Department of Electrical Engineering
Control Systems Research Group
Master Systems & Control

An Extended State Framework for Direct Data-Driven Control

Master Thesis

Eashwar Ravikkumar
1777033

Supervisors:
prof. dr. ir. Roland Tóth
dr. ir. Valentina Breschi
ir. Luuk M. Spin

Committee Members:

prof. dr. ir. Roland Tóth	Chair
dr. ir. Valentina Breschi	Member
dr. Mircea Lazar	Internal expert
dr. ir. A.A.J. Lefeber	External expert
ir. Luuk M. Spin	Advisor

October 22, 2024

This report was made in accordance with the TU/e Code of Scientific Conduct for the Master thesis.

Abstract

Direct data-driven control refers to the methodology of using data from an unknown dynamical system to design controllers directly, with the goal of enforcing a desired system behavior. The collected data consists of the input signals used for excitation and the output signals that are measured from the system as a response to the input. The measured data is organized into an artificial state, termed as the extended state, and consists of lagged input and output samples. This extended state encapsulates the system's behavior purely from the data, removing the need for an explicit model. This thesis presents a systematic control design framework based on the extended state to be utilized for direct data-driven control.

To ensure the practical applicability of the extended-state framework towards control of dynamical systems, this work proposes a novel method for order reduction of the extended state. The reduced-order extended state is developed by building upon existing theoretical results, improving the efficacy of the case of dynamic output-feedback controller design. Additionally, this study explicitly addresses the challenge of measurement noise, which is an inevitable factor in real-world systems. An Instrumental Variable-based approach is introduced to mitigate the effects of noise on the control process. This method is further extended to establish conditions for controller synthesis to ensure both system stabilization and performance optimization. The proposed framework is validated through extensive simulation studies, as well as through experimental implementation on a real-world system, demonstrating its robustness and practicality.

Preface

This thesis is submitted in partial fulfillment of the requirements for the degree Master of Science (MSc) in Systems & Control, and is the final product of my research conducted for the graduation project. This thesis is written in such a way that it encompasses all information about the data-driven methodology using an extended state, and I hope that it will be used as a reference in the future.

Before embarking on this project, I sought an assignment that would allow me to explore in depth two foundational ambitions I wanted for myself: conducting theoretical research and ensuring its relevance to real-world systems. This report reflects that aspiration, outlining the research process while addressing the practical implications and challenges encountered along the way.

The completion of this journey has been made possible by the incredible support from my supervisors, to whom I owe a deep sense of gratitude. Roland, with his extensive knowledge of both theoretical and practical aspects, helped me stay aligned with my vision for the project. Valentina provided me with constant guidance, and our insightful discussions in her office greatly enriched my understanding of the research process. Moreover, I cannot thank Luuk enough for his unwavering support throughout this journey; our (often endless) conversations always helped me find the right direction. I am truly grateful for your collective encouragement and constructive feedback.

I would also like to thank my committee members, Mircea Lazar and Erjen Lefeber, for graciously agreeing to be part of this committee. Your willingness to share your expertise and insights has significantly enriched this project. I am particularly grateful to Mircea for his assistance with the programming script, which enabled me to overcome my research block.

Additionally, I would also like to thank my MSc colleagues for the interesting discussions and mutual support on a daily basis throughout these two years. Lastly, I want to express my heartfelt thanks and gratitude to my parents for their unconditional support and encouragement.

Eashwar Ravikkumar, Eindhoven - 22nd October 2024

Contents

Chapter 1

Introduction

The process of designing a control system traditionally follows a two-step procedure: first, the development of a mathematical model of the system using system identification techniques, and then designing model-based control strategies based on the developed model [1]. Over the past decades, substantial advancements have been made in the field of system identification towards utilizing measured data to accurately develop a general parametric model that characterizes the behavior of the dynamical system. These obtained models are suitable for model-based controller design, such as pole-zero placement, Linear Quadratic Regulator (LQR), Lyapunov-based controller, and feedback linearization. This particular strategy of identifying a model from the data followed by a model-based control design strategy is often called indirect data-driven control [2].

The first step of system identification is a data-driven methodology, and a primary requirement for any data-driven procedure is that the data should be sufficiently informative. This means that the data collected from a system has sufficient information about the entire system dynamics. Data informativity is enforced in experiment design by considering an input signal that sufficiently excites the system such that the output response holds adequate knowledge of the system dynamics [3]. This places a stringent requirement on how to use the measured data for system identification. However, it has been observed that even if the data is not informative for identification, it can still be informative for controller design [4]. Also, in many practical scenarios, accurately identifying systems can be challenging or even infeasible due to system complexity, high-dimensionality, or presence of non-linearities [5].

To address these challenges, direct data-driven control has emerged as a one-step methodology that bypasses the need for explicit system identification [6]. This approach circumvents the necessity of modeling and attempts to eliminate the trade-off between informativity and complexity. The measured system data is directly utilized to design controllers that enforce the desired system behavior, making this approach highly relevant in situations where obtaining a reliable model is impractical or impossible. The key challenge is the development of simple and practically implementable *direct data-driven* (DDD) control techniques using data points that may be affected by noise. Furthermore, these methods should allow for a systematic analysis and development of guarantees such as closed-loop stability and performance. The existing state-of-the-art for the realization of these desired requirements are discussed in the following section.

1.1 State-of-the-art

A cornerstone result in the field of direct data-driven control was developed by Willems [7] in the context of behavioral systems theory, and is now called the *Fundamental Lemma* [8]. This result states that one or multiple sufficiently exciting trajectories of a discrete-time (DT) linear time-invariant (LTI) system can be used to parameterize all the trajectories that can be generated by the system [9, 10]. The Fundamental Lemma hence characterizes the informativity of the measured data, i.e., whether a finite collection of data can completely represent the LTI system. It has been utilized for the development of various methods of direct data-driven control [11]. These methods include feedback stabilization, optimal and robust control and predictive control [3, 12, 13, 14]. All these involve control strategies being directly obtained from the measured data with minimal prior knowledge about the model.

The DDD control strategies leverage data from the system, which can be of different types - input-state, input-state-output, and input-output data. There exist extensive results based on input-state data-driven control [3, 12] based on the assumption that the state of the system is available for measurement. However, in most real-life applications, only input-output data is available for analysis and controller synthesis. One common strategy to handle input-output data is to create a non-minimal extended state vector consisting of shifted input-output measurements [12]. Therefore, a corresponding non-minimal state-space representation is obtained, which can be used for data-driven control. Defining a state helps to establish stability and performance guarantees based on the comprehensive research about state-space representations [15]. Additionally, the well-defined methodologies for input-state DDD control can be directly utilized for input-output DDD control through the definition of extended state [16].

Nonetheless, there are certain limitations to using the extended state for control. The constructed extended state-space representation may not necessarily be controllable, especially when multiple outputs are considered [17]. The current methodology also requires prior knowledge of the order of the system. The extended state, therefore, requires modifications to be directly utilized alongside existing input-state control methodologies for these reasons, and this acts as the first key research direction of this work.

The methods of DDD control discussed here are based on the assumption of noise-free data, meaning that the measured data is not influenced by noise or disturbances. However, this assumption is theoretical, as real systems are always exposed to noise. This noise can be viewed as measurement noise arising from sensor errors. To gain a better understanding of both the deterministic aspects of the system and the stochastic noise components, a stochastic model class such as an Output-Error (OE) or Box-Jenkins (BJ) model structure can be defined and used for analysis. Additionally, the presence of noise can compromise stability guarantees and may introduce unwanted measurement biases in the system. Measurement noise has been considered for the DDD control using input-state data [18] by assuming prior knowledge in terms of the energy bounds of the noise samples. Input-output data is considered in [19] with prior knowledge of noise assumed as a sample cross-covariance type bound concerning a user-chosen instrumental signal. Additional assumptions of various types of noise bounds for data-driven control have been discussed in [20, 21, 22]. A pivotal point of research is to establish a general method for handling noise, which translates to practical applicability and, therefore, is the second key research direction of this thesis.

These ideas of ensuring stability using DDD control methodologies can be extended toward performance optimization as well. It is crucial to ensure robust performance through the designed DDD controller, which can be defined as achieving a desired closed-loop performance bound in the presence of noise. A DDD controller synthesis methodology using input-state is discussed in [18], where the closed-loop system achieves a desired $\mathcal{H}_\infty$ or $\mathcal{H}_2$ performance bound [15]. This has been extended to the case of noisy input-output data in [19] using the constructed non-minimal extended state. However, this approach also encounters the discussed limitations on the extended state, as discussed before. Moreover, there are no existing methodologies for the case of noise-free input-output data. These research gaps establish the third key research direction of this thesis.

1.2 Problem Statement

The extended state framework is proposed as a methodology for utilizing the non-minimal extended state constructed from collected input-output data for DDD stabilization and performance control. The central objective of this thesis is to develop the extended state framework towards ensuring data-based closed-loop stabilization and performance guarantees while addressing the three research directions explored in the existing state-of-the-art. This drives the formulation of the following three *research questions* (RQs).

RQ1) How can a non-minimal extended state based on collected input-output data be constructed so as to render input-state data-driven control techniques feasible with minimal prior knowledge of the system?

RQ2) How to ensure minimum stability and performance loss during direct data-driven control using an extended state while considering general noise conditions?

RQ3) How to develop tractable conditions for data-driven control using an extended state to ensure performance optimization?

