same sign symmetries as in (3.4) and (3.5). Thus applying the result of Theorem 3.6 to $\widehat{\mathcal{G}}$ and $\widehat{\mathcal{G}}_r$, we conclude

$$\left\|\mathcal{G} - \mathcal{G}_r^{sp}\right\|_{\mathcal{H}_\infty} = \left\|\widehat{\mathcal{G}} - \widehat{\mathcal{G}}_r\right\|_{\mathcal{H}_\infty} = 2(\sigma_{k+1} + \cdots + \sigma_q),$$

completing the proof. □

Chapter 4

Determining the Sign Parameters of Arrowhead Systems

Theorem 3.6 shows that the balanced truncation $\mathcal{H}_\infty$ error bound holds with equality for a family of systems that exhibit a generalized type of state-space symmetry based on the sign parameters corresponding to its Hankel singular values. In this chapter, we show how to determine these sign parameters, and thus the associated sign symmetry of the canonical balanced form, from the *arrowhead form* of a SISO system, when available. We first introduce the necessary preliminaries to show this fact, including a brief review of eigenvalue perturbation theory. We then motivate the result by providing a special example of an arrowhead system arising naturally in power systems modeling.

4.1 Arrowhead Realizations and Preliminary Tools

In general, we say a SISO LTI system $\mathcal{G}$ as in (1.3) has an *arrowhead realization* if it has a realization satisfying

$$A = \begin{bmatrix} d_1 & \alpha \\ \beta & D \end{bmatrix}, \quad b = \gamma e_1, \quad \text{and} \quad c = e_1^\top, \tag{4.1}$$

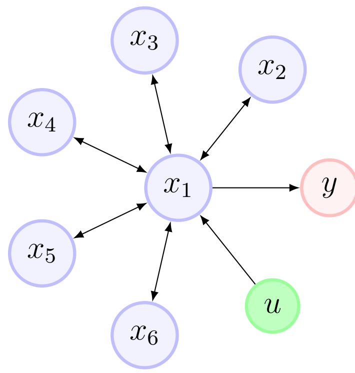

Figure 4.1: An arrowhead network with $n = 6$, input u and output y restricted to state x_1.

where $\alpha = [\alpha_2 \ \cdots \ \alpha_n] \in \mathbb{R}^{1\times(n-1)}$, $\beta = [\beta_2 \ \cdots \ \beta_n]^\top \in \mathbb{R}^{(n-1)\times 1}$, d_1, $\gamma \in \mathbb{R}$, and $D = \operatorname{diag}(d_2, \ldots, d_n) \in \mathbb{R}^{(n-1)\times(n-1)}$. Here $e_1 = [1 \ 0 \ \cdots \ 0]^\top \in \mathbb{R}^n$ is the first standard basis vector. Such an A is called an *arrowhead matrix*. Occasionally, we refer to systems having arrowhead realizations as *arrowhead systems*.

The arrowhead realization of a system describes a natural physical structure if we consider A to be the adjacency matrix of a weighted graph. Arrowhead realizations arise in systems with a physically meaningful coordinate system in which the evolution of the internal state variables are coupled as in Figure 4.1, with the input $u(t)$ and output $y(t)$ directly involving only the first state variable, $x_1(t)$. In Section 4.3, we provide an example of one such system used in modeling the aggregate response of a power network.

We first provide a formula for the inverse of an arrowhead matrix, as it will serve as a useful theoretical tool throughout the section. If $A \in \mathbb{R}^{n\times n}$ has the arrowhead form (4.1) with $d_i \neq 0$ for all $i = 2, \ldots, n$, and $d_1 - \alpha D^{-1}\beta \neq 0$, then the inverse of A exists and can be expressed as a diagonal matrix plus a rank-one update [21]:

$$A^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & D^{-1} \end{bmatrix} + \rho \begin{bmatrix} -1 \\ D^{-1}\beta \end{bmatrix} \begin{bmatrix} -1 & \alpha D^{-1} \end{bmatrix}, \tag{4.2}$$

where

$$\rho = \frac{1}{d_1 - \alpha D^{-1}\beta}.$$

(This formula can be derived via the Sherman–Morrison–Woodbury formula [22, p. 65]; though most easily derived when $d_1 \neq 0$, the formula holds even when $d_1 = 0$.) This formula also holds when the nonzero entries of A are complex-valued, but for simplicity of the presentation we state it for the real case. Using (4.2) we see that the transfer function $G(s) = c(sI - A)^{-1}b$ of a system with an arrowhead realization (4.1) will always have the form

$$G(s) = \frac{\gamma}{s - d_1 - \sum_{i=2}^{n} \frac{\alpha_i \beta_i}{s - d_i}}. \tag{4.3}$$

Any system having a transfer function $G(s)$ with the general form (4.3) has an arrowhead realization given by (4.1). Conversely, any system having an arrowhead realization (4.1) can have its transfer function $G(s)$ expressed in the form (4.3). Deriving rigorous conditions under which an LTI system $\mathcal{G}$ is guaranteed to have an arrowhead realization (4.1) is a topic of future work. Note that the poles of this system (and equivalently the eigenvalues of the arrowhead matrix A) are given by the zeros of the secular equation $h(s) = s - d_1 - \sum_{i=2}^{n} \frac{\alpha_i \beta_i}{s - d_i}$. Additionally, multiplying through by $\prod_{i=2}^{n}(s - d_i)$ in the numerator and denominator of $G(s)$ reveals that the zeros of $G(s)$ occur at $s = d_i$ for $i = 2, \ldots, n$.

We next prove a lemma regarding the minimality of arrowhead systems based on the entries of the corresponding arrowhead matrix. This will be necessary in the proof of Theorem 4.3.

Lemma 4.1. *Let $\mathcal{G}$ be an order-n asymptotically stable SISO system as in (1.3) with an arrowhead realization (4.1). Then $\mathcal{G}$ is minimal if and only if α_i and β_i are nonzero for all $i = 1 \ldots, n$ and $d_i \neq d_j$ for all $i \neq j$.*

Proof. Without loss of generality take $\gamma = 1$. We prove the forward implication by contradiction. Let $\mathcal{G}$ be minimal and suppose that $\alpha_i = 0$ for some $i = 2, \ldots, n$. From [9, Theorem 4.26], a system $\mathcal{G}$ is observable if and only if no right eigenvector of A is in the right kernel of $c = e_1^\top$. Because $\alpha_i = 0$ for some $i = 2, \ldots, n$, then $e_i = [0, \ldots, 1, \ldots, 0]^\top \in \mathbb{R}^n$ is an eigenvector of A corresponding to the eigenvalue d_i. But $e_1^\top e_i = 0$, which implies $\mathcal{G}$ is not observable, thus contradicting the the minimality assumption. So it must be that $\alpha_i \neq 0$ for all $i = 2, \ldots, n$. The same argument can be applied if $\beta_i = 0$ for some $i = 2, \ldots, n$, using the equivalent conditions for reachability given in [9, Theorem 4.16]. Next, suppose that $d_i = d_j$ for some $i \neq j$. Using the representation of $\mathcal{G}$'s transfer function (4.3), the fact that $d_i = d_j$ implies that the term $\sum_{i=2}^{n} \frac{\alpha_i \beta_i}{s - d_i}$ collapses down to a sum with at most $n - 1$ terms. So, $G(s)$ can be represented by an order $n - 1$ rational function, implying that the given realization is not minimal.

For the reverse implication, we again consider the representation of $\mathcal{G}$'s transfer function given in (4.3). Assuming that $\alpha_i, \beta_i \neq 0$ for all $i = 1, \ldots, n$ and $d_i \neq d_j$ for all $i \neq j$, then the term $\sum_{i=2}^{n} \frac{\alpha_i \beta_i}{s - d_i}$ is minimally represented by an order n strictly proper rational function, $P(s)/Q(s)$. In other words, $\sum_{i=2}^{n} \frac{\alpha_i \beta_i}{s - d_i} = P(s)/Q(s)$, where $P(s)$ and $Q(s)$ have no common factors. If this were not the case, we could express $P(s)/Q(s)$ as a partial fraction expansion with less than n terms, implying either that α_i or β_i are zero for some i, or $d_i = d_j$ for some $i \neq j$. So

$$G(s) = \frac{1}{s - d_1 - P(s)/Q(s)} = \frac{Q(s)}{(s - d_1)Q(s) - P(s)}.$$

Note that the zeros of $Q(s)$ (and thus $G(s)$) are given by d_i for $i = 2, \ldots, n$. Because $P(s)$ and $Q(s)$ have no common factors, $P(s) \neq 0$ for $s = d_i$, $i = 2, \ldots, n$. So $(s - d_1)Q(s) - P(s) \neq 0$ at these points as well. Defining $D(s) = (s - d_1)Q(s) - P(s)$, $G(s)$ can be written as $G(s) = Q(s)/D(s)$ where $Q(s)$ and $D(s)$ have no common factors. So $G(s)$ is minimally represented by an order n rational function, and so the arrowhead realization is a minimal one under these assumptions. □

4.1.1 The Cross Gramian

Similar to the reachability and observability Gramians of an LTI system $\mathcal{G}$, we can define a third related Gramian known as the cross Gramian. The cross Gramian is well-defined for any multi-input, multi-output system with equal input and output dimensions (for further details, see [9, Sec. 4.3.2]). We define the *cross Gramian* of an asymptotically stable LTI system $\mathcal{G}$ as the unique solution $\mathcal{X}$ to the Sylvester equation

$$A\mathcal{X} + \mathcal{X}A + bc = 0. \tag{4.4}$$

Given the assumption that $\mathcal{G}$ is asymptotically stable, $\mathcal{X}$ can be written in the frequency domain as

$$\mathcal{X} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(A - i\omega I)^{-1}bc\,(A + i\omega I)^{-1}\,d\omega, \tag{4.5}$$

(see, e.g., [9, sect. 6.1.2]). For SISO systems, the cross Gramian relates to the reachability and observability Gramians by the equality $\mathcal{X}^2 = \mathcal{P}\mathcal{Q}$. For our purposes, the cross Gramian is significant due to the following fact, which can be obtained by combining results from [16] with [9, Lemma 5.6].

Theorem 4.2. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal LTI SISO system as in (1.3). Let $\lambda_i(\mathcal{X})$, $i = 1, \ldots, n$, denote the eigenvalues of $\mathcal{X}$ (counted with multiplicity) such that*

$$|\lambda_1(\mathcal{X})| \geq |\lambda_2(\mathcal{X})| \geq \cdots \geq |\lambda_n(\mathcal{X})|.$$

Then, the eigenvalues of $\mathcal{X}$ are real, and the sign parameters s_i of $\mathcal{G}$ are given by

$$s_i = sign(\lambda_i(\mathcal{X})), \quad i = 1, \ldots, n.$$

The proof of Theorem 4.2 is outlined as follows. First one shows that the eigenvalues of the cross Gramian $\mathcal{X}$ are the same as the nonzero eigenvalues of a Hankel operator naturally associated with the system $\mathcal{G}$ [9, Lemma 5.6]. The nonzero singular values of this operator are the Hankel singular values of $\mathcal{G}$. Combining this with facts from [16] implies

$$\lambda_i(\mathcal{X}) = s_i \sigma_i, \quad i = 1, \ldots, n, \tag{4.6}$$

where $\{\sigma_i\}_{i=1}^n$ are the Hankel singular values of $\mathcal{G}$. Equation (4.6) implies that the Hankel singular values of a system $\mathcal{G}$ are the unsigned eigenvalues of its cross Gramian [16, Section 2.7], so the given ordering is consistent with that of the system's Hankel singular values. In conjunction with Theorem 3.1, this result implies that one can deduce the sign symmetry of a system's canonical balanced realization from the signs of the eigenvalues of its cross Gramian, and thus whether $\mathcal{G}$ obeys the hypotheses of Theorem 3.6. Rather than using the cross Gramian directly, we will leverage Theorem 4.2 in order to gain access to the sign parameters $\{s_i\}$ of a $\mathcal{G}$ from its arrowhead realization. In particular, we show that given any real arrowhead realization of a stable and minimal system $\mathcal{G}$, there exists a permutation π of $\{1, 2, \ldots, n\}$ such that

$$s_{\pi_1} = \text{sign}(\gamma), \quad \text{and} \quad s_{\pi_i} = \text{sign}(\gamma \alpha_i \beta_i).$$

In other words, we can read off the sign parameters of a system from the signs of the off-diagonal entries of a system's arrowhead realization. The ordering of these sign parameters (derived from the decreasing magnitude of the Hankel singular values they correspond to) varies. This follows from the fact that the arrowhead realization of a system is *not unique*. Indeed given one such realization, others can be obtained via symmetric permutations of the state space. This is akin to Remark 3.2: if one were to observe a balanced realization of a system $\mathcal{G}$ satisfying the sign-symmetry condition as in (3.1), it would not be clear whether Hankel singular values were ordered non-increasingly in that basis. We show later that there does exist an arrowhead realization in which the states are ordered corresponding to the canonical ordering of $\mathcal{G}$'s sign parameters (i.e., where the permutation π is the identity).

We next review the basics of eigenvalue perturbation theory for linear operators. These techniques will be used in the proof of Theorem 4.3.