1.3 Outline

The content of the chapters in this report is specified in the overview. Chapter 2 introduces the concept of an extended state, which will be used for the DDD control of a MIMO (multi-input multi-output) system. The proposed methodology for answering **RQ 1** is discussed in Chapter 2.3, along with the complete addressal of the existing limitations, directing towards the development of the extended state framework for control using noise-free data. Also, Chapter 2.5 provides a novel approach for performance optimization using noise-free data and partially answers **RQ 3**. Chapter 3 continues the discussion by establishing the framework for data-driven control in the presence of external noise. A first attempt at handling noise efficiently and in a closed-loop setting is discussed in Chapter 3.4, therefore answering **RQ 2**. Furthermore, suitable modifications to existing methodologies of data-driven performance for noisy data are proposed to completely answer **RQ 3**. Chapter 4 contains the conclusions based on the research questions and suggests future work.

1.4 Notation

The set of real numbers, integers, and non-negative integers are denoted by $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$, respectively. A sequence of samples $\begin{bmatrix} z(0) & z(1) & \cdots & z(N-1) \end{bmatrix}$ of length $N \in \mathbb{N}$ such that $z(k) \in \mathbb{R}^{n_z}$ is represented as $z_{[0,N-1]}$. A column vector representation of $z_{[0,N-1]}$ is defined as $\mathrm{col}(z(0),\ z(1),\ \cdots,\ z(N-1))$. The notation $Z^\top$ and $Z^\dagger$ denote the transpose and left inverse of a matrix Z, respectively. A quadratic term is written as $zA(*)^\top$, where $(*)$ denotes a symmetric term, i.e., $zA(*)^\top = zAz^\top$. The notations $A \succ 0, A \succeq 0, A \prec 0$, and $A \preceq 0$ denote that the symmetric matrix A is positive definite, positive semi-definite, negative definite, and negative semi-definite, respectively. Given a discrete-time signal $z \in (\mathbb{R}^{n_z})^{\mathbb{Z}}$, where $(\mathbb{R}^{n_z})^{\mathbb{Z}}$ defines the collection of all maps $z : \mathbb{Z} \to \mathbb{R}^{n_z}$, the restriction to the interval $[0, L-1] \cap \mathbb{Z}$ is denoted by $z|_L$. The Hankel matrix of depth L associated with a sequence of samples $z_{[0,N-1]}$ is defined as

$$\mathcal{H}_L(z) = \begin{bmatrix} z(0) & z(1) & \cdots & z(T-L) \\ z(1) & z(2) & \cdots & z(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ z(L-1) & z(L) & \cdots & z(T-1) \end{bmatrix}. \tag{1.1}$$

Lastly, $A \otimes B$ denotes the Kronecker product of two matrices A and B.

Chapter 2

Data-Driven Control Using Noise-free Data

The previous chapter mentioned that this thesis focuses on constructing and utilizing an extended state framework for the data-driven control of a MIMO system. The data-based framework necessitates an analysis of the system information present in the collected data. This chapter defines the data-driven system representation and establishes the necessary conditions for data informativity.

Data informativity conditions and data-based control methodology for noise-free input-state data are introduced. A maximal extended state is introduced to extend these approaches towards noise-free input-output data. To address the limitations of the maximal extended state, a novel state reduction technique is proposed for the constructed maximal extended state used in the data-driven system representation. Subsequently, it is shown that the reduction technique renders established input-state data-based controller synthesis techniques suitable for input-output data. Conditions for controller synthesis and a first-approach for controller performance are discussed and validated through simulation examples.

2.1 System Description and Data-Driven Representation

This section describes the assumed class of data-generating systems for which the data-driven methodologies are being developed.

2.1.1 System Description

Consider a DT-LTI system $\mathcal{G}$ with multiple inputs and multiple outputs defined in terms of the minimal state-space representation

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k) + Du(k), \end{aligned} \tag{2.1}$$

with $y(k) \in \mathbb{R}^p$ denoting the output, $u(k) \in \mathbb{R}^m$ denoting the input, and $x(k) \in \mathbb{R}^n$ denoting the state. The system $\mathcal{G}$ is only used for data generation, and thus, only the measured data signals are utilized for data-driven analysis and control.

Let the input-output behavior of $\mathcal{G}$ (denoted by $\mathfrak{B}$) be defined as the collection of all the possible trajectories that are admissible solutions of the dynamics of $\mathcal{G}$ such that

$$\mathfrak{B} := \left\{ (u,y) \in (\mathbb{R}^m \times \mathbb{R}^p)^{\mathbb{N}} \,\middle|\, \exists x \in (\mathbb{R}^n)^{\mathbb{N}} \text{ s.t. (2.1) holds } \forall k \in \mathbb{N} \right\}. \tag{2.2}$$

The state $x(k)$ in (2.1) can be eliminated using the Elimination Theorem [23] to get an input-output or kernel representation [24] of the system as

$$\mathfrak{B} = \left\{ (u,y) \in (\mathbb{R}^m \times \mathbb{R}^p)^{\mathbb{N}} \,\middle|\, P(q^{-1}) \begin{bmatrix} u \\ y \end{bmatrix} = 0 \right\}. \tag{2.3}$$

Here, q^{-1} denotes the backward shift operator, i.e., $q^{-1}y(k) = y(k-1)$, $P(\xi) \in \mathbb{R}^{p\times(m+p)}[\xi]$ denotes a polynomial with matrix-valued coefficients with an indeterminate ξ used to represent the polynomial structure, and $P(q^{-1})$ is a time-domain operator. Because the input-output partitioning is assumed to be known, the kernel-based representation can be partitioned into $P(q^{-1}) = \begin{bmatrix} B(q^{-1}) & -A(q^{-1}) \end{bmatrix}$ such that $\mathtt{A}(q^{-1})y(k) = \mathtt{B}(q^{-1})u(k)$ is a minimal input-output representation, i.e., $\mathtt{A}(q^{-1})$ and $\mathtt{B}(q^{-1})$ are co-prime. Here, $\mathtt{A}(q^{-1})$ is a monic polynomial while $\mathtt{B}(q^{-1})$ is a regular polynomial, such that

$$\begin{aligned} \mathtt{A}(q^{-1}) &= I + \mathtt{A}_1 q^{-1} + \mathtt{A}_2 q^{-2} + \cdots + \mathtt{A}_\ell q^{-\ell}, \\ \mathtt{B}(q^{-1}) &= \mathtt{B}_0 + \mathtt{B}_1 q^{-1} + \mathtt{B}_2 q^{-2} + \cdots + \mathtt{B}_\ell q^{-\ell}, \end{aligned} \tag{2.4}$$

with the polynomial matrices having dimensions $\mathtt{A}_i \in \mathbb{R}^{p\times p}$ $\forall i \in [1, 2, \cdots, \ell]$ and $\mathtt{B}_i \in \mathbb{R}^{p\times m}$ $\forall i \in [1, 2, \cdots, \ell]$, where ℓ is the lag [23] of the system.

Definition 2.1 (System lag or Observability index [25])**.** The lag (or observability index) (ℓ) of the system $\mathcal{G}$ (2.1) is defined as the minimal value of r for which

$$\operatorname{rank} \mathscr{O}(r) = n,$$

with $\mathscr{O}(r)$ being the r-observability matrix $\mathscr{O} = \begin{bmatrix} C^\top & \ldots & (CA^{r-1})^\top \end{bmatrix}$. □

Therefore, the model class of the system $\mathcal{G}$ has its complexity [26] defined by (i) the order of the system n, (ii) the number of inputs m, and (iii) the lag of the system ℓ, such that $\ell \leq n \leq p\ell$. This inequality is important when characterizing the finite-horizon data-based system behavior [23].

2.1.2 Data-based Representation

Having defined a model-based system representation, a finite-horizon data-based non-parametric representation can be formulated. The system data of the unknown system $\mathcal{G}$ refers to the data collected from the measurement of signals, specifically the input $(u(k))$ and the output $(y(k))$ sequences. A measured input-output data set is denoted as $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N-1}$, with N representing the total number of collected samples. Assuming that the state is accessible, the state data can also be measured, resulting in a measured input-state data set $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^{N-1}$. It should be noted that apart from the case of defining the existing data-driven input-state methodologies, this thesis considers that only the input-output trajectory is measurable, and hence, the data dictionary $\mathcal{D}_N$ is considered for analysis and control.

Data informativity represents the notion of the collected data containing sufficient information about the system dynamics to control the system using data-driven methods. In general, a primary condition for data informativity is that the input fed to the system for data collection is sufficiently exciting. The notion of persistency of excitation formalizes this.

Definition 2.2 (Persistence of Excitation [24]). A signal sequence $\{u_k\}_{k=0}^{N-1} \in \mathbb{R}^m$ is said to be persistently exciting of order L if its Hankel matrix $\mathcal{H}_L(u)$ has full row rank. □

Considering that such a persistently exciting input sequence is used to gather output data from the system $\mathcal{G}$ to obtain the data set $\mathcal{D}_N$, a data-based representation for the finite-horizon L can be formalized based on the *Fundamental Lemma* stated as follows.

Lemma 2.1 (Fundamental Lemma [7]). *Given a measured input-output trajectory $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N-1}$ of an LTI system $\mathcal{G}$ with the input sequence $\{u_k^d\}_{k=0}^{N-1}$ being persistently exciting of order $L+n$, then $\mathcal{D}_N$ is a data trajectory of $\mathcal{G}$ if and only if there exists $g \in \mathbb{R}^{N-L+1}$ such that*

$$\begin{bmatrix} \mathcal{H}_L(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix} g = \begin{bmatrix} u|_L \\ y|_L \end{bmatrix}, \tag{2.5}$$

holds.