4.1.2 Results from Eigenvalue Perturbation Theory

We briefly review the necessary results of eigenvalue perturbation theory for linear operators. For more on this topic, see [23]. For a more general overview, see [24].

Consider a matrix $B(z) \in \mathbb{C}^{n\times n}$ that is an analytic function of the parameter $z \in \mathbb{C}$. Let $B_0 = B(z_0)$ for some $z_0 \in \mathbb{C}$, and suppose $B(z_0)$ has a *simple* eigenvalue $\lambda_0 \in \mathbb{C}$ corresponding to left and right eigenvectors $w_0, v_0 \in \mathbb{C}^n$, respectively. Because $B(z)$ is an analytic function of z, its entries depend continuously on z, and thus so do its eigenvalues. Further, one can define an eigenvalue $\lambda(z)$ of $B(z)$ that is analytically dependent upon z with $\lambda(0) = \lambda_0$.

In order to see how $\lambda(z)$ behaves under small perturbations, we derive a formula for its first derivative with respect to z. First implicitly differentiate the eigenvalue equation $B(z)v(z) = \lambda(z)v(z)$ to obtain

$$B'(z)v(z) + B(z)v'(z) = \lambda'(z)v(z) + \lambda(z)v'(z). \tag{4.7}$$

Multiply (4.7) from the left with $w(z)^\top$ and evaluate at $z = z_0$ to obtain

$$w_0^\top B'(z_0)v_0 + \underbrace{w_0^\top B_0}_{=\lambda_0 w_0^\top} v'(z_0) = \lambda'(z_0)w_0^\top v_0 + \lambda_0 w_0^\top v'(z_0).$$

Rearranging and cancelling appropriate terms yields

$$\lambda'(z_0) = \frac{w_0^\top B'(z_0)v_0}{w_0^\top v_0}. \tag{4.8}$$

By taking a power series expansion of $\lambda(z)$ about z_0 up to first-order, one can see how $\lambda(z)$ behaves under small perturbations in z.

If one needs higher-order terms in this expansion, a similar technique can be used to derive a formula for the second derivative of $\lambda(z)$ with respect to the parameter z. This time, begin by implicitly differentiating the eigenproblem $B(z)v(z) = \lambda(z)v(z)$ twice to obtain

$$B''(z)v(z) + 2B'(z)v'(z) + B(z)v''(z) = \lambda''(z)v(z) + 2\lambda'(z)v'(z) + \lambda(z)v''(z). \tag{4.9}$$

Again we multiply this from the left with $w(z)^\top$ and evaluate at $z = z_0$:

$$w_0^\top B''(z_0)v_0 + 2w_0^\top B'(z_0)v'(z_0) + \underbrace{w_0^\top B_0}_{=\lambda_0 w_0^\top} v''(z_0) = w_0^\top \lambda''(z_0)v_0 + 2w_0^\top \lambda'(z_0)v'(z_0) + w_0^\top \lambda_0 v''(z_0).$$

Rearranging and appropriate cancellation yield

$$\lambda''(z_0) = \frac{w_0^\top B''(z_0)v_0 + 2w_0^\top B'(z_0)v'(z_0) - 2\lambda'(z_0)w_0^\top v'(z_0)}{w_0^\top v_0}. \tag{4.10}$$

These formulae will be necessary to prove the main result of Chapter 4.

4.2 Main Results for Arrowhead Systems

Before proving the main result of this section, we establish the definition of a *subsystem* of an order-n LTI system $\mathcal{G}$. For $k < n$, we define the *order-k* subsystem of $\mathcal{G}$, denoted $\mathcal{G}_k$, to be the system having the realization $A_k \in \mathbb{R}^{k\times k}$, $b_k \in \mathbb{R}^{k\times 1}$, and $c_k \in \mathbb{R}^{1\times k}$, where A_k, b_k, and c_k are the leading k-dimensional submatrices of A, b, and c respectively. These subsystems can also be obtained directly by truncating the trailing $n-k$ entries of the state vector $x(t) \in \mathbb{R}^n$ for $k < n$. If the realization of $\mathcal{G}$ given in (1.3) is an arrowhead realization, then the order-k subsystem of $\mathcal{G}_k$ of $\mathcal{G}$ can be interpreted as a system having the network structure of $\mathcal{G}$ shown in Figure 4.1 with the nodes corresponding to the trailing $n-k$ states removed. We are now prepared to state and prove the main result of this section.

Theorem 4.3. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal SISO system as in (1.3) with an arrowhead realization (4.1) such that $d_i < 0$ for all $i = 1, \ldots, n$, and such that every subsystem $\mathcal{G}_k$, for $k < n$, of $\mathcal{G}$ is asymptotically stable. Then $\alpha_i, \beta_i \neq 0$ for all $i, j = 1, 2, \ldots, n$, $d_i \neq d_j$ for $i \neq j$, and there exists a permutation $\pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$ of $\{1, 2, \ldots, n\}$ such that the sign parameters of $\mathcal{G}$ are given by*

$$s_{\pi_1} = sign\{\gamma\}, \quad s_{\pi_i} = sign\{\gamma\alpha_i\beta_i\}, \quad i = 2, \ldots, n. \tag{4.11}$$

Proof. The claims that α_i, $\beta_i \neq 0$ for all $i = 1, \ldots, n$ and $d_i \neq d_j$ for $i \neq j$ follow from Lemma 4.1. Without loss of generality, take $\gamma = 1$. (For a general $\gamma \in \mathbb{R}$, the result follows by carrying the constant through the subsequent calculations wherever b is involved.) Recall A is an arrowhead matrix, so for each $\omega \in \mathbb{R}$, $A \pm i\omega I$ is an arrowhead matrix as well.

Applying the formula for the inverse of an arrowhead matrix (4.2) to $A \pm i\omega I$ gives

$$(A \pm i\omega I)^{-1} = \begin{bmatrix} 0 & \\ & (D \pm i\omega I)^{-1} \end{bmatrix} + \rho(i\omega) \begin{bmatrix} -1 \\ (D \pm i\omega I)^{-1}\beta \end{bmatrix} \begin{bmatrix} -1 & \alpha(D \pm i\omega I)^{-1} \end{bmatrix},$$

and so, in particular

$$(A - i\omega I)^{-1} e_1 = \rho(-i\omega) \begin{bmatrix} 1 \\ -(D - i\omega I)^{-1}\beta \end{bmatrix},$$

and

$$e_1^\top (A + i\omega I)^{-1} = \rho(i\omega) \begin{bmatrix} 1 & -\alpha(D + i\omega I)^{-1} \end{bmatrix},$$

where

$$\rho(\pm i\omega) = \frac{1}{(d_1 \pm i\omega) - \alpha(D \pm i\omega I)^{-1}\beta}. \tag{4.12}$$

Note that assuming A is asymptotically stable requires all of its eigenvalues lie in the left half-plane. Because the eigenvalues of A are given by the poles of $\rho(-z)$, $z \in \mathbb{C}$, this asymptotic stability assumption implies $\rho(\pm i\omega)$ is well defined for all $\omega \in \mathbb{R}$, implying $(A \pm i\omega I)^{-1}$ is well-defined for all $\omega \in \mathbb{R}$ as well. Letting $F(\pm i\omega) = (D \pm i\omega I)^{-1}$, and using the fact that $b = c^\top = e_1$, plugging these formulae into the integral representation of the cross Gramian (4.5) gives a useful representation of $\mathcal{X}$ for the associated arrowhead system:

$$\mathcal{X} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\rho(i\omega)|^2 \begin{bmatrix} 1 & -\alpha F(i\omega) \\ -F(-i\omega)\beta & F(-i\omega)\beta\alpha F(i\omega) \end{bmatrix} d\omega. \tag{4.13}$$

Note that $\rho(z)$ is a rational function with all real coefficients, and so $\rho(i\omega)\rho(-i\omega) = |\rho(i\omega)|^2$.

In order to ultimately say something about the eigenvalues of $\mathcal{X}$, and thus the sign parameters of $\mathcal{G}$, we invoke a homotopy in the arrowhead realization of $\mathcal{G}$. Define

$$A(\tau) = A_0 + \tau\beta_n e_n e_1^\top + \tau\alpha_n e_1 e_n^\top, \quad A_0 = \begin{bmatrix} d_1 & \alpha_2 & \cdots & \alpha_{n-1} & 0 \\ \beta_2 & d_2 & & & \\ \vdots & & \ddots & & \\ \beta_{n-1} & & & \ddots & \\ 0 & & & & d_n \end{bmatrix}, \tag{4.14}$$

for $\tau \in [0, 1]$. (We have simply added a linear dependency on the variable τ in the bottom left and upper right corner entries of A). For each $\tau \in (0, 1]$, $A(\tau)$ corresponds to an order-n minimal system $\mathcal{G}(\tau)$ having an arrowhead realization (4.1) with A replaced by $A(\tau)$. Note that $\mathcal{G}(1) = \mathcal{G}$ returns the original system, and $\mathcal{G}(0)$ corresponds to a *non*-minimal, asymptotically stable system. This non-minimality claim follows from Lemma 4.1 given the fact that the bottom-left and upper-right corner entries of $A(0)$ are 0. The asymptotic stability of $\mathcal{G}(0)$ follows from the assumptions that d_n is negative, and that the $n-1$ dimensional subsystem $\mathcal{G}_{n-1}$ of $\mathcal{G}$ is asymptotically stable. $A(0)$ has $n-1$ eigenvalues given by those of its principal $n-1$ submatrix (that we denote A_{n-1}) and one eigenvalue given by d_n. Because A_{n-1} is an arrowhead realization of $\mathcal{G}_{n-1}$, it has all of its eigenvalues in the left half-plane. This, in addition to the assumption that $d_n < 0$, implies that $A(0)$ has eigenvalues in the left half-plane, and so $\mathcal{G}(0)$ is asymptotically stable. The fact that $\mathcal{G}(\tau)$ is minimal for all $\tau \in (0, 1]$ follow from Lemma 4.1. Because $\mathcal{G}(1)$ and $\mathcal{G}(0)$ are asymptotically stable, without loss of generality we may assume that $\mathcal{G}(\tau)$ is asymptotically stable for all $\tau \in (0, 1)$ as well. If this were not the case, one can define an alternative family of minimal and asymptotically stable arrowhead systems $\widetilde{\mathcal{G}}(\tau)$ that agrees with $\mathcal{G}(\tau)$ for values of τ close to 0 and 1, and moves continuously from $\widetilde{\mathcal{G}}(0) = \mathcal{G}(0)$ to $\widetilde{\mathcal{G}}(1) = \mathcal{G}(1)$. For ease of presentation, we state and prove this formally in Lemma 4.4 following this proof of Theorem 4.3.

Because $\mathcal{G}(\tau)$ is asymptotically stable for all $\tau \in [0,1]$, $A(\tau)$ and $-A(\tau)$ will always have disjoint spectra, so it follows that the cross Gramian $\mathcal{X}(\tau)$ of $\mathcal{G}(\tau)$ is the unique solution to the Sylvester equation

$$A(\tau)\mathcal{X}(\tau) + \mathcal{X}(\tau)A(\tau) + e_1 e_1^\top = 0. \tag{4.15}$$

For sufficiently small $\tau > 0$, by making the corresponding change of variables from $\alpha_n \mapsto \tau\alpha_n$ and $\beta_n \mapsto \tau\beta_n$ in (4.13), $\mathcal{X}(\tau)$ is given by the integral formula

$$\mathcal{X}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,\tau)|^2\, W(\tau)\, d\omega, \tag{4.16}$$

where

$$W(\tau) = \begin{bmatrix} 1 & -\frac{\alpha_2}{d_2+i\omega} & \cdots & -\frac{\alpha_{n-1}}{d_{n-1}+i\omega} & -\tau\frac{\alpha_n}{d_n+i\omega} \\ -\frac{\beta_2}{d_2-i\omega} & \frac{\alpha_2\beta_2}{|d_2+i\omega|^2} & \cdots & \frac{\alpha_{n-1}\beta_2}{(d_2-i\omega)(d_{n-1}+i\omega)} & \tau\frac{\alpha_n\beta_2}{(d_2-i\omega)(d_n+i\omega)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{\beta_{n-1}}{d_{n-1}-i\omega} & \frac{\alpha_2\beta_{n-1}}{(d_{n-1}-i\omega)(d_2+i\omega)} & \cdots & \frac{\alpha_{n-1}\beta_{n-1}}{|d_{n-1}+i\omega|^2} & \tau\frac{\alpha_n\beta_{n-1}}{(d_{n-1}-i\omega)(d_n+i\omega)} \\ -\tau\frac{\beta_n}{d_n-i\omega} & \tau\frac{\alpha_2\beta_n}{(d_2-i\omega)(d_n+i\omega)} & \cdots & \tau\frac{\alpha_{n-1}\beta_n}{(d_n-i\omega)(d_{n-1}+i\omega)} & \tau^2\frac{\alpha_n\beta_n}{|d_n+i\omega|^2} \end{bmatrix},$$

and

$$\rho(\pm i\omega,\tau) = \frac{1}{(d_1 \pm i\omega) - \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i\pm i\omega} - \tau^2\frac{\alpha_n\beta_n}{d_n\pm i\omega}}.$$

(Note that we can only invoke this integral representation for small values of $\tau > 0$ due to the previous claim that $\mathcal{G}(\tau)$ is asymptotically stable for all $\tau \in (0,1)$. This reduction only guarantees that $\mathcal{G}(\tau)$ has the particular arrowhead realization given in (4.1) for small values of $\tau > 0$, and thus $\mathcal{X}(\tau)$ has the form (4.13) for small values of $\tau > 0$ as well. This is acceptable, because in the arguments that follow we only use this representation for values of τ close to zero.) We write out $W(\tau)$ entry-by-entry in order to appreciate fully how the homotopy in $A(\tau)$ affects the structure of $\mathcal{X}(\tau)$. (In particular, note the resulting quadratic dependence on τ in the bottom-right entry of $\mathcal{X}(\tau)$, and the nonlinear dependence on τ in $|\rho(\pm i\omega, \tau)|^2$.) Let the eigenvalues and corresponding left and right eigenvectors of $\mathcal{X}(\tau)$ be denoted by

$$\lambda_i(\tau) \in \mathbb{R}, \quad w_i(\tau),\, v_i(\tau) \in \mathbb{R}^{n\times 1}, \quad i = 1, 2, \ldots, n,$$

respectively. The formula (4.16) implies that for any system order n, $\mathcal{X}(0)$ is a rank $n-1$ matrix with a simple eigenvalue $\lambda_n(0) = 0$ corresponding to left and right eigenvectors $v_n(0) = w_n(0) = e_n$.