Proof. See [24] and the references therein for the proof. □

Furthermore, it is known [24] that the *Fundamental Lemma* can be generalized so that requirements of controllability or persistence of excitation are not required. Consider the finite-horizon manifest behavior of the system $\mathcal{G}$

$$\mathfrak{B}|_L^{\mathcal{G}} := \left\{ (u_p, y_p) \in (\mathbb{R}^m \times \mathbb{R}^p)^L \mid \exists (u_f, y_f) \in (\mathbb{R}^m \times \mathbb{R}^p)^{\mathbb{N}} : (u_p, y_p) \wedge (u_f, y_f) \in \mathfrak{B} \right\}. \tag{2.6}$$

The finite-horizon manifest behavior $\mathfrak{B}|_L^{\mathcal{G}}$ is equal to the image of $\begin{bmatrix} \mathcal{H}_L(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix}$ if and only if the rank condition

$$\operatorname{rank}\left(\begin{bmatrix} \mathcal{H}_L(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix} \right) = mL + n, \tag{2.7}$$

is satisfied for $L \geq \ell$. This implies the necessary condition that a data-driven non-parametric representation of the system can be formulated using input-output Hankel matrices only when the considered length of the finite-horizon (L), which is also taken as the depth of the Hankel matrices, is greater than or equal to the lag (ℓ) of the system. This is an important condition to ensure that the data encapsulates the system behavior. Throughout the rest of this thesis, it is considered that $L \geq \ell$ unless otherwise specified. Practically, it can be considered that an upperbound on the system lag (ℓ^u) is known, so that the depth can be selected as $L = \ell^u$.

It is also known that if the state data is accessible for measurement, then the *Fundamental Lemma* can be used to define the data-informativity conditions for the state-space representation (2.1).

Lemma 2.2 (Input-State Fundamental Lemma [7])**.** *Given a measured input-state data trajectory $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^{N-1}$ of an LTI system $\mathcal{G}$ with the input sequence $\{u_k^d\}_{k=0}^{N-1}$ being persistently exciting of order $n+1$, then $\mathcal{D}_{NS}$ is a input-state data trajectory of $\mathcal{G}$ if and only if the following rank condition*

$$\operatorname{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d)\\ \mathcal{H}_1(x^d)\end{bmatrix}\right) = m+n. \tag{2.8}$$

holds.

This rank condition ensures that the data encodes all information necessary for control and has been used in [12] to develop a methodology for state-feedback controllers. Before exploring the development of control strategies for input-output data, an overview of the existing strategies based on collected input-state data is presented.

2.1.3 Data-Driven Control using Input-State Data

Consider the state space system representation (2.1). It can be equivalently represented using data by the following result.

Theorem 2.1 (Input-State Data-based System Representation [12])**.** *Consider the system (2.1) and let the data dictionary $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^{N}$ be a measured input-state trajectory of the system. Furthermore, let the condition (2.8) hold. Then, the system has the following finite-horizon data-based representation*

$$x(k+1) = \mathcal{H}_1(x^d)_+ \begin{bmatrix}\mathcal{H}_1(u^d)\\ \mathcal{H}_1(x^d)\end{bmatrix}^\dagger \begin{bmatrix}u(k)\\ x(k)\end{bmatrix}, \tag{2.9}$$

where $\mathcal{H}_1(x^d)_+ = x^d_{[1,N]}$.

Proof. See Theorem 1 from [12] for the proof. □

Note that an extra input-state sample in the data dictionary $\mathcal{D}_{NS}$ (representing the current input and state) is collected to ensure the creation of the state data matrix $\mathcal{H}_1(x^d)_+$. For this data-based open-loop system, a state-feedback controller can be designed which ensures asymptotic stability. The notion of (asymptotic) quadratic stabilization is defined as follows.

Definition 2.3. Consider the system (2.1). The system can be stabilized by the state-feedback control law $u(k) = Kx(k)$ if there exists a K and a symmetric $P \succ 0$ such that for any state and input system matrices A and B, the condition

$$(A+BK)P(A+BK)^\top - P \prec 0, \tag{2.10}$$

holds. □

Using this definition, the stabilizing state-feedback controller can be designed with the following result.

Theorem 2.2 (Stabilization for noise-free Input-State Data [12])**.** *Consider the system* (2.1) *and let the data dictionary* $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^N$ *be a measured input-state trajectory of the system. Furthermore, let the condition* (2.8) *hold. Then the following statements hold.*

1. *The system* (2.1) *in a closed-loop with a state feedback* $u = Kx$ *has the following finite-horizon data-based representation*
$$x(k+1) = \mathcal{H}_1(x^d)_+ G_K x(k)$$
with G_K *satisfying*
$$\begin{bmatrix} K \\ I_n \end{bmatrix} = \begin{bmatrix} \mathcal{H}_1(u^d) \\ \mathcal{H}_1(x^d) \end{bmatrix} G_K,$$
and
$$u(k) = \mathcal{H}_1(u^d) G_K x(k).$$

2. *Any matrix* $\mathcal{Q}$ *satisfying*
$$\begin{bmatrix} \mathcal{H}_1(x^d)\mathcal{Q} & \mathcal{H}_1(x^d)_+\mathcal{Q} \\ \mathcal{Q}^\top \mathcal{H}_1(x^d)_+^\top & \mathcal{H}_1(x^d)\mathcal{Q} \end{bmatrix} \succ 0,$$
such that
$$K = \mathcal{H}_1(u^d)\mathcal{Q}(\mathcal{H}_1(x^d)\mathcal{Q})^{-1},$$
stabilizes the system (2.1).

Proof. See Theorems 2 and 3 from [12] for the proof. □

Theorem 2.2 gives the results for designing a stabilizing state-feedback controller using the finite-horizon data-based system representation created from the collected input-state data. However, this methodology is applicable only when noise-free state data is accessible for measurement, either through direct state measurement or through output measurement with full-state feedback and a unity output matrix ($C = I$). As discussed, this is often not the case, and only input-output data can be measured from the system in practice. This motivates a methodology that allows for control of a system based on the collected input-output data and which possibly utilizes the existing results for input-state data.

2.2 Maximal Extended State

An extended state is an artificially constructed state consisting of lagged input and output samples leading to the creation of an augmented input-output state-space system to which techniques of control using input-state data (e.g., see Section 2.1.3) are applicable. From

[12], the collected input-output data can be used to construct a non-minimal extended state $\chi(k) \in \mathbb{R}^{(m+p)L}$ as

$$\begin{aligned}\chi(k) = \operatorname{col}(y^d(k-1),\ y^d(k-2),\ \cdots,\ y^d(k-L),\\ \cdots u^d(k-1),\ u^d(k-2),\ \cdots,\ u^d(k-L)).\end{aligned} \tag{2.11}$$

The corresponding augmented state-space representation, henceforth called an extended state-space representation, is given by

$$\begin{aligned}\chi(k+1) &= \mathcal{A}\chi(k) + \mathcal{B}u(k),\\ y(k) &= \begin{bmatrix}\bar{A} & \bar{B}\end{bmatrix}\chi(k) + \mathtt{B}_0 u(k),\end{aligned} \tag{2.12}$$

with

$$\mathcal{A} = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\bar{A}} & \multicolumn{2}{c}{\bar{B}} \\ \hline I_{p(L-1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_{m(L-1)} & 0 \end{array}\right], \quad \mathcal{B} = \left[\begin{array}{c} B_0 \\ \hline 0 \\ I_m \\ 0 \end{array}\right]. \tag{2.13}$$

Here, $\bar{A}$ and $\bar{B}$ are the matrix coefficients of polynomials $A(q^{-1})$ and $B(q^{-1})$ such that

$$\begin{aligned}\bar{A} &= \begin{bmatrix}-\mathtt{A}_1 & -\mathtt{A}_2 & \cdots & -\mathtt{A}_{L-1} & -\mathtt{A}_L\end{bmatrix},\\ \bar{B} &= \begin{bmatrix}\mathtt{B}_1 & \mathtt{B}_2 & \cdots & \mathtt{B}_{L-1} & \mathtt{B}_L\end{bmatrix}.\end{aligned} \tag{2.14}$$

The primary advantage of such an extended state is that it can be directly constructed from the input-output measurements of the system $\mathcal{G}$. Considering the initial condition $\chi_0 \in \mathfrak{B}|_L^{\mathcal{G}}$, with $L \geq \ell$, the finite-horizon manifest behavior of the extended state-space representation (2.12) is identical to the finite-horizon manifest behavior of the original system (2.1). Furthermore, the input-state rank condition (2.8) (as described in Lemma 2.2) can be reformulated for the extended state (2.11) as

$$\operatorname{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d)\\ \mathcal{H}_1(\chi_L^d)\end{bmatrix}\right) = m(L+1) + pL. \tag{2.15}$$

This essentially implies that if the condition (2.15) is satisfied, then a static state-feedback controller synthesized for the extended state-space realization (2.12) can be equivalently realized as a dynamic output-feedback controller for the original system $\mathcal{G}$. However, considering the *Fundamental Lemma*, it is visible that this rank condition (2.15) is only satisfied when $p\ell = n$ with $L = \ell$, as highlighted below.

Theorem 2.3 (Unsatisfied Rank Condition)**.** *Consider the system* (2.12)*, and let the data dictionary* $\mathcal{D}_{NX} = \{u^d, \chi^d\}_{k=0}^{N-1}$ *be a constructed input-extended-state data set of the system based on the measured input-output data of the original data generating system* (2.1)*, with the input sequence* $\{u_k^d\}_{k=0}^{N-1}$ *being persistently exciting of order* $n + L$*. The matrix* $\begin{bmatrix}\mathcal{H}_1(u^d)\\ \mathcal{H}_1(\chi^d)\end{bmatrix}$ *satisfies the rank condition* (2.15) *if and only if* n *satisfies* $p\ell = n$ *and* L *is chosen such that* $L = \ell$.

Proof. By the definition of the extended state (2.11)

$$\begin{bmatrix} \mathcal{H}_L(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix} \equiv \mathcal{H}_1(\chi_L^d).$$

From the *Fundamental Lemma* and specifically utilizing the input-output rank condition (2.7), it is known that under the condition of a persistently exciting input sequence

$$\operatorname{rank}\left(\begin{bmatrix} \mathcal{H}_{L+1}(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} \mathcal{H}_1(u^d) \\ \mathcal{H}_1(\chi_L^d) \end{bmatrix}\right) = m(L+1) + n. \tag{2.16}$$

with n being the true order of the system $\mathcal{G}$. To ensure that the matrix $\begin{bmatrix} \mathcal{H}_1(u^d) \\ \mathcal{H}_1(\chi^d) \end{bmatrix}$ satisfies the rank condition (2.15), it is required that $pL = n$, or specifically, $L = \frac{n}{p}$.

It is known that $L \geq \ell$ to ensure that the collected data is informative. Furthermore, it is also known that $p\ell \geq n \geq \ell$. Therefore, the condition (2.15) is only satisfied when $p\ell = n$ and $L = \ell$. □

Theorem 2.3 effectively shows that the maximal extended state can only be utilized to control systems with $p\ell = n$ and $L = \ell$). These cases include single-input-single-output (SISO) systems (with $L = \ell$) and MIMO systems with $p\ell = n$ and $L = \ell = 1$. The above theorem also shows that there is no equivalent condition for the data informativity of constructed input-maximal-extended-state data as compared with the data informativity condition (2.8) on input-state data. So, it is not possible to completely establish the informativity of the input-maximal-extended-state data. Therefore, the extended state space (2.12) is uncontrollable, and hence, the maximally extended state cannot be used for data-driven control (apart from the specified cases). This is a current limitation of the existing methodology.

Furthermore, the lag (ℓ) and the order (n) of a real system are not known in practice. This necessitates the consideration of the depth L as large as possible. But, such a consideration of L is directly restricted by Theorem 2.3. Thereby, from Theorem 2.3, it is not possible to directly utilize the extended state for practical applications.

The discussed points are highlighted in the following example.

Example 2.1. Consider a randomly generated DT LTI system of the form (2.1), which is regarded as the data-generating system. The random system has complexity $(n, m, \ell) = (20, 2, 7)$ and $p = 3$, and hence it is a MIMO system. The data dictionary $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N-1}$ with $N = 1000$ is used to define the finite-horizon Hankel data matrices (both the input-output and input-extended state data matrices). The rank of the actual data matrices and their respective desired ranks are shown in Figure 2.1 for the depths $L = \{1, 2, \cdots, 15\}$.

It can be observed that the rank condition on the input-output data (2.7) holds when the depth is greater than or equal to the lag of the system ($L \geq \ell$). Therefore, the collected input-output data is always informative. However, $p\ell \neq n$, as $p\ell = 21$ and $n = 20$. It can also be seen that the rank condition on the input-extended state data (2.15) does not hold for $L = \ell = 7$, validating Theorem 2.3. □

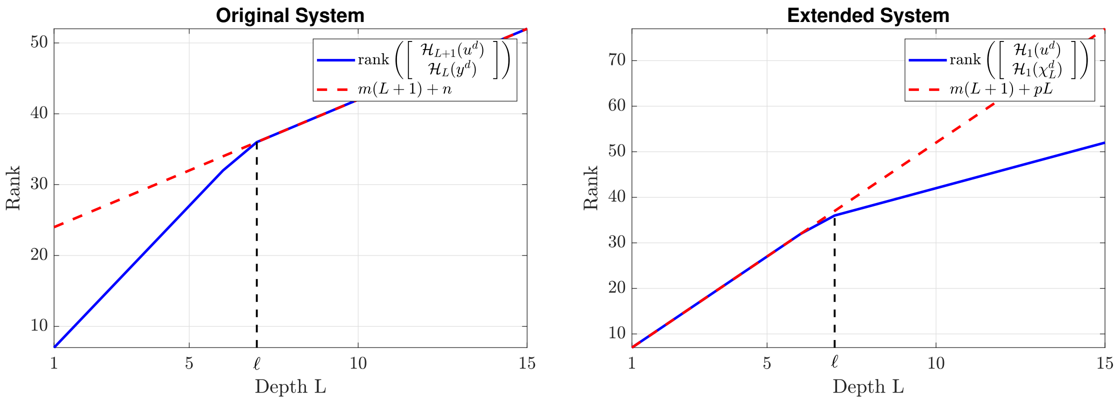


Figure 2.1: Comparision of numerical and mathematical rank conditions

The example shows that the data dictionary $\mathcal{D}_{NX}$ collected from the system (2.12) will not represent the finite-horizon system behavior $\mathfrak{B}|_L^{\mathcal{G}}$. Hence, as discussed, Theorem 2.3 effectively means that the maximal extended state, as currently defined, cannot be utilized for data-driven control for MIMO systems. Consequently, the existing methodology has to be modified to ensure its applicability. The proposed methodology for handling the mentioned limitations is described as extended state reduction.

Remark 2.1. It should be noted that to correctly define the extended state data dictionary $\mathcal{D}_{NX}$, it is essential to define the initial conditions χ_0 of the extended state. To keep the notation constant, it is henceforth considered that $N+L$ data samples are measured, with the first L data samples defining the initial conditions and the rest N data samples utilized for control. This results in the precise *required* rank condition

$$\operatorname{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d_{[L,N+L-1]})\\ \mathcal{H}_1(\chi^d_{[L,N+L-1]})\end{bmatrix}\right) = m(L+1)+n. \tag{2.17}$$

2.3 Extended State Reduction

As discussed in the previous section, the extended state should satisfy (2.17) to satisfy data informativity conditions according to the *Fundamental Lemma*. One method to ensure this is by a technique called extended state reduction [17]. State reduction is a technique that allows the creation of a non-minimal extended state with a reduced-order (as compared with the maximal extended state). This reduced-order extended state χ_r needs to be designed such that $\chi_r = \Theta\chi$, where $\Theta \in \mathbb{R}^{mL+n\times(m+p)L}$ is a transformation matrix guaranteeing that

$$\operatorname{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d_{[L,N+L-1]})\\ \mathcal{H}_1(\chi^d_{r[L,N+L-1]})\end{bmatrix}\right) = m(L+1)+n, \tag{2.18}$$

holds. Satisfaction of the above rank condition ensures that the collected data completely represents the system behavior, which is desired and essential for the development of the extended state framework for control, as discussed in the previous section.

2.3.1 Creating a Reduced-Order State

Having defined the requirements for the reduced state, it is now necessary to create a suitable reduced state (or more accurately, a transformation matrix Θ) such that (2.18) is satisfied. Consider the measured input-output data dictionary $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N+L-1}$. A numerical method has to be developed now to ensure the following operations

$$\begin{bmatrix} \mathcal{H}_{L+1}(u^d_{[0,N+L-1]}) \\ \mathcal{H}_L(y^d_{[0,N+L-2]}) \end{bmatrix} \xrightarrow[\text{Construction}]{\text{Extended State}} \begin{bmatrix} \mathcal{H}_1(u^d_{[L,N+L-1]}) \\ \mathcal{H}_1(\chi^d_{[L,N+L-1]}) \end{bmatrix} \xrightarrow{\text{State Reduction}} \begin{bmatrix} \mathcal{H}_1(u^d_{[L,N+L-1]}) \\ \mathcal{H}_1(\chi^d_{r[L,N+L-1]}) \end{bmatrix}. \tag{2.19}$$

The first step consists of the rearrangement of the input-output Hankel data matrix resulting in the maximal extended state data matrix. This can be performed by a simple transformation

$$\begin{bmatrix} \mathcal{H}_1(u^d_{[L,N+L-1]}) \\ \mathcal{H}_1(\chi^d) \end{bmatrix} = \Pi_1 \begin{bmatrix} \mathcal{H}_{L+1}(u^d_{[0,N+L-1]}) \\ \mathcal{H}_L(y^d_{[0,N+L-2]}) \end{bmatrix}, \tag{2.20}$$

with $\Pi_1 \in \mathbb{R}^{(m(L+1)+pL)\times(m(L+1)+pL)}$ being a binary permutation matrix directly computable from the Hankel data matrices. Based on the definition of the extended state (2.11), it is calculated as

$$\Pi_1 = \begin{bmatrix} 0_{m\times mL} & I_m & 0_{m\times pL} \\ 0_{pL\times mL} & 0_{pL\times m} & I_{pL} \\ I_{mL} & 0_{mL\times m} & 0_{mL\times pL} \end{bmatrix}. \tag{2.21}$$

The second step of the conversion $\mathcal{H}_1(\chi^d) \xrightarrow{\text{State Reduction}} \mathcal{H}_1(\chi_r^d)$ is the more important step in the process of state reduction. It can be thought of as a rearrangement of the data matrix $\mathcal{H}_1(\chi^d)$ using Gaussian elimination to eliminate all the data rows that are unobservable (and hence linearly dependent) [17]. This can be carried out using different possible algorithms.

2.3.2 State Reduction Algorithms

The aim of state reduction is to reduce the maximal extended state such that the reduced state results in a data matrix satisfying (2.18). This subsection proposes two algorithms for achieving state reduction. A crucial point to note here is that both algorithms utilize only the data and do not require any prior knowledge about the system, specifically about the order n.