We will argue how these eigenvalues and their signs change as τ goes from 0 to 1, to ultimately say something about the sign parameters of $\mathcal{G}$. In order to do this, we first must show that the eigenvalues of $\mathcal{X}(\tau)$ change continuously with τ. By vectorizing $\mathcal{X}(\tau)$, the parameter dependent Sylvester equation (4.15) can be written in the form of a parameter-dependent linear equation, namely

$$\big(I \otimes A(\tau) + A(\tau) \otimes I\big)\, \mathrm{vec}(\mathcal{X}(\tau)) = -\mathrm{vec}(e_1 e_1^\top).$$

Solutions to nonsingular parameter dependent linear equations with continuously differentiable coefficient matrices are continuously differentiable themselves, and hence $\mathcal{X}(\tau)$ inherits this property [25, Sec. 7.8]. The entries of $\mathcal{X}(\tau)$ must then be continuous with respect to τ as well, implying that the eigenvalues of $\mathcal{X}(\tau)$ change continuously with τ [23, Ch. 2].

The rest of the proof is organized as follows: We proceed inductively on the system order n, so as to deal with a single eigenvalue of $\mathcal{X}$ (and thus a single sign parameter of $\mathcal{G}$) at a time. Using techniques from eigenvalue perturbation theory, we leverage the structure of the homotopy to show that $\lambda_n(\tau)$ moves from 0 in the direction of $\mathrm{sign}(\alpha_n\beta_n)$ in response to a small enough perturbation in τ. We then argue by contradiction that once $\lambda_n(\tau)$ is nonzero, it must stay nonzero for all $\tau \in (0,1]$, and in particular for $\tau = 1$. Taking $\tau \to 1$, in combination with the induction hypothesis provides the desired formula for the sign parameters of $\mathcal{G}$, as in (4.11).

We now proceed by induction. The base case of $n = 2$ is similar in spirit to the induction step (without requiring the induction hypothesis) so we handle the induction step first. Suppose the result holds for all systems of order $n-1$ satisfying the hypotheses of the theorem. First note that for $\tau = 0$, the cross Gramian $\mathcal{X}(\tau)$ of $\mathcal{G}(\tau)$ has the structure

$$\mathcal{X}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} 1 & -\frac{\alpha_2}{d_2+i\omega} & \cdots & -\frac{\alpha_{n-1}}{d_{n-1}+i\omega} & 0 \\ -\frac{\beta_2}{d_2-i\omega} & \frac{\alpha_2\beta_2}{|d_2+i\omega|^2} & \cdots & \frac{\alpha_{n-1}\beta_2}{(d_2-i\omega)(d_{n-1}+i\omega)} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{\beta_{n-1}}{d_{n-1}-i\omega} & \frac{\alpha_2\beta_{n-1}}{(d_{n-1}-i\omega)(d_2+i\omega)} & \cdots & \frac{\alpha_{n-1}\beta_{n-1}}{|d_{n-1}+i\omega|^2} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} d\omega$$

$$=: \begin{bmatrix} \mathcal{X}_{n-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

Using (4.13) it is easily verifiable that $\mathcal{X}_{n-1}$ is the integral representation of the cross Gramian of $\mathcal{G}$'s $n-1$ dimensional principal subsystem $\mathcal{G}_{n-1}$. This follows from the fact that $\mathcal{G}_{n-1}$ is an asymptotically stable system by hypothesis, having an arrowhead realization given by A_{n-1} (the upper-left $n-1$ dimensional block of $A(1)$ as in (4.14)) and $b = c^\top = e_1 \in \mathbb{R}^{n-1}$. As noted before, $\mathcal{X}(0)$ has a simple eigenvalue $\lambda_n(0) = 0$ corresponding to the left and right eigenvectors $w_n(0) = v_n(0) = e_n \in \mathbb{R}^n$. Further, the fact that $\alpha_i, \beta_i \neq 0$ for all $i = 2, \ldots, n-1$ and $d_i \neq d_j$ for all $i \neq j$ implies that $\mathcal{G}_{n-1}$ is minimal, and thus satisfies the induction hypotheses. This implies that the $n-1$ nonzero eigenvalues of $\mathcal{X}(0)$ (not necessarily ordered in terms of magnitude) are those of $\mathcal{X}_{n-1}$, and thus satisfy

$$\operatorname{sign}(\lambda_1(0)) = +1, \quad \operatorname{sign}(\lambda_i(0)) = \operatorname{sign}(\alpha_i\beta_i), \quad i = 2,\dots,n-1,$$

Next, we show how $\mathcal{X}(\tau)$'s eigenvalues behave for small $\tau > 0$ using techniques from eigenvalue perturbation theory. The aforementioned nonlinear dependence of $|\rho(i\omega,\tau)|^2$ on τ makes it difficult to deal with $\mathcal{X}(\tau)$ directly. We instead deal with the second order power series approximation of $\mathcal{X}(\tau)$ about $\tau = 0$, namely

$$M(\tau) = \mathcal{X}(0) + \mathcal{X}'(0)\tau + \mathcal{X}''(0)\frac{\tau^2}{2}. \tag{4.17}$$

Let the eigenvalues and corresponding left and right eigenvectors of $M(\tau)$ be denoted by $\widetilde{\lambda}_i(\tau) \in \mathbb{R}$ and $\widetilde{w}_i(\tau), \widetilde{v}_i(\tau) \in \mathbb{R}^{n\times 1}$ respectively, for $i = 1,2,\dots,n$. Because $\mathcal{X}(\tau) = M(\tau) + \mathcal{O}(\tau^3)$, the eigenvalues of $\mathcal{X}(\tau)$ behave like those of $M(\tau)$ for sufficiently small τ, as can be verified explicitly by taking power series expansions of both $\lambda_i(\tau)$ and $\widetilde{\lambda}_i(\tau)$, and comparing terms with like orders. In particular, the fact that $M(0) = \mathcal{X}(0)$, $M'(0) = \mathcal{X}'(0)$, and $M''(0) = \mathcal{X}''(0)$, along with formulas (4.8) and (4.10), imply that all terms up to and including $\mathcal{O}(\tau^2)$ in these expansions for $\lambda_i(\tau)$ and $\widetilde{\lambda}_i(\tau)$ match. Further, the fact that $M(0) = \mathcal{X}(0)$ implies that $\widetilde{\lambda}_n(0) = \lambda_n(0) = 0$, and so $\widetilde{w}_n(0) = \widetilde{v}_n(0) = e_n$. So, the second order power series expansion of $\widetilde{\lambda}_n(\tau)$ about $\tau = 0$ is given by

$$\begin{aligned}\widetilde{\lambda}_n(\tau) &= \widetilde{\lambda}_n(0) + \widetilde{\lambda}_n'(0)\tau + \widetilde{\lambda}_n''(0)\frac{\tau^2}{2} + \mathcal{O}(\tau^3)\\ &= \widetilde{\lambda}_n'(0)\tau + \widetilde{\lambda}_n''(0)\frac{\tau^2}{2} + \mathcal{O}(\tau^3).\end{aligned}$$

Noting that $M'(0) = \mathcal{X}'(0)$ and $M''(0) = \mathcal{X}''(0)$, applying formulas (4.8) and (4.10) to $M(\tau)$

and evaluating at $\tau = 0$ gives

$$\widetilde{\lambda}_n'(0) = e_n^\top \mathcal{X}'(0) e_n,$$

and

$$\widetilde{\lambda}_n''(0) = e_n^\top \mathcal{X}''(0) e_n + 2 e_n^\top \mathcal{X}'(0) \widetilde{v}_n'(0) - 2\widetilde{\lambda}_n'(0) e_n^\top \widetilde{v}_n'(0).$$

So we can say something about the sign $\widetilde{\lambda}_n(\tau)$ for small τ by investigating the terms in its power series expansion about $\tau = 0$, based on $\mathcal{X}'(0)$ and $\mathcal{X}''(0)$.

We claim that

$$\mathcal{X}'(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\rho(i\omega, 0)|^2 \, Z d\omega, \tag{4.18}$$

where

$$Z = \begin{bmatrix} 0 & 0 & \dots & 0 & \frac{-\alpha_n}{d_n + i\omega} \\ 0 & 0 & \dots & 0 & \frac{\alpha_n \beta_2}{(d_2 - i\omega)(d_n + i\omega)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \frac{\alpha_n \beta_{n-1}}{(d_{n-1} - i\omega)(d_n + i\omega)} \\ \frac{-\beta_n}{d_n - i\omega} & \frac{\alpha_2 \beta_n}{(d_n - i\omega)(d_2 + i\omega)} & \dots & \frac{\alpha_{n-1}\beta_n}{(d_n - i\omega)(d_{n-1} + i\omega)} & 0 \end{bmatrix}$$

and

$$e_n^\top \mathcal{X}''(0) e_n = \frac{1}{\pi} \int_{-\infty}^{\infty} |\rho(i\omega, 0)|^2 \frac{\alpha_n \beta_n}{|d_n + i\omega|^2} \, d\omega. \tag{4.19}$$

To derive the formula (4.18) for $\mathcal{X}'(0)$, observe that implicitly differentiating the Sylvester equation $A(\tau)\mathcal{X}(\tau) + \mathcal{X}(\tau)A(\tau) + e_1 e_1^\top = 0$ with respect to τ and evaluating at $\tau = 0$ gives

$$A'(0)\mathcal{X}(0) + \mathcal{X}'(0)A(0) + A(0)\mathcal{X}'(0) + \mathcal{X}(0)A'(0) = 0, \tag{4.20}$$

which amounts to another Sylvester equation for the unknown $\mathcal{X}'(0)$:

$$\mathcal{X}'(0)A(0) + A(0)\mathcal{X}'(0) = -\mathcal{X}(0)A'(0) - A'(0)\mathcal{X}(0). \tag{4.21}$$

Since $\mathcal{G}(0)$ is asymptotically stable, the coefficient matrices $A(0)$ and $-A(0)$ have disjoint spectra, and so the solution $\mathcal{X}'(0)$ to this equation is guaranteed to be unique. Thus it suffices to show that the formula for $\mathcal{X}'(0)$ claimed in (4.18) satisfies the Sylvester equation (4.21). We will show this term-by-term, hitting (4.21) with $e_i^\top \in \mathbb{R}^{1\times n}$ and $e_j \in \mathbb{R}^{n\times 1}$ from the left and right, respectively, and showing that the terms sum up to 0. This is done by cases, based on the different parts of $\mathcal{X}'(0)$. For the specific calculations, see Appendix A.