Algorithm 1 performs a QR decomposition on the maximal extended state data matrix $(\mathcal{H}_1(\chi^d_{[L,N+L-1]}))$ and takes advantage of the fact that it should have full row rank. Knowing this allows for the numerical selection of the desired states to create the reduced state data matrix.

Algorithm 1 State Reduction - QR Decomposition [27]

Require: Maximal extended state data Hankel data matrix $\mathcal{H}_1(\chi^d)$, tolerance value `tol = 1e-13`

1. Perform QR decomposition $\mathcal{H}_1(\chi^d) = Q_H R_H$.
2. Select rows of R_H with diagonal values greater than tolerance and note the corresponding indices in I_H^n with $n_{red} = \text{length}\ (I_H^n)$.
3. Create transformation matrix $\Theta \in \mathbb{R}^{n_{red}\times(m+p)L}$ to arrange data matrix $\mathcal{H}_1(\chi^d)$ using the indices I_H^n, such that $\mathcal{H}_1(\chi_r^d) = \Theta\mathcal{H}_1(\chi^d)$.

return Reduced-order data matrix $\mathcal{H}_1(\chi_r^d)$, transformation matrix Θ

Remark 2.2. Note that the QR decomposition is performed in `MATLAB` with `econ` and `vector` parameters enabled such that an economy-size decomposition is performed. This enables the decomposition according to the specified tolerance and allows the definition of the indices matrix I_H^n. Also, R_H is obtained as an upper diagonal matrix arranged in descending order of absolute magnitude of elements.

Algorithm 1 creates the required transformation matrix Θ, which allows the creation of a reduced state such that (2.18) is satisfied. However, even though QR decomposition is efficient, there are certain drawbacks. There are no insights to be gained about the reduced state from the QR decomposition, for example, as compared with the singular values obtained from a Singular Value Decomposition (SVD). Furthermore, the algorithm does not allow for any noise or disturbance to be accounted for due to the numerical dependence on the tolerance value.

This motivates the development of an algorithm based on the SVD, which is widely applicable for rank-deficient and ill-conditioned matrices. Again, it is highly suitable for rank-reduction and the singular values of the data matrix provide insights about the data representing the system. This means that it can also be extended to the case when noise is present. Algorithm 2 performs an SVD on the extended state data matrix and calculates the number of states from the data. Specifically, the order of the system can be directly calculated from the data as

$$n_{sys} = \text{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d_{[L,N+L-1]}) \\ \mathcal{H}_1(\chi^d)\end{bmatrix}\right) - m(L+1). \tag{2.22}$$

Notice that by definition, $n_{sys} = n$, and hence the correct order is always obtained.

Algorithm 2 State Reduction - SVD

Require: Maximal extended state data Hankel data matrix $\mathcal{H}_1(\chi^d)$

1. Perform a singular value decomposition (SVD) $\mathcal{H}_1(\chi^d) = U_f S_f V_f^\top$.
2. Calculate system order as $n_{sys} = \text{rank}\left(\begin{bmatrix} \mathcal{H}_1(u^d_{[L,N+L-1]}) \\ \mathcal{H}_1(\chi^d) \end{bmatrix}\right) - m(L+1)$.
3. Select reduced-order as $n_{red} = mL + n_{sys}$.
4. Take $U_r = U_f(:, 1 : n_{red})$ and $S_r = S_f(1 : n_{red}, 1 : n_{red})$.
5. Create transformation matrix $\Theta = S_r^{-\frac{1}{2}} U_r \in \mathbb{R}^{n_{red}\times(m+p)L}$ to arrange data matrix $\mathcal{H}_1(\chi^d)$ such that $\mathcal{H}_1(\chi_r^d) = \Theta\mathcal{H}_1(\chi^d)$

return Reduced-order data matrix $\mathcal{H}_1(\chi_r^d)$, transformation matrix Θ

Theorem 2.4 (State Reduction). *Consider the system* (2.12), *and let the data dictionary* $\mathcal{D}_{NX} = \{u^d, \chi^d\}_{k=0}^{N-1}$ *be a measured input-extended state trajectory of the system. If state reduction is performed on the extended state* $\chi(k)$ *such that* $\chi_r(k) = \Theta\chi(k)$ *according to Algorithm 2, then the obtained reduced-order extended state* $\chi_r(k)$ *will satisfy the rank condition* (2.18).

Proof. By design, the transformation matrix Θ ensures that the obtained reduced-order extendxed state $\chi_r \in \mathbb{R}^{mL+n}$. This ensures that the rank condition (2.18) is satisfied. □

By Theorem (2.4), a reduced-order state ($\chi_r \in \mathbb{R}^{n_{red}}$) can be created, which successfully addresses the existing limitations of the existing methodology based on the maximal extended state, such as handling multiple outputs, unknown lag ℓ and unknown order n. The significance of this result is that all the existing methodologies for input-state control can be equivalently applied to the case of control using input-output data by utilizing the extended state framework and the concept of state reduction. Likewise, it also supports further development of the framework to tackle other control problems (such as boosting closed-loop system performance).

2.4 Synthesis for Stabilization

The extended state framework (along with the state reduction methodology) can now be developed towards direct data-driven control of systems. This section explores closed-loop asymptotic quadratic stabilization utilizing noise-free input-output data from the system using a model-based controller.

Considering the reduced state, the equivalent reduced-order extended state representation can be defined as,

$$\chi_r(k+1) = \mathcal{A}_r\chi_r(k) + \mathcal{B}_r u(k) \tag{2.23}$$

with the relations

$$\chi_r = \Theta\chi, \quad \mathcal{A}_r = \Theta\mathcal{A}\Theta^\dagger, \quad \mathcal{B}_r = \Theta\mathcal{B}. \tag{2.24}$$

Notice that the reduced state equation is still in a model-based representation. With the aim of reduced state-feedback, a data-based representation has to be considered. To simplify the notation, the following short-hand notations for the data matrices are used.

$$\begin{aligned} U_- &= \mathcal{H}_1(u^d_{[L,N+L-1]}) \\ Z_- &= \mathcal{H}_1(\chi^d_{r[L,N+L-1]}) \\ Z_+ &= \mathcal{H}_1(\chi^d_{r[L+1,N+L]}) \end{aligned} \tag{2.25}$$

Note that to define Z_+, an extra input-output data sample has to be measured. This again redefines the data dictionary as $\mathcal{D}_{Nr} = \{u^d_k, \chi^d_{rk}\}_{k=L}^{N+L}$. Now, the data-based system representation can be highlighted by the following theorem.

Theorem 2.5 (Data-based System Representation [12])**.** *Consider the reduced state representation* (2.23) *and let the data dictionary* $\mathcal{D}_{Nr} = \{u^d_k, \chi^d_{rk}\}_{k=L}^{N+L}$ *be a constructed input-reduced-state data set based on the measured input-output trajectory of the original data generating system* (2.1)*. Furthermore, let the condition* (2.18) *hold. Then, the reduced state representation has the following equivalent data-based representation*

$$\chi_r(k+1) = Z_+ \begin{bmatrix} U_- \\ Z_- \end{bmatrix}^\dagger \begin{bmatrix} u(k) \\ \chi_r(k) \end{bmatrix} \tag{2.26}$$

Proof. The result follows from the relations (2.24) using similar arguments as in Theorem 7 from [12]. □

Having defined the required system to be controlled using data, a model-based controller representation is formulated in the following subsection.

2.4.1 Interconnection with Model-based Controller

Initially considering the maximal extended state, a dynamic output-feedback controller can be designed as

$$y^c(k) + \mathtt{C}_1 y^c(k-1) + \cdots + \mathtt{C}_L y^c(k-L) = \mathtt{D}_1 u^c(k-1) + \cdots + \mathtt{D}_L u^c(k-L), \tag{2.27}$$

where $u_c \in \mathbb{R}^p$ and $y_c \in \mathbb{R}^m$ are the controller input and output respectively, with the polynomial matrices $\mathtt{C}_i \in \mathbb{R}^{m\times m}$ $\forall i \in [1, 2, \cdots, L]$ and $\mathtt{D}_i \in \mathbb{R}^{m\times p}$ $\forall i \in [1, 2, \cdots, L]$. The controller is connected to the system (2.1) through the interconnection,

$$u^c(k) = y(k), \quad y^c(k) = u(k), \quad k \in \mathbb{N}. \tag{2.28}$$

Notice that the interconnection results in the dynamic output-feedback controller being parameterized as a static extended-state feedback controller having the form

$$u(k) = \mathcal{K}\chi(k),$$

where $\mathcal{K}$ is a vector comprising of the controller coefficients and given by

$$\mathcal{K} = \begin{bmatrix} \mathtt{D}_1 & \mathtt{D}_2 & \cdots & \mathtt{D}_L & -\mathtt{C}_1 & -\mathtt{C}_2 & \cdots & -\mathtt{C}_L \end{bmatrix} \tag{2.29}$$

.

When the reduced state is considered with $\chi_r(k) = \Theta\chi(k)$, the static reduced state-feedback controller can be designed as

$$u(k) = \mathcal{K}_r\chi_r(k). \tag{2.30}$$

Therefore, the designed controller can be utilized for controlling the system (2.23). The next subsection defines the conditions for closed-loop stability while considering the interconnection of the data-based system representation together with the model-based controller representation.

2.4.2 Data-based Quadratic Stabilization using noise-free Data

The system (2.23) can be considered to be stabilized by a static state-feedback controller of the form $u(k) = \mathcal{K}_r\chi_r(k)$ according to the Definition 2.3 as illustrated in the below theorem.