To prove the formula (4.19) for $e_n^\top \mathcal{X}''(0)e_n$, observe that implicitly differentiating $A(\tau)\mathcal{X}(\tau)+\mathcal{X}(\tau)A(\tau) + e_1e_1^\top = 0$ twice with respect to τ and evaluating at $\tau = 0$ gives

$$A(0)\mathcal{X}''(0) + 2A'(0)\mathcal{X}'(0) + 2\mathcal{X}'(0)A'(0) + \mathcal{X}''(0)A(0) = 0, \tag{4.22}$$

which can be viewed as a Sylvester equation having $\mathcal{X}''(0)$ as its solution:

$$A(0)\mathcal{X}''(0) + \mathcal{X}''(0)A(0) = -2A'(0)\mathcal{X}'(0) - 2\mathcal{X}'(0)A'(0). \tag{4.23}$$

By the same logic as above, the solution $\mathcal{X}''(0)$ is unique. We can derive the formula (4.19) for $e_n^\top \mathcal{X}''(0)e_n$ from the (n, n) entry of the Sylvester equation (4.23). First, multiply (4.23) by $e_n^\top \in \mathbb{R}^{1\times n}$ and $e_n \in \mathbb{R}^{n\times 1}$ from the left and right, respectively. Term by term, the left-hand side of (4.23) becomes

$$e_n^\top A(0)\mathcal{X}''(0)e_n = d_n e_n^\top \mathcal{X}''(0)e_n,$$

$$e_n^\top \mathcal{X}''(0)A(0)e_n = d_n e_n^\top \mathcal{X}''(0)e_n.$$

Similarly, the right-hand side of (4.23) involves the two terms

$$\begin{aligned}-2e_n^\top A'(0)\mathcal{X}'(0)e_n =& 2\beta_n e_1^\top \mathcal{X}'(0)e_n \\ =& -2\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{-\alpha_n}{d_n+i\omega}\,d\omega \\ =& d_n\alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{|d_n+i\omega|^2}\,d\omega \\ & -\alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{|d_n+i\omega|^2}\,d\omega,\end{aligned}$$

$$\begin{aligned}-2e_n^\top \mathcal{X}'(0)A'(0)e_n =& 2\alpha_n e_n^\top \mathcal{X}'(0)e_1 \\ =& -2\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{-\beta_n}{d_n-i\omega}\,d\omega \\ =& d_n\alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{|d_n+i\omega|^2}\,d\omega \\ & +\alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{|d_n+i\omega|^2}\,d\omega.\end{aligned}$$

Using these expressions, the (n,n) entry of (4.23) becomes

$$2d_n e_n^\top \mathcal{X}''(0)e_n = 2d_n\alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{|d_n+i\omega|^2}\,d\omega,$$

and so

$$e_n^\top \mathcal{X}''(0)e_n = \alpha_n\beta_n \frac{1}{\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{|d_n+i\omega|^2}\,d\omega,$$

verifying formula (4.19).

Using the formulae (4.18) and (4.19), the formulae (4.8) and (4.10) for $\widetilde{\lambda}'_n(0)$ and $\widetilde{\lambda}''_n(0)$ reduce further to

$$\widetilde{\lambda}'_n(0) = 0 \quad \text{and} \quad \widetilde{\lambda}''_n(0) = e_n^\top \mathcal{X}''(0)e_n + 2e_n^\top \mathcal{X}'(0)\widetilde{v}'_n(0).$$

So, the power series expansion of $\widetilde{\lambda}_n(\tau)$ about 0 reduces further to

$$\widetilde{\lambda}_n(\tau) = \widetilde{\lambda}''_n(0)\frac{\tau^2}{2} + \mathcal{O}(\tau^3). \tag{4.24}$$

Recall that we want to prove that $\operatorname{sign}(\lambda_n(\tau)) = \operatorname{sign}(\alpha_n\beta_n)$ for small τ. To this end, we will show that $\operatorname{sign}\left(\widetilde{\lambda}_n(\tau)\right) = \operatorname{sign}(\alpha_n\beta_n)$ for small τ by showing that $\operatorname{sign}\left(\widetilde{\lambda}''_n(0)\right) = \operatorname{sign}(\alpha_n\beta_n)$, and applying this to (4.24) for a small enough $\tau > 0$. Noting that (4.19) is $\alpha_n\beta_n$ times $1/\pi > 0$ times the integral of a positive quantity, it follows that

$$\operatorname{sign}(e_n^\top \mathcal{X}''(0)e_n) = \operatorname{sign}(\alpha_n\beta_n).$$

Thus it will suffice to show that

$$\operatorname{sign}(2e_n^\top \mathcal{X}'(0)\widetilde{v}'_n(0)) = \operatorname{sign}(\alpha_n\beta_n),$$

as well. First note that by implicitly differentiating the eigenvalue equation $M(\tau)\widetilde{v}_n(\tau) = \widetilde{\lambda}_n(\tau)\widetilde{v}_n(\tau)$ with respect to τ and evaluating at $\tau = 0$ gives

$$(M'(0) - \lambda'_n(0)I_n)\widetilde{v}_n(0) = -(M(0) - \widetilde{\lambda}_n(0))\widetilde{v}'_n(0)$$

after some rearrangement. Recalling that $M'(0) = \mathcal{X}'(0)$, $M(0) = \mathcal{X}(0)$, $\widetilde{v}_n(0) = e_n$, and

$\widetilde{\lambda}_n(0) = \widetilde{\lambda}'_n(0) = 0$, we obtain the useful identity

$$\mathcal{X}'(0)e_n = -\mathcal{X}(0)\widetilde{v}'_n(0). \tag{4.25}$$

Let

$$\mathcal{X}'(0) = \begin{bmatrix} 0 & y \\ w & 0 \end{bmatrix}, \quad \text{and} \quad \widetilde{v}'_n(0) = \begin{bmatrix} z \\ \alpha \end{bmatrix},$$

where $y, w^\top, z \in \mathbb{R}^{(n-1)\times 1}$, $\alpha \in \mathbb{R}$, and y and w are defined as in (4.18). Expanding upon equation (4.25) and substituting in for $\mathcal{X}(0)$ gives

$$\begin{bmatrix} 0 & y \\ w & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = - \begin{bmatrix} \mathcal{X}_{n-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z \\ \alpha \end{bmatrix}.$$

Multiplying through in the first block reveals that

$$y = -\mathcal{X}_{n-1}z. \tag{4.26}$$

Because $\mathcal{X}_{n-1}$ is the Gramian of the minimal $n-1$ dimensional system $\mathcal{G}_{n-1}$, it is necessarily full rank [9], and thus invertible. Solving (4.26) for z gives $z = -\mathcal{X}_{n-1}^{-1}y$, and at this point $\widetilde{v}'_n(0)$ becomes

$$\widetilde{v}'_n(0) = \begin{bmatrix} -\mathcal{X}_{n-1}^{-1}y \\ \alpha \end{bmatrix},$$

where we recall that y is the vector containing the first $n-1$ entries of the nth column of $\mathcal{X}_{n-1}$. Expanding $e_n^\top \mathcal{X}'(0)\widetilde{v}'_n(0)$ using this form gives

$$e_n^\top \mathcal{X}'(0)\widetilde{v}'_n(0) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & y \\ w & 0 \end{bmatrix} \begin{bmatrix} -\mathcal{X}_{n-1}^{-1}y \\ \alpha \end{bmatrix} = -w\mathcal{X}_{n-1}^{-1}y.$$

By definition of w and y, we can write

$$\begin{aligned} w &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} \frac{-\beta_n}{d_n - i\omega} & \frac{\alpha_2\beta_n}{(d_2+i\omega)(d_n-i\omega)} & \cdots & \frac{\alpha_{n-1}\beta_n}{(d_{n-1}+i\omega)(d_n-i\omega)} \end{bmatrix} d\omega \\ &= \beta_n \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} \frac{-1}{d_n - i\omega} & \frac{1}{(d_2+i\omega)(d_n-i\omega)} & \cdots & \frac{1}{(d_{n-1}+i\omega)(d_n-i\omega)} \end{bmatrix} d\omega \right) \widehat{A} \\ &= \beta_n \widehat{w}\widehat{A}, \end{aligned}$$

where $\widehat{A} = \operatorname{diag}(1, \alpha_2, \ldots, \alpha_{n-1}) \in \mathbb{R}^{(n-1)\times(n-1)}$. Similarly

$$\begin{aligned} y &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} \frac{-\alpha_n}{d_n + i\omega} & \frac{\alpha_n\beta_2}{(d_n+i\omega)(d_2-i\omega)} & \cdots & \frac{\alpha_n\beta_{n-1}}{(d_n+i\omega)(d_{n-1}-i\omega)} \end{bmatrix}^\top d\omega \\ &= \alpha_n \widehat{B} \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} \frac{-1}{d_n + i\omega} & \frac{1}{(d_n+i\omega)(d_2-i\omega)} & \cdots & \frac{1}{(d_n+i\omega)(d_{n-1}-i\omega)} \end{bmatrix}^\top d\omega \right) \\ &= \alpha_n \widehat{B}\widehat{y}, \end{aligned}$$

where $\widehat{B} = \operatorname{diag}(1, \beta_2, \ldots, \beta_{n-1}) \in \mathbb{R}^{(n-1)\times(n-1)}$. Note that $-\widehat{w}^\top = \widehat{y}$ under the change of variable $\zeta = -\omega$. Making this substitution gives

$$e_n^\top \mathcal{X}'(0)\widetilde{v}_n'(0) = -w\mathcal{X}_{n-1}^{-1}y = \alpha_n\beta_n \left(\widehat{y}^\top \widehat{A}\mathcal{X}_{n-1}^{-1}\widehat{B}\widehat{y} \right) = \alpha_n\beta_n \left(\widehat{y}^\top \left(\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1} \right)^{-1} \widehat{y} \right).$$

We can thus verify the claim that $\operatorname{sign}\left(e_n^\top \mathcal{X}'(0)\widetilde{v}_n'(0)\right) = \operatorname{sign}(\alpha_n\beta_n)$ if we can show the above quadratic form is a positive quantity. Equivalently, we show that $\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1}$ is symmetric positive definite, as this implies $\left(\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1}\right)^{-1}$ is symmetric positive definite as well, proving the claim. Observe that by definition of $\widehat{B}$, $\mathcal{X}_{n-1}$, and $\widehat{A}$, we have

$$\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} 1 & -\frac{1}{d_2+i\omega} & \cdots & -\frac{1}{d_{n-1}+i\omega} \\ -\frac{1}{d_2-i\omega} & \frac{1}{|d_2+i\omega|^2} & \cdots & \frac{1}{(d_2-i\omega)(d_{n-1}+i\omega)} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{d_{n-1}-i\omega} & \frac{1}{(d_{n-1}-i\omega)(d_2+i\omega)} & \cdots & \frac{1}{|d_{n-1}+i\omega|^2} \end{bmatrix} d\omega. \tag{4.27}$$

Recall that the reachability Gramian $\mathcal{P}$ of an LTI system $\mathcal{G}$ is a symmetric positive definite matrix having the integral representation (2.4)

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (A - i\omega)^{-1} bb^{\top}(A^{\top} + i\omega)\, d\omega.$$

So, applying this definition along with formula (4.2) in the case of the subsystem $\mathcal{G}_{n-1}$ having an arrowhead realization with $A_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$ and $b = e_1 \in \mathbb{R}^{(n-1)\times 1}$, the corresponding reachability Gramian (that we denote $\mathcal{P}_{n-1}$) has the integral representation

$$\mathcal{P}_{n-1} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \begin{bmatrix} 1 & -\frac{\alpha_2}{d_2+i\omega} & \cdots & -\frac{\alpha_{n-1}}{d_{n-1}+i\omega} \\ -\frac{\alpha_2}{d_2-i\omega} & \frac{\alpha_2^2}{|d_2+i\omega|^2} & \cdots & \frac{\alpha_2\alpha_{n-1}}{(d_2-i\omega)(d_{n-1}+i\omega)} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\alpha_{n-1}}{d_{n-1}-i\omega} & \frac{\alpha_2\alpha_{n-1}}{(d_{n-1}-i\omega)(d_2+i\omega)} & \cdots & \frac{\alpha_{n-1}^2}{|d_{n-1}+i\omega|^2} \end{bmatrix} d\omega.$$

Now, note that (4.27) implies

$$\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1} = \widehat{A}^{-1}\mathcal{P}_{n-1}\widehat{A}^{-1},$$

i.e., $\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1}$ is a congruence transformation of the reachability Gramian $\mathcal{P}_{n-1}$. By Sylvester's law of inertia [9, Prop. 6.15], both $\widehat{A}^{-1}\mathcal{P}_{n-1}\widehat{A}^{-1}$ and $\mathcal{P}_{n-1}$ have the same number of positive eigenvalues. Because $\mathcal{P}_{n-1}$ is symmetric positive definite, this implies that $\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1} = \widehat{A}^{-1}\mathcal{P}_{n-1}\widehat{A}^{-1}$ is symmetric positive definite as well. From the definition of positive definiteness, it follows that $\widehat{y}^\top \left(\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1}\right)^{-1}\widehat{y}$ is a positive quantity, and so it follows that

$$\operatorname{sign}(e_n^\top \mathcal{X}'(0)\widetilde{v}_n'(0)) = \operatorname{sign}\left(\alpha_n\beta_n\left(\widehat{y}^\top\left(\widehat{B}^{-1}\mathcal{X}_{n-1}\widehat{A}^{-1}\right)^{-1}\widehat{y}\right)\right) = \operatorname{sign}(\alpha_n\beta_n),$$

as claimed. Recalling formula (4.10) for $\widetilde{\lambda}_n''(0)$, this implies

$$\operatorname{sign}\left(\widetilde{\lambda}_n''(0)\right) = \operatorname{sign}\left(e_n^\top\mathcal{X}''(0)e_n + e_n^\top\mathcal{X}'(0)\widetilde{v}_n'(0)\right) = \operatorname{sign}(\alpha_n\beta_n).$$

Applying this to the second order power series expansion of $\widetilde{\lambda}_n(\tau)$ about $\tau = 0$ as given in (4.24), for sufficiently small $\tau > 0$, it follows that

$$\operatorname{sign}(\widetilde{\lambda}_n(\tau)) = \operatorname{sign}(\widetilde{\lambda}_n''(0)) = \operatorname{sign}(\alpha_n\beta_n).$$

By the previous perturbation series argument, the eigenvalues of $M(\tau)$ and $\mathcal{X}(\tau)$ agree for perturbations up to $\mathcal{O}(\tau^2)$. Because $\widetilde{\lambda}_n''(0)$ is an $\mathcal{O}(\tau^2)$ term, and the eigenvalues of $\mathcal{X}(\tau)$ and $M(\tau)$ behave the same under perturbations of this order, we may conclude that

$$\operatorname{sign}(\lambda_n(\tau)) = \operatorname{sign}(\widetilde{\lambda}_n(\tau)) = \operatorname{sign}(\alpha_n\beta_n)$$

for small enough values of τ. By taking τ sufficiently small, we can also guarantee that

$$\operatorname{sign}(\lambda_1(\tau)) = +1, \quad \operatorname{sign}(\lambda_i(\tau)) = \operatorname{sign}(\alpha_i\beta_i), \quad i = 2, \dots, n-1.$$

This follows from our previous claim involving the eigenvalues of $\mathcal{X}(0)$ and the induction hypothesis, along with the fact that the eigenvalues of $\mathcal{X}(\tau)$ are continuous functions of τ. By taking τ small enough, we can guarantee these eigenvalues remain nonzero, and thus maintain their sign.