Theorem 2.6 (Stabilization for noise-free data [17])**.** *Consider the system* (2.23) *and let the data dictionary* $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ *be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (2.1)*. Furthermore let the input sequence* $u_{[0,N+L]}$ *be persistently exciting of order* $L + n + 1$*. Then the following results hold true.*

1. *The system* (2.23) *in a closed-loop with a reduced state-feedback controller* $u(k) = \mathcal{K}_r\chi_r(k)$ *has the following finite-horizon data-based representation*

$$\chi_r(k+1) = Z_+\mathcal{G}_\mathcal{K}\chi_r(k)$$

with $\mathcal{G}_\mathcal{K}$ *satisfying*

$$\begin{bmatrix} \mathcal{K} \\ I \end{bmatrix} = \begin{bmatrix} U_- \\ Z_- \end{bmatrix} \mathcal{G}_\mathcal{K},$$

and

$$u(k) = U_-\mathcal{G}_\mathcal{K}\chi_r(k).$$

2. *The data dictionary* $\mathcal{D}_{Nr}$ *is said to be informative for quadratic stabilization by a static state feedback controller* (2.30) *if there exists any matrix* $\mathcal{Q}$ *satisfying*

$$\begin{bmatrix} Z_-\mathcal{Q} & Z_+\mathcal{Q} \\ \mathcal{Q}^\top Z_-^\top & Z_-\mathcal{Q} \end{bmatrix} \succ 0 \tag{2.31}$$

with $(Z_-\mathcal{Q}) = (Z_-\mathcal{Q})^\top$*. Moreover, this results in a controller* $u(k) = \mathcal{K}_r\chi_r(k)$ *with* $\mathcal{K}_r = U_-\mathcal{Q}(Z_-\mathcal{Q})^{-1}$ *which stabilizes the system* (2.23)*.*

Proof. See Theorem 5 from [17] for the proof. □

Theorem 2.6 provides the required conditions for the reduced state to be utilized for quadratic stabilization. Hence, it shows that the non-minimal reduced state representation (2.23) can be used to design a dynamic output-feedback controller from the measured data trajectory of the original system $\mathcal{G}$. This essentially means that the static reduced state-feedback controller can be implemented as a dynamic output feedback controller for the system $\mathcal{G}$.

Remark 2.3. Notice the condition (2.31) is affine with respect to the decision variable $\mathcal{Q}$, and hence it can be found by appropriately defining and solving a semi-definite program (SDP). Also, note that it is necessary to numerically implement the condition $(Z_-\mathcal{Q}) = (Z_-\mathcal{Q})^\top$ to solve this SDP and obtain a feasible $\mathcal{Q}$. This can be done by considering one of the constraints as $c_1 = 0.5((Z_-\mathcal{Q}) + (Z_-\mathcal{Q})^\top)$. Otherwise, the feasibility problem may not be defined correctly and may result in a numerically incorrect solution.

Remark 2.4. The decision variable $\mathcal{Q}$ is dependent on the number of data samples (N) and the depth of the Hankel matrix (L), which can be seen from Theorem 2.6. This means that as the values of either N or L are increased, the time taken by the solver to find a feasible value of $\mathcal{Q}$ will also increase. This means that this methodology can handle only a limited number of data samples before becoming computationally inefficient.

Having established formal guarantees to ensure closed-loop stability, the framework can also be extended to design output-feedback controllers for ensuring performance.

2.5 Synthesis for Performance

A performance control problem is defined as finding a stabilizing reduced-state feedback controller (2.30) that achieves an $\mathcal{H}_\infty/\mathcal{H}_2$ performance bound on the system. This section proposes a novel method for performance control of a data-based system representation in the case of noise-free data.

2.5.1 Closed-loop System

Define a generalized disturbance $w(k)$ and a (controlled) performance output $z(k)$

$$z(k) = y(k) = \mathcal{C}_r\chi_r(k) + \mathcal{D}_r u(k) + D_K w(k),$$

with $\mathcal{C}_r = \begin{bmatrix}\bar{A} & \bar{B}\end{bmatrix}\Theta^\dagger$, $\quad \mathcal{D}_r = B_0$. Defining a matrix $\mathcal{V} := \begin{bmatrix}I_p & 0 & 0 & 0\end{bmatrix}^\top \in \mathbb{R}^{(m+p)L\times p}$, the following relations hold.

$$\begin{aligned}
\mathcal{C}_r &= \begin{bmatrix}\bar{A} & \bar{B}\end{bmatrix}\Theta^\dagger = \mathcal{V}^\top\mathcal{A}\Theta^\dagger = \mathcal{V}^\top\Theta^\dagger\Theta\mathcal{A}\Theta^\dagger = \mathcal{V}^\top\Theta^\dagger\mathcal{A}_r\\
\mathcal{D}_r &= B_0 = \mathcal{V}^\top\Theta^\dagger\Theta\mathcal{B} = \mathcal{V}^\top\Theta^\dagger\mathcal{B}_r
\end{aligned}$$

This results in the following open-loop system representation.

$$\begin{aligned}
\chi_r(k+1) &= \mathcal{A}_r\chi_r(k) + \mathcal{B}_r u + B_K w(k),\\
z(k) &= \mathcal{C}_r\chi_r(k) + \mathcal{D}_r u(k) + D_K w(k).
\end{aligned} \tag{2.32}$$

Here, note that the matrices B_K and D_K are known reduced state-space matrices selected based on the control configuration. Assume that the disturbance $w(k)$ acts as an input disturbance, i.e., $w(k) \in \mathbb{R}^m$ with $B_K = \Theta \begin{bmatrix} 0 & 0 & I_m & 0 \end{bmatrix}^\top$ and $D_K = 0_{p \times m}$.

The closed-loop system is created using the reduced state-feedback controller $u = \mathcal{K}_r \chi_r$ (2.30) as

$$\begin{aligned} \chi_r^+ &= (\mathcal{A}_r + \mathcal{B}_r \mathcal{K}_r)\chi_r + B_K w = A_K \chi_r + B_K w \\ z &= (\mathcal{C}_r + \mathcal{D}_r \mathcal{K}_r)\chi_r + D_K w = C_K \chi_r + D_K w \end{aligned} \tag{2.33}$$

Hence, the transfer matrix from the generalized disturbance $w(k)$ to the performance output $z(k)$ is given by

$$T(z) = C_K (zI - A_K)^{-1} B_K + D_K, \tag{2.34}$$

for which the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms are defined as $||T||_\infty$ and $||T||_2$ respectively. Using this closed-loop representation, the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control problems are investigated.

2.5.2 Data-based $\mathcal{H}_\infty$ control using noise-free Data

$\mathcal{H}_\infty$ control refers to ensuring a desired $\mathcal{H}_\infty$ norm performance guarantee for the closed-loop system (2.33) under the worst-case gain from input to output. Firstly, the idea of model-based $\mathcal{H}_\infty$ control is defined.

Definition 2.4 (Model-based $\mathcal{H}_\infty$ control [28])**.** Consider the system (2.33) and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system (2.1). The data $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_\infty$ control ($||T||_\infty < \gamma$) by a reduced state-feedback controller (2.30) with a performance metric γ if there exist a $\mathcal{K}_r$ and $X \succ 0$ such that

$$\begin{bmatrix} X & 0 & A_K^\top X & C_K^\top \\ 0 & \gamma I & B_K^\top X & D_K^\top \\ X A_K & X B_K & X & 0 \\ C_K & D_K & 0 & \gamma I \end{bmatrix} \succ 0 \tag{2.35}$$

holds, with A_K and C_K as defined in (2.33). □

Based on this definition of $\mathcal{H}_\infty$ control, the necessary conditions on the data towards its informativity for $\mathcal{H}_\infty$ control can be defined.

Theorem 2.7 (Data-based $\mathcal{H}_\infty$ control)**.** *Consider the system* (2.33) *and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (2.1)*. Furthermore let the input sequence $u_{[0,N+L]}$ be persistently exciting of order $L+n+1$. Then the data dictionary $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_\infty$ control (such that $||T||_\infty \prec \gamma$) with performance metric γ using a static reduced state-feedback controller* (2.30) *if there exists a matrix $\mathcal{Q}$ such that*

$$\begin{bmatrix} Z_- Q & 0 & (Z_+ Q)^\top & (\mathcal{V}^\top \Theta^\dagger Z_+ Q)^\top \\ 0 & \gamma I & B_K^\top & D_K^\top \\ Z_+ Q & B_K & Z_- Q & 0 \\ \mathcal{V}^\top \Theta^\dagger Z_+ Q & D_K & 0 & \gamma I \end{bmatrix} \succ 0 \tag{2.36}$$

holds. For such informative data, the controller $u = \mathcal{K}_r \chi_r$ *which achieves the performance* γ *can be designed with* $\mathcal{K}_r = U_- \mathcal{Q}(Z_- \mathcal{Q})^{-1}$.

Proof. Consider the definition of $\mathcal{H}_\infty$ control and the corresponding condition (2.35). Perform a congruence transformation with diag (X^{-1}, I, X^{-1}, I) to obtain

$$\begin{bmatrix} P & 0 & PA_K^\top & PC_K^\top \\ 0 & \gamma I & B_K^\top & D_K \\ A_K P & B_K & P & 0 \\ C_K P & D_K & 0 & \gamma I \end{bmatrix} \succ 0. \tag{2.37}$$

Using Theorem 2.6, the following relations can be defined.

i $I = Z_- G_K \implies P = Z_- Q$ with $Q = G_K P$,

ii $A_K = Z_+ G_K \implies A_K P = Z_+ Q$ and $(A_K P)^\top = (Z_+ Q)^\top$,

iii $C_K = \mathcal{V}^\top \Theta^\dagger (\mathcal{A}_r + \mathcal{B}_r \mathcal{K}_r) = \mathcal{V}^\top \Theta^\dagger A_K \implies (C_K P) = \mathcal{V}^\top \Theta^\dagger Z_+ Q$ and $(C_K P)^\top = (\mathcal{V}^\top \Theta^\dagger Z_+ Q)^\top$.