We claim that these statements are sufficient to show that the given equalities hold for all $\tau \in (0, 1]$. For each such τ, $\mathcal{X}(\tau)$ is the cross Gramian of a minimal system $\mathcal{G}(\tau)$ having an arrowhead realization (4.1) with $A = A(\tau)$ as in (4.14). Suppose that for some $\tau \in (0, 1]$, $\lambda_i(\tau) = 0$ for some $i = 1, 2, \ldots, n$. This would imply that $\mathcal{X}(\tau)$ is rank-deficient. For SISO systems, it is known that the system Gramians satisfy the relationship $\mathcal{X}^2 = \mathcal{P}\mathcal{Q}$ [26], so $\mathcal{X}(\tau)$ being rank-deficient implies that either $\mathcal{P}(\tau)$ or $\mathcal{Q}(\tau)$ is as well. This contradicts either the full reachability or observability of $\mathcal{G}(\tau)$, and thus the minimality of $\mathcal{G}(\tau)$ for all $0 < \tau \leq 1$. So it must be that all eigenvalues $\lambda_i(\tau)$ of $\mathcal{X}(\tau)$ remain nonzero for all $\tau \in (0, 1]$, else this would contradict the minimality of the corresponding system. Because these eigenvalues can never be 0 for any such τ, they can never change sign. So, taking $\tau \to 1$, we may conclude that

$$s_{\pi_1} = \operatorname{sign}(\lambda_{\pi_1}(1)) = +1, \quad s_{\pi_i} = \operatorname{sign}(\lambda_{\pi_i}(1)) = \operatorname{sign}(\alpha_i \beta_i),$$

where $\pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$ is some permutation of $\{1, 2, \ldots, n\}$. This permutation stems from the fact that while we can determine the signs of the eigenvalues of $\mathcal{X}(\tau)$, we cannot determine their eventual magnitude as $\tau \to 1$, and thus cannot deduce which sign parameter $\operatorname{sign}(\alpha_i \beta_i)$ corresponds to. Thus we have proved the induction argument.

Lastly, we return to the verification of the base case of the induction argument. This is done by applying the general argument in the induction step to a system of order 2. Consider a minimal and asymptotically stable system $\mathcal{G}$ having an arrowhead realization

$$A = \begin{bmatrix} d_1 & \alpha_2 \\ \beta_2 & d_2 \end{bmatrix}, \quad b = c^\top = e_1 \in \mathbb{R}^2.$$

Similarly, we invoke a homotopy on the realization given above, defined by (4.14), and explicitly given by

$$A(\tau) = A_0 + \tau\beta_2 e_2 e_1^\top + \tau\alpha_2 e_1 e_2^\top, \quad A_0 = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}.$$

Then, the cross Gramian (4.16) in this case is

$$\mathcal{X}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,\tau)|^2 \begin{bmatrix} 1 & -\tau\frac{\alpha_2}{d_2+i\omega} \\ -\tau\frac{\beta_2}{d_2-i\omega} & \tau^2\frac{\alpha_2\beta_2}{|d_2+i\omega|^2} \end{bmatrix} d\omega,$$

where

$$\rho(\pm i\omega,\tau) = \frac{1}{(d_1 \pm i\omega) - \tau^2\frac{\alpha_2\beta_2}{d_2\pm i\omega}}.$$

So, for $\tau = 0$ this becomes

$$\mathcal{X}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{|d_1+i\omega|^2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} d\omega,$$

which is a rank-1 matrix having one eigenvalue given by $\lambda_1(0) = e_1^\top\mathcal{X}(0)e_1$ with left and right eigenvectors $w_1(0) = v_1(0) = e_1 \in \mathbb{R}^2$, and a second eigenvalue given by $\lambda_2(0) = 0$, with left and right eigenvectors $w_2(0) = v_2(0) = e_2 \in \mathbb{R}^2$. Note that $\lambda_1(0)$ is necessarily positive because it is the integral of a positive quantity. As in the previous argument, we consider the second order power series approximation about $\tau = 0$ to $\mathcal{X}(\tau)$, namely $\mathcal{X}(\tau) = \mathcal{X}(0)+\mathcal{X}'(0)\tau+\mathcal{X}''(0)(\tau^2/2)$, where $\mathcal{X}'(0)$ and $\mathcal{X}''(0)$ are given as in (4.18) and (4.19) for $n = 2$. Applying the same eigenvalue derivative calculations as in the induction step, we get that

$$\widetilde{\lambda}_2'(0) = 0,$$

and

$$\widetilde{\lambda}_2''(0) = e_2^\top \mathcal{X}''(0)e_2 + 2e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0).$$

As before, $\operatorname{sign}(e_2^\top \mathcal{X}''(0)e_2) = \operatorname{sign}(\alpha_2\beta_2)$. We can again use relationship (4.25) to solve for $e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0)$. In this case (4.25) becomes

$$\mathcal{X}'(0)e_2 = -\mathcal{X}(0)\widetilde{v}_2'(0),$$

which, plugging in for $\mathcal{X}'(0)$ and $\mathcal{X}(0)$, becomes

$$\begin{bmatrix} 0 & y \\ w & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = - \begin{bmatrix} e_1^\top \mathcal{X}(0)e_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z \\ a \end{bmatrix},$$

where $w, y, z, \alpha \in \mathbb{R}$. Multiplying this out and solving for z gives

$$z = -\frac{y}{e_1^\top \mathcal{X}(0)e_1},$$

because $e_1^\top \mathcal{X}(0)e_1$ is a nonzero scalar. Plugging this into $e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0)$, we get

$$e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0) = -\frac{wy}{e_1^\top \mathcal{X}(0)e_1},$$

where w and y are the nonzero entries of $\mathcal{X}'(0)$. By expanding this further, this equation becomes

$$\begin{aligned}
-\frac{wy}{e_1^\top \mathcal{X}(0) e_1} &= -\frac{\frac{1}{(2\pi)^2}\left(\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{-\beta_2}{d_2 - i\omega}\, d\omega\right)\left(\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{-\alpha_2}{d_2 + i\omega}\, d\omega\right)}{\frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2\, d\omega} \\
&= -\alpha_2\beta_2 \frac{\frac{1}{2\pi}\left(\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_2 - i\omega}\, d\omega\right)\left(\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_2 + i\omega}\, d\omega\right)}{\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2\, d\omega} \\
&= \alpha_2\beta_2 \frac{\frac{1}{2\pi}\left(\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_2 + i\omega}\, d\omega\right)^2}{\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2\, d\omega},
\end{aligned}$$

where the last equality follows from noting that

$$\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_2 + i\omega}\, d\omega = -\int_{-\infty}^{\infty} |\rho(i\zeta,0)|^2 \frac{1}{d_2 - i\zeta}\, d\zeta,$$

for the change of variable $\omega = -\zeta$. The integral in the denominator of the above expression for $e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0)$ is that of a positive quantity, and thus itself positive. It can be shown that the integral in the numerator of the above expression for $e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0)$ is a real quantity via the Residue Theorem, and so its square is positive as well. Thus it follows that $\operatorname{sign}(e_2^\top \mathcal{X}'(0)\widetilde{v}_2'(0)) = \operatorname{sign}(\alpha_2\beta_2)$, and so $\operatorname{sign}(\widetilde{\lambda}_2''(0)) = \operatorname{sign}(\alpha_2\beta_2)$. Applying the same perturbation argument as in the induction step, from this it follows that

$$\operatorname{sign}(\lambda_2(\tau)) = \operatorname{sign}(\alpha_2\beta_2)$$

for small enough $\tau > 0$. We can apply the same state theoretic contradiction argument at the end of the induction step to conclude that the signs of the eigenvalues of $\mathcal{X}(\tau)$ must stay the same for all $\tau \in (0,1]$. This gives us that

$$s_{\pi_1} = \operatorname{sign}(\lambda_1(1)) = +1, \quad s_{\pi_2} = \operatorname{sign}(\lambda_2(1)) = \operatorname{sign}(\alpha_2\beta_2),$$

where $\pi = \{\pi_1, \pi_2\}$ is a permutation of $\{1, 2\}$, thus verifying the base case, and completing the proof by induction. □

We now justify the previous claim that $\mathcal{G}(\tau)$ can without loss of generality be taken as asymptotically stable in the following lemma.

Lemma 4.4. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal SISO system as in (1.3) with an arrowhead realization (4.1) such that $d_i < 0$ for all $i = 1, \dots, n$. Let $A(\tau)$ be the homotopy defined in (4.14), and define $\mathcal{G}(\tau)$ to be the linear system having the arrowhead realization (4.1) with A replaced by $A(\tau)$. If $\mathcal{G}(0)$ and $\mathcal{G}(1)$ are asymptotically stable, then there exists a continuous path of matrices $\widetilde{A}(\tau)$ from $\widetilde{A}(0) = A(0)$ to $\widetilde{A}(1) = A(1)$, such that the system $\widetilde{\mathcal{G}}(\tau)$ having the arrowhead realization (4.1) with A replaced by $A(\tau)$ is asymptotically stable for all $\tau \in [0, 1]$. Further, $\widetilde{\mathcal{G}}(\tau)$ is minimal for $\tau \in (0, 1]$.*

Proof. If for all $\tau \in [0, 1]$ $A(\tau)$ as defined in (4.14) has all of its eigenvalues in the left half-plane, the result follows trivially. Thus, suppose the contrary, namely that for some $\tau \in [\tau_1, \tau_2]$ $A(\tau)$ has an eigenvalue in the closed right half-plane, where $\tau_1, \tau_2 \in (0, 1)$. Without loss of generality we can choose $\tau_1 > 0$ and $\tau_2 < 1$ such that $A(\tau)$ has its eigenvalues in the open left half-plane for all $\tau \notin [\tau_1, \tau_2]$. Thus $\mathcal{G}(\tau)$ is asymptotically stable for all $\tau \notin [\tau_1, \tau_2]$. Let the eigenvalues of $A(\tau)$ be given by $\lambda_i(\tau)$ for $i = 1, \dots, n$ and define:

$$\alpha = \max_{\tau \in [\tau_1, \tau_2]} \text{real}\left(\lambda_i(\tau)\right).$$

Note that $\alpha \geq 0$ by definition since $A(\tau)$ has an eigenvalue in the closed right half-plane for some $\tau \in [\tau_1, \tau_2]$. Choose $\tau_0 > 0$ small enough and $\tau_3 < 1$ big enough such that $A(\tau)$ has its eigenvalues in the left half-plane for all $\tau \in [0, \tau_0]$ and $[\tau_3, 1]$, i.e., choose $\tau_0 < \tau_1$ and $\tau_3 > \tau_2$.

For some small $\epsilon > 0$, let $\alpha_\epsilon = \alpha + \epsilon$, and define the following function:

$$h(\tau) = \begin{cases} 0, & 0 \leq \tau \leq \tau_0 \\ h_1(\tau), & \tau_0 < \tau < \tau_1 \\ \alpha_\epsilon, & \tau_1 \leq \tau \leq \tau_2 \\ h_2(\tau) & \tau_2 < \tau < \tau_3 \\ 0 & \tau_3 \leq \tau \leq 1 \end{cases}, \tag{4.28}$$

where $h_1(\tau)$ and $h_2(\tau)$ are chosen as monotonically increasing and monotonically decreasing linear functions, respectively, such that $h_1(\tau) \geq 0$ for all $\tau \in [\tau_0, \tau_1]$ and $h_2(\tau) \geq 0$ for all $\tau \in [\tau_2, \tau_3]$. Further, define $h_1(\tau)$ and $h_2(\tau)$ such that $h(\tau)$ is continuous for all τ. Note that this definition implies that $h(\tau)$ is real-valued and such that $h(\tau) \geq 0$ for all $\tau \in [0, 1]$. Define $\widetilde{A}(\tau) = A(\tau) - h(\tau)I$. By definition of $h(\tau)$, $\widetilde{A}(\tau)$ is a continuous function of τ with $\widetilde{A}(0) = A(0)$ and $\widetilde{A}(1) = A(1)$. Let $\widetilde{\mathcal{G}}(\tau)$ be the system having the arrowhead realization with $\widetilde{A}(\tau)$, and claim that $\widetilde{\mathcal{G}}(\tau)$ is asymptotically stable for all $\tau \in [0, 1]$. Let the eigenvalues of $\widetilde{A}(\tau)$ be given by $\widetilde{\lambda}_i(\tau)$ for $i = 1, \ldots, n$. Because $\widetilde{\mathcal{G}}(\tau)$ agrees with $\mathcal{G}(\tau)$ for $\tau \in [0, \tau_0]$ and $[\tau_3, 1]$, $\widetilde{\mathcal{G}}(\tau)$ is asymptotically stable for such τ by our previous choices of τ_0 and τ_3. For each $\tau \in (\tau_0, \tau_3)$, by definition of $\widetilde{A}(\tau)$ each of its diagonal entries are given by those of $A(\tau)$ minus $h(\tau)$, i.e., the diagonal entries of $\widetilde{A}(\tau)$ are obtained by perturbing the diagonal entries of $A(\tau)$ by the same value. Thus, for each $\tau \in (\tau_0, \tau_3)$ the eigenvalues of $\widetilde{A}(\tau)$ are given by

$$\widetilde{\lambda}_i(\tau) = \lambda_i(\tau) - h(\tau),$$

and so in particular

$$\text{real}(\widetilde{\lambda}_i(\tau)) = \text{real}(\lambda_i(\tau)) - h(\tau). \tag{4.29}$$

For $\tau \in (\tau_0, \tau_1)$ and (τ_2, τ_3), as stated previously real$(\lambda_i(\tau)) < 0$ for all i, thus real$(\widetilde{\lambda}_i(\tau)) < 0$ by equation (4.29). Thus $\widetilde{\mathcal{G}}(\tau)$ is asymptotically stable for these values of τ. For $\tau \in [\tau_1, \tau_2]$, by equation (4.29) we have

$$\text{real}(\widetilde{\lambda}_i(\tau)) = \text{real}(\lambda_i(\tau)) - h(\tau) = \text{real}(\lambda_i(\tau)) - \alpha_\epsilon < 0,$$

because by definition of α, real$(\lambda_i(\tau)) \leq \alpha < \alpha_\epsilon$ for all $i = 1, \ldots, n$. Thus $\widetilde{\mathcal{G}}(\tau)$ is asymptotically stable for these values of τ, and so $\mathcal{G}(\tau)$ is asymptotically stable for all $\tau \in [0, 1]$ as claimed.

We next show that $\widetilde{\mathcal{G}}(\tau)$ is minimal for all τ. Because $\mathcal{G}$ is assumed to be minimal, $\alpha_i, \beta_i \neq 0$ for all $i = 1, \ldots n$, and $d_i \neq d_j$ for all $i \neq j$. For $i = 1, \ldots, n$, let $\widetilde{d}_i(\tau)$ be the diagonal entries of $\widetilde{A}(\tau)$. Because the off-diagonal entries of $A(\tau)$ and $\widetilde{A}(\tau)$ are the same, we need only verify that $\widetilde{d}_i \neq \widetilde{d}_j$ for $i \neq j$. By hypotheses $d_i < 0$ for all $i = 1, \ldots, n$, so $\widetilde{d}_i(\tau) = d_i - h(\tau) < 0$ for all $i = 1, \ldots, n$ as well. Further the fact that $d_i \neq d_j$ implies that for any given τ $d_i - h(\tau) \neq d_j - h(\tau)$, and so $\widetilde{d}_i \neq \widetilde{d}_j$ for $i \neq j$. Thus, it follows by definition of $\widetilde{A}(\tau)$ that the system $\widetilde{\mathcal{G}}(\tau)$ is minimal for all $\tau \in (0, 1]$ according to Lemma 4.1.

□

Remark 4.5. We note that the assumptions that $\mathcal{G}$ be asymptotically stable with an arrowhead realization (4.1) such that $d_i < 0$ for all $i = 1, \ldots, n$ implies collectively that the zeros and poles of $G(s)$ lie in the left half-plane. Systems having this property are called *minimum-phase* systems. In other words, the hypotheses of Theorem 4.3 can be restated by assuming that $\mathcal{G}$ is a minimum-phase arrowhead system.

We provide an elementary example to illustrate how different arrowhead realizations of a system $\mathcal{G}$ correspond to different orderings of the system's sign parameters.

Example 4.6. We construct a system $\mathcal{G}$ of order $n = 3$ having the arrowhead realizations

$$A_1 = \begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 0 \\ 1 & 0 & -3 \end{bmatrix}, \quad \text{and} \quad A_2 = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -3 & 0 \\ -1 & 0 & -2 \end{bmatrix},$$

with $b_1 = c_1^\top = e_1$ and $b_2 = c_2^\top = e_1$. These realizations correspond to an equivalent system, as they differ up to a symmetric permutation via $P = \begin{bmatrix} e_1 & e_3 & e_2 \end{bmatrix}$, where $e_2 = [0\ 1\ 0]^\top$, $e_3 = [0\ 0\ 1]^\top \in \mathbb{R}^3$. Calculating the eigenvalues of the corresponding cross Gramian, we see

$$(\lambda_1, \lambda_2, \lambda_3) = (4.46 \times 10^{-1}, -1.81 \times 10^{-2}, 6.35 \times 10^{-4}),$$

and so Theorem 4.2 implies that the sign parameters of $\mathcal{G}$ are given by

$$(s_1, s_2, s_3) = (+1, -1, +1).$$

Applying Theorem 4.3 to the given realizations (A_i, b_i, c_i) for $i = 1, 2$ of $\mathcal{G}$, the case of $i = 1$ implies that $\mathcal{G}$'s sign parameters are such that $s_{\pi_1} = +1$, $s_{\pi_2} = -1$, and $s_{\pi_3} = +1$, with $\pi = \{1, 2, 3\}$. For $i = 2$, the result implies $s_{\nu_1} = +1$, $s_{\nu_2} = +1$, and $s_{\nu_3} = -1$, for the permutation $\nu = \{1, 3, 2\}$.

This example suggests that if a system $\mathcal{G}$ has an arrowhead realization (4.1), and $s_1 = \text{sign}(\gamma)$, there exists a permutation of the realization such that the resulting one corresponds to the canonical ordering of $\mathcal{G}$'s sign parameters; that is, π in Theorem 4.3 is identity. We show this fact next.

Corollary 4.7. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal SISO system as in (1.3) with an arrowhead realization (4.1) satisfying the hypotheses of Theorem 4.3. Then* *there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$ such that*

$$s_k = sign(\gamma) \quad and \quad s_i = sign(\gamma\widetilde{\alpha}_i\widetilde{\beta}_i), \quad i = 1,\ldots,n, \quad i \neq k,$$

where $\widetilde{\alpha}_i, \widetilde{\beta}_i$ are the off-diagonal entries of the permuted arrowhead matrix given by $\widetilde{A} = P^{\top}AP$, with $\widetilde{b} = P^{\top}b = \gamma e_k$ and $c = cP = e_k$.

Proof. From Theorem 4.3, the sign parameters of $\mathcal{G}$ are given by (4.11). By rearranging the sign parameters into their canonical order, (4.11) implies that there is a permutation $\nu = \{\nu_1, \nu_2, \ldots, \nu_{n-1}\}$ of $\{2,\ldots,n\}$ such that $s_k = \operatorname{sign}(\gamma)$ for some $k = 1,2,\ldots,n$ and $s_i = \operatorname{sign}(\gamma\alpha_{\nu_i}\beta_{\nu_i})$ for all $i = 2,\ldots,n$ with $i \neq k$. Let $P \in \mathbb{R}^{n\times n}$ be the permutation matrix given by $P = [e_{\nu_1} \cdots e_k \cdots e_{\nu_n}] \in \mathbb{R}^{n\times n}$. Applying this permutation to $\mathcal{G}$'s arrowhead realization produces a *permuted* arrowhead realization of $\mathcal{G}$ given by $\widetilde{A} = P^{\top}AP$, $\widetilde{b} = \gamma e_k$, and $\widetilde{c} = e_k$, where $e_k \in \mathbb{R}^n$ is the kth standard basis vector. Note that $\widetilde{A}$ no longer necessarily has the structure (4.1); it can be given by any symmetric permutation of this form. In this permuted realization, denote the off-diagonal entries $\widetilde{\alpha}_i = \alpha_{\nu_i}$ and $\widetilde{\beta}_i = \beta_{\nu_i}$ for $i = 1,\ldots,n$, $i \neq k$. Then, applying (4.11) under this relabeling gives

$$s_k = \operatorname{sign}(\gamma), \quad s_i = \operatorname{sign}(\gamma\widetilde{\alpha}_i\widetilde{\beta}_i), \quad i = 1,\ldots,n \quad i \neq k.$$

□

We note that with an appropriate diagonal scaling, this select permuted arrowhead realization obeys the same sign symmetry condition as the system's canonical balanced form (3.1). In the particular case where a system $\mathcal{G}$ that satisfies the hypotheses of Theorem 4.3 has an arrowhead realization (4.1) satisfying $\gamma = +1$, and $\operatorname{sign}(\alpha_i\beta_i) = +1$ for $i = 2,\ldots,n$, Theorem 4.3 implies that $s_i = +1$ for all $i = 1,\ldots,n$. In other words, $\mathcal{G}$ is state-space symmetric according to Definition 3.3. This agrees with the observation that the given realization of $\mathcal{G}$ satisfies $A = A^\top$ and $b = c^\top$.

We finish this discussion with a special case of Theorem 4.3 applied to a certain class of systems, in which the canonical ordering of a system's sign parameters can be determined *a priori* from any arrowhead realization of the system. We present an example of one such system arising naturally in power systems modeling.

4.3 A Special Class of Systems, with an Example from Power Systems Modeling

We next apply Theorem 4.3 to a special class of systems, namely those having an arrowhead realization for which the signs of the products of the off-diagonal entries are consistent. For these systems, we show that this specific sign structure of the arrowhead realization in conjunction with Theorem 4.3 reveals the canonical ordering on the system's sign parameters. To motivate our discussion, we first provide an example of one such system arising in power systems modeling. We first discovered the phenomenon characterized in Theorems 3.6 and 4.3 while studying this model.

4.3.1 Derivation of the Model

We begin by providing the relevant discussion information necessary to derive the model. A common technique for modeling the frequency response of a power network containing n coherent generators is to aggregate the system response and treat the network as a single effective machine. A group of generators is said to be *coherent* if they exhibit similar frequency response under power deviations. It is shown in [27, 28] that for a network of N coherent generators, the aggregate frequency dynamics of the network are well-approximated by an order $N+1$ system $\mathcal{G}$ having the transfer function

$$G(s) = \frac{1}{\widehat{m}s + \widehat{d} + \sum_{i=1}^{N} \frac{r_i^{-1}}{\tau_i s+1}}, \tag{4.30}$$

for generators given by the swing model with first-order turbine control. Here $\widehat{m}$ and $\widehat{d}$ denote the aggregate inertia and damping coefficients of the generators in the network, while τ_i and r_i^{-1} denote the time constant and droop coefficient of the ith generator. For the theoretical justification that $G(s)$ sufficiently approximates the network response, see [27, Sec. 2].

It is readily verified using (4.2) that the corresponding system $\mathcal{G}$ has an arrowhead realization given in terms of the physical parameters of the network:

$$A = \begin{bmatrix} \frac{-\widehat{d}}{\widehat{m}} & \frac{1}{\widehat{m}} & \cdots & \frac{1}{\widehat{m}} \\ \frac{-r_1^{-1}}{\tau_1} & \frac{-1}{\tau_1} & & \\ \vdots & & \ddots & \\ \frac{-r_N^{-1}}{\tau_N} & & & \frac{-1}{\tau_N} \end{bmatrix}, \quad b = \frac{1}{\widehat{m}}e_1, \quad c = e_1^\top. \tag{4.31}$$

This realization can also be derived from the network dynamics of the individual generators. If each generator in the network is modeled by the swing equation with first order turbine control, the dynamics of each individual machine are governed by

$$w_i{}' = -\frac{d_i}{m_i}w_i + \frac{1}{m_i}q_i + \left(\frac{1}{m_i}\right)u_i - p_i^e, \tag{4.32}$$

$$q_i{}' = -\frac{r_i^{-1}}{\tau_i}w_i - \frac{1}{\tau_i}q_i, \tag{4.33}$$

for $i = 1, \ldots N$, where m_i and d_i are the inertia and damping of each generator, respectively, q_i is the (variation of) turbine power, $u_i - p_i^e$ is the net power deviation at the ith generator, and w_i is the angular frequency deviation of the generator rotor. Each generator can be modeled by a transfer function $G_i(s)$ mapping net power deviation to the frequency response w_i. For a group of coherent generators, equal frequency response is assumed, i.e., $w_i(s) = \widehat{w}(s)$ for all $i = 1, \ldots, N$. If one takes the state vector to be $x = [\widehat{w}\ q_1\ \cdots\ q_N]^\top$, then summing up equations (4.32) over all i and keeping each equation in (4.33) leads to a linear system of $N+1$ ODEs as in (1.3), mapping the aggregate power deviation $\sum_{i=1}^{N}(u_i - p_i^e)$ to the angular deviation $\widehat{w}$. In state-space form, this system is given by the arrowhead realization (4.31).