Their substitution in (2.37) results in the condition (2.36). □

Notice that the condition (2.36) is affine with respect to the decision variable $\mathcal{Q}$, and hence it can be found by appropriately defining and solving an SDP by minimizing γ.

2.5.3 Data-based $\mathcal{H}_2$ control using noise-free Data

$\mathcal{H}_2$ control refers to ensuring a desired $\mathcal{H}_2$ norm performance guarantee for the closed-loop system under an average-case gain from input to output. Similar to the previous subsection, model-based $\mathcal{H}_2$ control is first defined.

Definition 2.5 (Model-based $\mathcal{H}_2$ control [19])**.** Consider the system (2.33) and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system (2.1). The data $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_2$ control ($||T||_2 < \gamma$) by a reduced state-feedback controller (2.30) with a performance metric γ if there exist a $\mathcal{K}_r$ and $X \succ 0$ such that

$$\operatorname{trace}(B_K^\top X B_K + D_K^\top D_K) < \gamma^2 \quad \text{and} \quad X \succ A_K^\top X A_K + C_K^\top C_K, \tag{2.38}$$

hold, with A_K and C_K as defined in (2.33). □

Based on this definition of $\mathcal{H}_2$ control, the necessary informativity conditions for data-based $\mathcal{H}_2$ control can be defined.

Theorem 2.8 (Data-based $\mathcal{H}_2$ control)**.** *Consider the system* (2.33) *and let the data dictionary* $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ *be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (2.1)*. Furthermore let the input sequence* $u_{[0,N+L]}$ *be persistently exciting of order* $L+n+1$*. Then* *the data dictionary $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_2$ control (such that $||T||_2 \prec \gamma$) with performance metric γ using a static reduced state-feedback controller* (2.30) *if there exists a matrix $\mathcal{Q}$ such that*

$$\begin{gathered}\operatorname{trace}(Z) < \gamma^2, \quad \begin{bmatrix} Z & B_K^\top & D_K^\top \\ B_K & Z_- Q & 0 \\ D_K & 0 & I \end{bmatrix} \succeq 0 \quad \text{and} \\ \begin{bmatrix} Z_- Q & (Z_+ Q)^\top & (\mathcal{V}^\top \Theta^\dagger Z_- Q)^\top \\ Z_+ Q & Z_- Q & 0 \\ \mathcal{V}^\top \Theta^\dagger Z_- Q & 0 & I \end{bmatrix} \succ 0,\end{gathered} \tag{2.39}$$

hold. For such informative data, the controller $u = \mathcal{K}_r \chi_r$ which achieves the performance γ can be designed with $\mathcal{K}_r = U_- \mathcal{Q}(Z_- \mathcal{Q})^{-1}$.

Proof. Consider the definition of $\mathcal{H}_2$ control and the corresponding condition (2.38). Take the first equation and consider $X = P^{-1}$ such that

$$\operatorname{trace}(B_K^\top P^{-1} B_K + D_K^\top D_K) < \gamma^2.$$

Taking $Z := B_K^\top P^{-1} B_K + D_K^\top D_K$ results in the first condition of (2.39) ($\operatorname{trace}(Z) < \gamma^2$), and $Z - B_K^\top P^{-1} B_K - D_K^\top D_K \succeq 0$, whose Schur complement results in

$$\begin{bmatrix} Z - D_K^\top D_K & B_K^\top \\ B_K & P \end{bmatrix} \succeq 0 \quad \equiv \begin{bmatrix} Z & B_K^\top & D_K^\top \\ B_K & P & 0 \\ D_K & 0 & I \end{bmatrix} \succeq 0.$$

Substitute with the relation $P = Z_- Q$ to obtain the second condition of (2.39). Perform a congruence transformation with $P = X^{-1}$ on the second equation of(2.38) so that

$$P - (A_K P)^\top P^{-1} (A_K P) - (C_K P)^\top (C_K P) \succ 0.$$

Taking the Schur complement results in

$$\begin{bmatrix} P - (C_K P)^\top (C_K P) & (A_K P)^\top \\ A_K P & P \end{bmatrix} \succ 0,$$

whose substitution with the known relations from Theorem 2.6 results in the third condition of (2.39). □

Notice that the conditions (2.39) are affine with respect to the decision variable $\mathcal{Q}$ in the case of $\mathcal{H}_2$ control as well, and hence it can be found by appropriately defining and solving an SDP for minimizing γ.

The discussed methodologies for performance control are a novel extension to the existing extended state framework available in the literature for noise-free data, which focuses only on quadratic stabilization. Furthermore, the closed-loop transfer is between a generalized input and output signal. This allows for the shaping of the closed-loop transfers to achieve the desired frequency and time-domain characteristics, e.g., through the inclusion of weighting filters in a mixed-sensitivity approach [15], which is a potential future work.

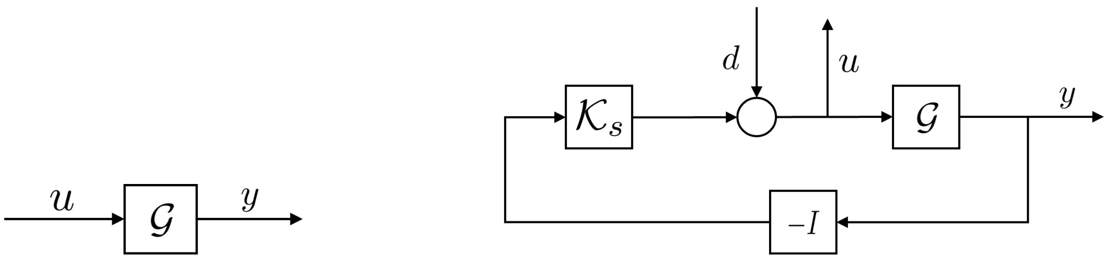
Having established the methodology of stabilization and enforcing performance using the extended state framework, it is necessary to show the implementability and validity of the obtained results.

2.6 Implementation

The extended state framework allows for the utilization of the constructed input-reduced state data set $\mathcal{D}_{NX}$ to create a static reduced-state feedback controller. This requires the generation and collection of the input-output data set $\mathcal{D}_N$. Also, it has been mentioned that an equivalent dynamic output feedback controller can be designed based on the obtained controller coefficient. This section discusses the input-output data collection and the controller implementation.

2.6.1 Data Collection

The data is collected from a system in two ways - open-loop or closed-loop. In an open-loop setting, the desired input signal to the system (u_k^d) and the output response (y_k^d) are measured. For the closed-loop setting, a model-based controller $\mathcal{K}_s$ with a low bandwidth is designed to ensure closed-loop stability with the data-generation system. In this case, an input disturbance signal (d_k^d) is provided as the external disturbance signal satisfying the persistence of excitation conditions. Then, the actual system input signal (u_k^d) and the output response (y_k^d) are measured, as shown in Figure 2.2.



(a) Open loop setting (b) Closed loop setting

Figure 2.2: Data Generation System Setting

In certain cases, i.e., when there is a significant relative difference in the orders of magnitudes of the measured input and output data, the SDP might face numerical errors while obtaining a feasible solution. This can be resolved by scaling the data, i.e., ensuring that the magnitude of the data samples is bounded by an absolute value of 1. This is done by division of the input and output data samples by the corresponding maximum values. Note that this is essentially a multiplication with a diagonal scaling matrix with each term being the inverse of the maximum signal value obtained from the data.

2.6.2 Controller Implementation

The controller implementation is performed on `Simulink` according to the block diagram representation shown in Figure 2.3. The matrices Di and Do are state-space representations of the delays, or specifically shift registers, which are used for the creation of the lagged input and output samples. Their design is highlighted below.

$$
\begin{aligned}
D_{delay} &= \left(\begin{array}{c|c} A_{mat} & B_{mat} \\ \hline C_{mat} & D_{mat} \end{array}\right) = \left(\begin{array}{cc|c} 0_{1\times 1} & 0_{1\times L-1} & I_{1\times 1} \\ 0_{(L-1)\times 1} & I_{L-1} & 0_{(L-1)\times 1} \\ \hline \multicolumn{2}{c|}{I_L} & 0_{L\times 1} \end{array}\right), \\
Di &= \left(\begin{array}{c|c} A_{mat}\otimes I_m & B_{mat}\otimes I_m \\ \hline C_{mat}\otimes I_m & D_{mat}\otimes I_m \end{array}\right), \\
Do &= \left(\begin{array}{c|c} A_{mat}\otimes I_p & B_{mat}\otimes I_p \\ \hline C_{mat}\otimes I_p & D_{mat}\otimes I_p \end{array}\right).
\end{aligned}
\tag{2.40}
$$

The resulting lagged input-output samples can be used to create the maximal extended state. `Theta` is the transformation matrix Θ designed from the knowledge of the lagged input-output data samples according to Theorem 2.4, which results in the reduced state such that $\chi_r(k) = \Theta\chi(k)$. Furthermore, when scaling is considered, the signals are pre- and post-multiplied with an appropriate diagonal scaling matrix such that the diagonal terms are the inverse and actual maximum values of the measured signals respectively. This ensures that the scaled signals are utilized for the generation of the input signal according to the control law $u(k) = \mathcal{K}_r\chi_r(k)$.