The physical parameters $\widehat{m}$, $\widehat{d}$, r_i^{-1}, and τ_i associated with the model are guaranteed to be positive for all $i = 1, \ldots, N$. The construction of the model additionally assumes that all subsystems are asymptotically stable. Thus Theorem 4.3 applied to the given system $\mathcal{G}$ having the arrowhead realization (4.31) guarantees that

$$s_{\pi_1} = \operatorname{sign}\left(\frac{1}{\widehat{m}}\right) = +1, \quad s_{\pi_i} = \operatorname{sign}\left(-\frac{r_i^{-1}}{\widehat{m}^2\tau_i}\right) = -1, \quad i = 2, \ldots, N+1.$$

For these systems, we observed numerically that $s_1 = +1$ and $s_i = -1$ for all $i = 2, \ldots, N+1$, suggesting that the permutation π given in the statement of Theorem 4.3 is the identity. In the result that follows, we prove this for a general class of arrowhead systems having this property. Namely, we show that if a system has an arrowhead realization (4.1) such that the products of the signs of the off-diagonal entries are negative, i.e. $\operatorname{sign}(\alpha_i\beta_i) = -1$ for all $i = 2, \ldots, n$, then the trailing sign parameters of $\mathcal{G}$ are consistent.

4.3.2 A Special Class of Arrowhead Systems

Corollary 4.8. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal SISO system as in (1.3) with an arrowhead realization as in (4.1) such that $\mathcal{G}$ satisfies the hypotheses of Theorem 4.3. Suppose further that the arrowhead realization (4.1) is such that sign$(\alpha_i\beta_i) =$*

-1 *for all* $i = 2, \ldots, n$. *Then the sign parameters of* $\mathcal{G}$ *are given by*

$$s_1 = sign(\gamma), \quad s_i = -sign(\gamma), \quad i = 2, \ldots, n.$$

Additionally, $\mathcal{G}$ *has distinct Hankel singular values.*

Proof. Without loss of generality suppose that $\text{sign}(\gamma) = +1$. Then $\text{sign}(\gamma\alpha_i\beta_i) = -1$ for all $i = 2, \ldots, n$. If the given arrowhead realization of $\mathcal{G}$ is the one corresponding to the permutation of $\mathcal{X}$ with nonincreasing absolute eigenvalues as in Theorem 4.3, then the result follows trivially. Otherwise, Theorem 4.3 only guarantees that $\mathcal{G}$ has one positive sign parameter and $n-1$ negative sign parameters. Recalling Theorem 4.2, the sign parameters of a system are the signs of the eigenvalues of its cross Gramian. Then in this case Theorem 4.3 implies $\mathcal{X}$ has $n-1$ negative eigenvalues and one positive eigenvalue. We claim that the positive sign parameter corresponds to the dominant eigenvalue of the cross Gramian; that is, $s_1 = +1$. As noted by [9, Remark 5.4.3], it can be shown that $2\text{tr}(\mathcal{X}) = -\text{tr}(cA^{-1}b)$. Using the fact that $b = \gamma e_1$ and $c^\top$, formula (4.2) gives

$$cA^{-1}b = \frac{\gamma}{d_1 - \sum_{i=2}^{n} \frac{\alpha_i\beta_i}{d_i}}.$$

Because this is a scalar quantity, $cA^{-1}b = \text{tr}(cA^{-1}b)$. By the assumption that $\mathcal{G}$ satisfies the hypotheses of Theorem 4.3, the diagonal entries of its arrowhead realization satisfy $d_i < 0$ for all $i = 1, \ldots, n$. Further, because $\text{sign}(\alpha_i\beta_i) = -1$ for all $i = 2, \ldots, n$, it follows that $\sum_{i=2}^{n} \frac{\alpha_i\beta_i}{d_i} > 0$. This, along with the assumption that $\text{sign}(\gamma) = +1$, implies

$$\text{tr}(cA^{-1}b) = cA^{-1}b = \frac{\gamma}{d_1 - \sum_{i=2}^{n} \frac{\alpha_i\beta_i}{d_i}} < 0.$$

Thus $\text{tr}(\mathcal{X}) = -\frac{1}{2}\text{tr}(cA^{-1}b) > 0$. Because the trace of $\mathcal{X}$ is equal to the sum of its eigenvalues, it must be that the dominant eigenvalue of $\mathcal{X}$ is the positive one. By Theorem 4.2, we conclude that $s_1 = +1$ and $s_i = -1$, for all $i = 2, \ldots, n$. By Theorem 3.6, this further implies that the $n-1$ trailing Hankel singular values of $\mathcal{G}$, and thus all Hankel singular values of $\mathcal{G}$, are distinct. For the case of $\text{sign}(\gamma) = -1$, the proof follows nearly identically. Theorem 4.3 implies $\mathcal{X}$ has one negative eigenvalue and $n-1$ positive ones. From the hypotheses of the theorem it follows that $\text{tr}(\mathcal{X}) = -\frac{1}{2}\text{tr}(cA^{-1}b) < 0$, implying that the dominant eigenvalue of $\mathcal{X}$ is negative. Thus $s_1 = -1$ and $s_i = +1$ for all $i = 2, \ldots, n$. □

For systems having an arrowhead realization in which the products of the signs of the off-diagonal entries are all negative, Corollary 4.8 implies that the trailing sign parameters of the system are consistent, and can be determined directly from the off-diagonal entries of the arrowhead matrix using the formula

$$s_1 = \text{sign}(\gamma), \quad s_i = \text{sign}(\gamma\alpha_i\beta_i), \quad i = 2, \ldots, n.$$

In other words, the sign parameters of systems having an arrowhead realization satisfying the hypotheses of Corollary 4.8 undergo the single sign flip after the first parameter. This implies directly that these systems obey necessary sign consistency conditions in Theorem 3.6 for *all* orders of reduction; and so the balanced truncation $\mathcal{H}_\infty$ error bound (2.15) will always be hold with equality. This is stated formally with the following corollary. In Subsection 4.3.3, we illustrate this property with a specific example of the power systems model introduced in Subsection 4.3.1.

Corollary 4.9. *Let $\mathcal{G}$ be an order-n asymptotically stable and minimal SISO system as in (1.3) satisfying the hypotheses of Theorem 4.3 with an arrowhead realization (4.1) satisfying the hypotheses of Corollary 4.8. Let $\mathcal{G}_r$ denote the order-r approximation of $\mathcal{G}$ via balanced truncation as in (2.11). Then, the trailing sign parameters are the same, and $\mathcal{G}_r$ achieves* *the balanced truncation error bound* (2.15).

4.3.3 Numerical Results

We illustrate Theorem 3.6 and Corollary 4.8 with a specific example of the power systems model described previously in Section 4.3. Consider the case of a network with $N = 4$ coherent generators, which leads to a SISO LTI system of order $n = N + 1 = 5$. Take $\widehat{m} = 0.044$, $\widehat{d} = 0.038$,

$$(r_1^{-1}, r_2^{-1}, r_3^{-1}, r_4^{-1}) = (0.013, 0.014, 0.022, 0.025),$$

and

$$(\tau_1, \tau_2, \tau_3, \tau_4) = (5.01, 6.82, 7.38, 7.79).$$

Recalling that the physical parameters associated with the system are all positive, the realization of $\mathcal{G}$ given by (4.31) in conjunction with Corollary 4.8 implies that the sign parameters of $\mathcal{G}$ are given by

$$s_1 = \operatorname{sign}\left(\frac{1}{\widehat{m}}\right) = +1, \quad s_i = \operatorname{sign}\left(-\frac{r_i^{-1}}{\widehat{m}^2 \tau_i}\right) = -1, \quad i = 2, \ldots, 5.$$

This is verified by computing a balanced realization of $\mathcal{G}$ satisfying the symmetry condition (3.1) with the sign matrix S defined as

$$S = \operatorname{diag}(1, -1, -1, -1, -1).$$

We next calculate the Hankel singular values of the system $\mathcal{G}$ with transfer function $G(s)$ in (4.30):

$$\Sigma = \text{diag}(11.63, 7.13, 3.53 \times 10^{-2}, 8.48 \times 10^{-5}, 4.12 \times 10^{-8}).$$

Now we construct the canonical balanced realization of $\mathcal{G}$ using formula (3.2) and compute order-r balanced truncation approximations to $\mathcal{G}$ for $r = 2$, 3, and 4. Under these conditions the truncated systems obey the sign consistency in (3.8). As in the previous example we highlight this symmetry by partitioning the system for $r = 3$:

$$A = \left[\begin{array}{ccc|cc} -0.9913 & 0.5924 & -0.0467 & 0.0020 & 0.0000 \\ -0.5924 & -0.0216 & 0.0087 & -0.0004 & -0.0000 \\ 0.0467 & 0.0087 & -0.1800 & 0.0157 & 0.0003 \\ \hline -0.0020 & -0.0004 & 0.0157 & -0.1437 & -0.0062 \\ -0.0000 & -0.0000 & 0.0003 & -0.0062 & -0.1372 \end{array}\right],$$

$$b = \left[\begin{array}{c} -4.8009 \\ -0.5552 \\ 0.1126 \\ \hline -0.0049 \\ -0.0001 \end{array}\right], \quad c^\top = \left[\begin{array}{c} -4.8021 \\ 0.5552 \\ -0.1126 \\ \hline 0.0049 \\ 0.0001 \end{array}\right], \quad d = 0.$$

Table 4.1 compares the $\mathcal{H}_\infty$ norm of the error system to the balanced truncation upper bound (2.15). Because the trailing sign parameters of $\mathcal{G}$ are all -1, we can perform truncation at any order $r \geq 1$ and the truncated system will satisfy the sign requirements (3.8) of Theorem 3.6. Thus the balanced truncation error bound holds with equality for approximations of all orders.

Table 4.1: $\mathcal{H}_\infty$ norm of the error system, compared to the balanced truncation upper bound (2.15) for a power system, $N = 4$.

	$\|\mathcal{G} - \mathcal{G}_r\|_{\mathcal{H}_\infty}$	$2(\sigma_{r+1} + \cdots + \sigma_n)$
$r = 1$	1.747×10^{1}	1.747×10^{1}
$r = 2$	7.067×10^{-2}	7.067×10^{-2}
$r = 3$	1.697×10^{-4}	1.697×10^{-4}
$r = 4$	8.248×10^{-8}	8.248×10^{-8}

Chapter 5

Conclusions

In this work, we have shown through analysis and numerical study that the balanced truncation error bound (2.15) holds with equality for SISO systems satisfying the sign consistency condition (3.8), providing an explicit formula for the approximation error in terms of the system's Hankel singular values. This result generalizes Theorem 3.5 for state-space symmetric systems. In Theorem 4.3 we proved an explicit formula for the sign parameters corresponding to a system's Hankel singular values in terms of the entries of a system's arrowhead realization. We then strengthened this result for a special class of arrowhead systems in Corollary 4.8. From these results, one can verify whether the sign consistency condition (3.8) in Theorem 3.6 holds, and thus whether or not the corresponding order-r balanced truncation approximation achieves the $\mathcal{H}_\infty$ error bound (2.15). We motivated our discussion with an example of one such arrowhead system arising naturally in power systems modeling. Moving forward, there are a few topics for future work. We hope to find more examples of arrowhead systems arising from real world problems, like the power systems example described in Section 4.3. As stated previously, we hope to derive more rigorous conditions under which a SISO LTI system is guaranteed to have an arrowhead realization (4.1).

Appendices

Appendix A

Verification of Formula (4.18)

Proof. First note, that from definition of $A(\tau)$ given in (4.14), we have that

$$A'(0) = \beta_n e_n e_1^\top + \alpha_n e_1 e_n^\top$$

and $A(0) = A_0$ as defined in (4.14).

(Case 1: $i = j = n$) Because $A(0)e_n = d_n e_n$, $e_n^\top A(0) = d_n e_n^\top$, and $e_n^\top A'(0) = \beta_n e_1^\top$,

$$e_n^\top A'(0)\mathcal{X}(0)e_n = \beta_n e_1^\top \mathcal{X}(0)e_n = 0,$$

$$e_n^\top A(0)\mathcal{X}'(0)e_n = d_n e_n^\top \mathcal{X}'(0)e_n = 0,$$

$$e_n^\top \mathcal{X}'(0)A(0)e_n = d_n e_n^\top \mathcal{X}'(0)e_n = 0,$$

$$e_n^\top \mathcal{X}(0)A'(0)e_n = \alpha_n e_n^\top \mathcal{X}(0)e_n = 0.$$

So the (n, n) entry of (4.21) holds trivially.

(Case 2: $i = n$, $2 \leq j \leq n-1$) Noting $e_n^\top A'(0) = \beta_n e_1^\top$ and $A'(0)e_j = 0$,

$$
\begin{aligned}
e_n^\top A'(0)\mathcal{X}(0)e_j &= \beta_n e_1^\top \mathcal{X}(0)e_j \\
&= -\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_j + i\omega}\, d\omega \\
&= -\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{(d_n - i\omega)}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&= -d_n\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&\quad +\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{(d_j + i\omega)(d_n - i\omega)}\, d\omega,
\end{aligned}
$$

$$
\begin{aligned}
e_n^\top A(0)\mathcal{X}'(0)e_j &= d_n e_n^\top \mathcal{X}'(0)e_j \\
&= d_n\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega,
\end{aligned}
$$

$$
\begin{aligned}
e_n^\top \mathcal{X}'(0)A(0)e_j &= e_n^\top \mathcal{X}'(0)(d_j e_j + \alpha_j e_1) \\
&= d_j e_n^\top \mathcal{X}'(0)e_j + \alpha_j e_n^\top \mathcal{X}'(0)e_1 \\
&= d_j\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&\quad - \alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_n - i\omega}\, d\omega \\
&= d_j\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&\quad - \alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{(d_j + i\omega)}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&= d_j\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&\quad - d_j\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&\quad - \alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{(d_j + i\omega)(d_n - i\omega)}\, d\omega \\
&= -\alpha_j\beta_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{(d_j + i\omega)(d_n - i\omega)}\, d\omega,
\end{aligned}
$$

$$
e_n^\top \mathcal{X}(0)A'(0)e_j = 0.
$$

Summing these terms up leads to 0.

(Case 3: $2 \leq i \leq n-1$, $j = n$) This case is similar to the previous. Each term ends up being

$$e_i^\top A'(0)\mathcal{X}(0)e_n = 0$$

$$\begin{aligned} e_i^\top A(0)\mathcal{X}'(0)e_n &= (d_i e_i^\top + \beta_i e_i^\top)\mathcal{X}'(0)e_n \\ &= \alpha_n \beta_i \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{(d_n+i\omega)(d_i-i\omega)}\,d\omega \end{aligned}$$

$$\begin{aligned} e_i^\top \mathcal{X}'(0)A(0)e_n &= d_n e_i^\top \mathcal{X}'(0)e_n \\ &= d_n\alpha_n \beta_i \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_n+i\omega)(d_i-i\omega)}\,d\omega \end{aligned}$$

$$\begin{aligned} e_i^\top \mathcal{X}(0)A'(0)e_n &= \alpha_n e_i^\top \mathcal{X}(0)e_1 \\ &= -d_n\alpha_n \beta_i \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{(d_n+i\omega)(d_i-i\omega)}\,d\omega \\ &\quad -\alpha_n \beta_i \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{i\omega}{(d_n+i\omega)(d_i-i\omega)}\,d\omega. \end{aligned}$$

The terms sum up to zero here as well.

(Case 4: $i = j = 1$) In this case we simply have

$$e_1^\top A'(0)\mathcal{X}(0)e_1 = \alpha_n e_n^\top \mathcal{X}(0)e_1 = 0,$$

$$e_1^\top A(0)\mathcal{X}'(0)e_1 = d_1 e_1^\top \mathcal{X}'(0)e_1 + \sum_{j=2}^{n-1} \alpha_j e_j^\top \mathcal{X}'(0)e_1 = 0,$$

$$e_1^\top \mathcal{X}'(0)A(0)e_1 = d_1 e_1^\top \mathcal{X}'(0)e_1 + \sum_{j=2}^{n-1} \beta_j e_1^\top \mathcal{X}'(0)e_j = 0,$$

$$e_1^\top \mathcal{X}(0)A'(0)e_1 = \beta_n e_1^\top \mathcal{X}(0)e_n = 0.$$

(Case 5: $2 \leq i, j \leq n-1$) Noting that $e_i^\top \mathcal{X}'(0)e_j = 0$ for all $1 \leq i, j \leq n-1$ gives

$$e_i^\top A'(0)\mathcal{X}(0)e_j = 0,$$

$$e_i^\top A(0)\mathcal{X}'(0)e_j = 0,$$

$$e_i^\top \mathcal{X}'(0)A(0)e_j = 0,$$

$$e_i^\top \mathcal{X}(0)A'(0)e_j = 0.$$

Thus, this case is trivially 0 as well.

(Case 6: $i = 1,\ j = n$) In this case

$$e_1^\top A'(0)\mathcal{X}(0)e_n = \alpha_n e_n^\top \mathcal{X}'(0)e_n = 0,$$

$$\begin{aligned} e_1^\top A(0)\mathcal{X}'(0)e_n &= d_1 e_1^\top \mathcal{X}'(0)e_n + \sum_{i=2}^{n-1} \alpha_j e_i^\top \mathcal{X}'(0)e_n \\ &= -d_1\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_n + i\omega}\, d\omega \\ &\quad + \alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_n + i\omega} \sum_{i=2}^{n-1} \frac{\alpha_i\beta_i}{d_i - \omega}\, d\omega, \end{aligned}$$

$$\begin{aligned} e_i^\top \mathcal{X}'(0)A(0)e_j &= d_n e_1^\top \mathcal{X}'(0)e_n \\ &= -d_n\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \frac{1}{d_n + i\omega}\, d\omega, \end{aligned}$$

$$\begin{aligned} e_i^\top \mathcal{X}(0)A'(0)e_j &= \alpha_n e_1^\top \mathcal{X}(0)e_1 \\ &= \alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2\, d\omega. \end{aligned}$$

Summing up the terms inside the integrands and pulling out common factors gives

$$\begin{aligned} &\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \left(1 - \frac{d_1}{d_n + i\omega} - \frac{d_n}{d_n + i\omega} + \frac{1}{d_n + i\omega}\sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i - i\omega}\right) d\omega \\ =&\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \left(\frac{(d_n + i\omega) - d_1 - d_n + \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i - i\omega}}{d_n + i\omega}\right) d\omega \\ =&\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \left(\frac{-d_1 + i\omega + \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i - i\omega}}{d_n + i\omega}\right) d\omega \\ =&-\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty} |\rho(i\omega,0)|^2 \left(\frac{(d_1 - i\omega) - \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i - i\omega}}{d_n + i\omega}\right) d\omega. \end{aligned}$$

Recalling the definition of $\rho(\pm i\omega, 0)$, this becomes

$$-\alpha_n \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{(d_1+i\omega)-\sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i+i\omega}}\frac{1}{d_n+i\omega}\right)d\omega = -\alpha_n\frac{1}{2\pi}\int_{-\infty}^{\infty}\rho(i\omega,0)\eta(-i\omega)\,d\omega,$$

where $\eta(z) = 1/(d_n - z)$. This integral can be interpreted as an inner product over $\mathcal{L}_2(i\mathbb{R})$. This Hilbert space can be separated into a direct sum of orthogonal subspaces, namely $\mathcal{L}_2(i\mathbb{R}) = \mathcal{H}_2(\mathbb{C}^+)\bigoplus\mathcal{H}_2(\mathbb{C}^-)$, where $\mathcal{H}_2(\mathbb{C}^+)$ is the space of finite $\mathcal{L}_2$-norm functions analytic in the open right half-plane, and $\mathcal{H}_2(\mathbb{C}^-)$ is the space of finite $\mathcal{L}_2$-norm functions analytic in the open left half-plane [29, Section 2]. Because $\eta(z)$ only has a pole at $d_n < 0$, $\eta(z) \in \mathcal{H}_2(\mathbb{C}^+)$. Similarly, because the poles of $\mathcal{G}_{n-1}$ are given by $\rho(-z,0) = 1/\left(d_1 - z - \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i-z}\right)$, and $\mathcal{G}_{n-1}$ is asymptotically stable, the poles of $\rho(z,0)$ are in the right half-plane, and so $\rho(z,0)\in\mathcal{H}_2(\mathbb{C}^-)$. Then the above expression can be written as

$$\langle\rho(z),\eta(z)\rangle_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\rho(i\omega)\eta(-i\omega)\,d\omega = 0,$$

because these functions belong to orthogonal subspaces of $\mathcal{L}_2(i\mathbb{R})$. Thus the claim is verified this case.

(Case 7: $i = n$, $j = 1$) This case is nearly identical to the previous. One can show that the terms sum up to

$$-\beta_n\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\frac{1}{(d_1-i\omega)-\sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i-i\omega}}\frac{1}{d_n-i\omega}\right).$$

Similarly, this can be interpreted as an inner product over $\mathcal{L}_2(i\mathbb{R})$. Taking $\gamma(z) = 1/(d_n + i\omega)$, $\gamma(z)$ has a single pole at $-d_n > 0$, so $\gamma(z)\in\mathcal{H}_2(\mathbb{C}^-)$. The poles of $\mathcal{G}_{n-1}$ are given by $\rho(-z,0) = 1/\left(d_1 - z - \sum_{i=2}^{n-1}\frac{\alpha_i\beta_i}{d_i-z}\right)$, so $\rho(z,0)\in\mathcal{H}_2(\mathbb{C}^+)$. So, the above integral can be written as

$$\langle \rho(z), \gamma(z) \rangle_{\mathcal{L}_2(i\mathbb{R})} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \rho(i\omega)\gamma(-i\omega)\, d\omega = 0.$$

This proves the formula (4.18) for $\mathcal{X}'(0)$.

$\square$

Bibliography

[1] U. Baur, P. Benner, and L. Feng, "Model order reduction for linear and nonlinear systems: A system-theoretic perspective," *Archives of Computational Methods in Engineering*, vol. 21, no. 4, pp. 331–358, 2014.

[2] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced basis methods for partial differential equations: an introduction.* Springer, 2015, vol. 92.

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

[4] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, *Model Reduction and Approximation.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 2017.

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

[6] C. Mullis and R. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits Syst.*, vol. 23, no. 9, pp. 551–562, 1976.

[7] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Trans. Auto. Control*, vol. 26, no. 1, pp. 17–32, 1981.

[8] W. Q. Liu, V. Sreeram, and K. L. Teo, "Model reduction for state-space symmetric systems," *Systems Control Lett.*, vol. 34, pp. 209–215, 1998.

[9] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems.* Philadelphia: SIAM, 2005.

[10] L. Pernebo and L. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Auto. Control*, vol. 27, no. 2, pp. 382–387, 1982.

[11] D. F. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," in *The 23rd IEEE Conference on Decision and Control.* IEEE, 1984, pp. 127–132.

[12] D. Wilson and A. Kumar, "Symmetry properties of balanced systems," *IEEE Trans. Auto. Control*, vol. 28, no. 9, pp. 927–929, 1983.

[13] R. J. Ober, "Balanced realizations: canonical form, parametrization, model reduction," *Int. J. Control*, vol. 46, no. 2, pp. 643–670, 1987.

[14] S. Reiter, M. Embree, and S. Gugercin, "The balanced truncation bound is tight for SISO systems when the truncated system is state-space symmetric," *arXiv preprint arXiv:2011.07170*, 2020.

[15] R. Ober and D. McFarlane, "Balanced canonical forms for minimal systems: A normalized coprime factor approach," *Linear Algebra Appl.*, vol. 122, pp. 23–64, 1989.

[16] J. Maciejowski and R. Ober, "Balanced parametrizations and canonical forms for system identification," *IFAC Proceedings Volumes*, vol. 21, no. 9, pp. 701–708, 1988.

[17] M. R. Opmeer and T. Reis, "A lower bound for the balanced truncation error for MIMO systems," *IEEE Trans. Auto. Control*, vol. 60, no. 8, pp. 2207–2212, 2015.

[18] B. Srinivasan and P. Myszkorowski, "Model reduction of systems with zeros interlacing the poles," *Systems Control Lett.*, vol. 30, no. 1, pp. 19–24, 1997.

[19] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control.* Upper Saddle River, NJ: Prentice Hall, 1996.

[20] Y. Liu and B. D. Anderson, “Singular perturbation approximation of balanced systems,” *Int. J. Control*, vol. 50, no. 4, pp. 1379–1405, 1989.

[21] D. K. Salkuyeh and F. P. A. Beik, “An explicit formula for the inverse of arrowhead and doubly arrow matrices,” *International Journal of Applied and Computational Mathematics*, vol. 4, no. 3, pp. 1–8, 2018.

[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore: Johns Hopkins University Press, 2012.

[23] T. Kato, *Perturbation Theory for Linear Operators.* Berlin: Springer, 1980.

[24] A. Greenbaum, R.-C. Li, and M. L. Overton, “First-order perturbation theory for eigenvalues and eigenvectors,” *SIAM Review*, vol. 62, no. 2, pp. 463–482, 2020.

[25] N. J. Higham, *Accuracy and Stability of Numerical Algorithms.* SIAM, 2002.

[26] K. Fernando and H. Nicholson, “On the cross-gramian for symmetric MIMO systems,” *IEEE Trans. Circuits Syst.*, vol. 32, no. 5, pp. 487–489, 1985.

[27] H. Min, F. Paganini, and E. Mallada, “Accurate reduced order models for coherent synchronous generators,” in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton).* IEEE, 2019, pp. 316–317.

[28] ——, “Accurate reduced-order models for heterogeneous coherent generators,” 2019, extended preprint, arXiv:1909.12864.

[29] C. Magruder, C. Beattie, and S. Gugercin, "Rational Krylov methods for optimal $\mathcal{L}_2$ model reduction," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 6797–6802.