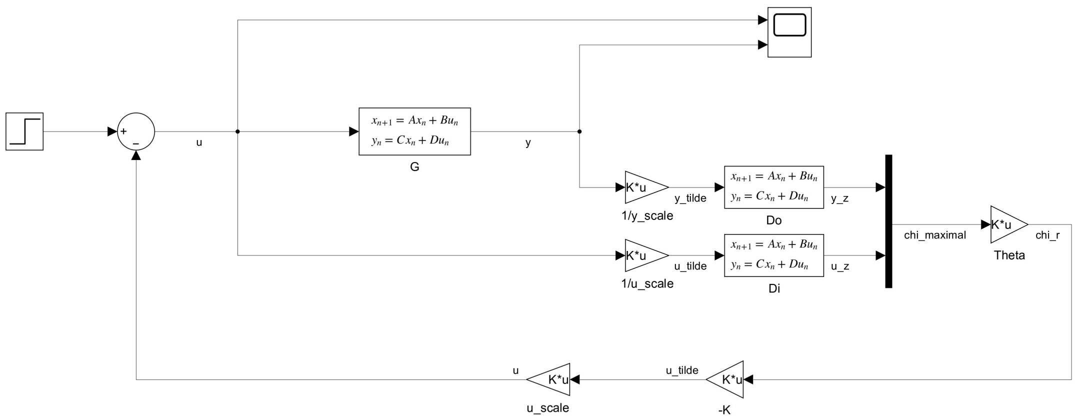


Figure 2.3: Controller Implementation - MIMO System

2.7 Simulation Examples

Noise-free data is a theoretic assumption, and hence, the validation of the developed framework is limited to a simulation environment. The aim is to utilize input-output data from a system for developing controllers enforcing closed-loop stability and performance.

2.7.1 Experiment Setting

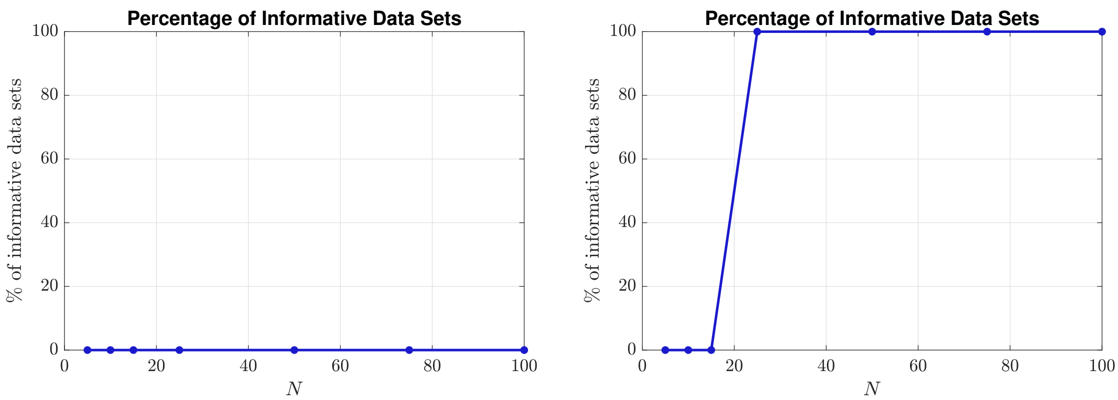
An academic example from [18] is considered as the data-generating system $\mathcal{G}_D$ (2.41). This is an unstable system in its minimal representation with two inputs ($m = 2$) and two outputs ($p = 2$), with the observer matrix C chosen randomly and the sampling time as $T_s = 0.1$ seconds. Furthermore, it can also be seen that the order $n = 3$ and the lag $\ell = 2$, though this prior knowledge is not utilized for controller design.

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 0.850 & -0.038 & -0.380 \\ 0.735 & 0.815 & 1.594 \\ -0.664 & 0.697 & -0.064 \end{bmatrix} x(k) + \begin{bmatrix} 1.431 & 0.705 \\ 1.620 & -1.129 \\ 0.913 & 0.369 \end{bmatrix} u(k) \\ y(k) &= \begin{bmatrix} 0.6458 & 1.1183 & -1.1694 \\ -1.4651 & -0.0050 & 0.5865 \end{bmatrix} x(k) \end{aligned} \tag{2.41}$$

The number of data samples N has varying values taken as $N \in [5, 10, 15, 25, 50, 75, 100]$. For each value of N, 10 Monte Carlo simulation runs are performed. Since the data-generating system is unstable, input-output data is collected in a closed-loop setting with a model-based stabilizing controller having a low bandwidth, which allows for the system dynamics to be captured effectively with reduced interference by the controller. The input disturbance ($d(k)$) consists of samples drawn from a Gaussian white noise disturbance with zero mean and unit standard deviation (i.e., $d_k^d \sim \mathcal{N}(0, 1)$). The system input ($u_k^d$) and output ($y_k^d$) responses are measured. For the creation of the data matrices, the depth is considered as $L = \ell + 6 = 8$ The solver used for solving the SDP for a feasible $\mathcal{Q}$ is MOSEK [29]. These are the experiment settings considered for the following examples.

2.7.2 Example - Data Informativity for Quadratic Stabilization

Consider the system (2.41) with the discussed experiment settings. Initially, the problem of data informativity for quadratic stabilization is analyzed (i.e., if the collected data represents the finite-horizon system behavior $\mathfrak{B}|_L^{\mathcal{G}_D}$ and can be used for designing a stabilizing controller) based on the methodology proposed in Theorem 2.6.



(a) Maximal Extended State

(b) Reduced State

Figure 2.4: Data Informativity for Quadratic Stabilization

Initially, the extended state is not reduced, and the data informativity results are shown in Figure 2.4a. It shows that the collected data dictionary $\mathcal{D}_{Nr}$ does not represent the system behavior $\mathfrak{B}|_L^{\mathcal{G}_D}$ and hence a stabilizing controller of the form (2.30) cannot be designed.

The reduced state is designed according to Theorem 2.4 with the transformation matrix Θ designed in each experiment based on the collected data. The data informativity is plotted in Figure 2.4b. It can be seen that when a sufficient number of data points are collected, then the data dictionary $\mathcal{D}_{Nr}$ can be used to design a stabilizing controller of the form (2.30).

Having established data informativity, the data-based controller is synthesized for $N = 100$ and is implemented on `Simulink`, as shown in 2.3. The system is simulated with random initial conditions and the resulting input and output trajectories are shown in Figure 2.5. The feasibility of condition (2.31), which allows the design of a stabilizing controller $\mathcal{K}_r$, as well as the asymptotic system behavior, validate the closed-loop stability of the system $\mathcal{G}_D$.

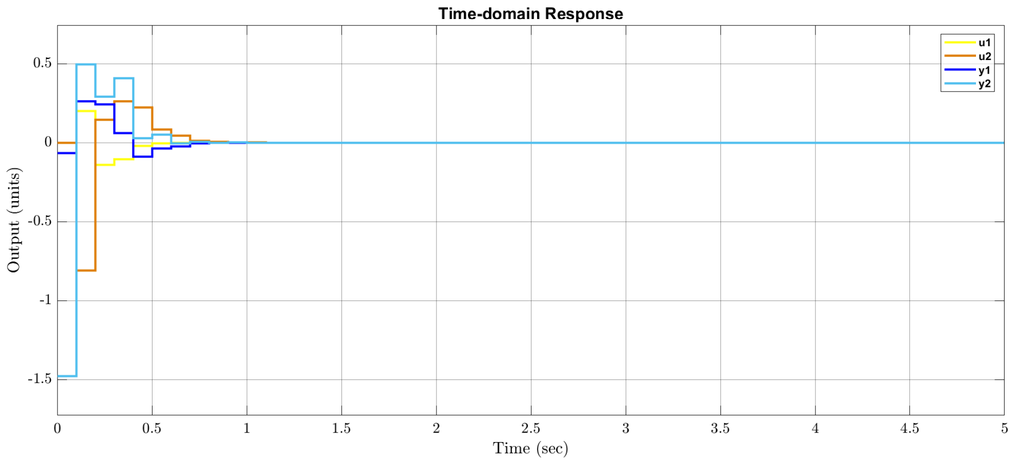
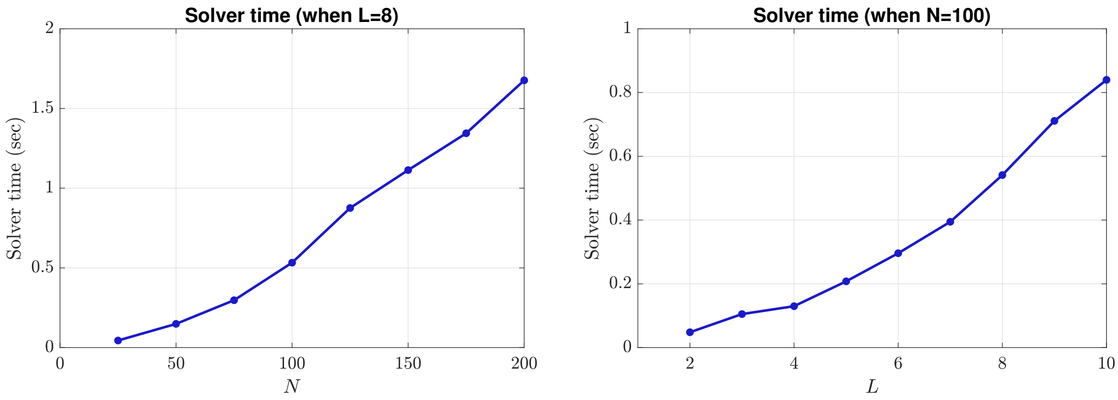


Figure 2.5: Time-domain Response for a MIMO System

2.7.3 Example - Solver Time

Consider the system (2.41) with the discussed experiment settings. Theorem 2.6 is applied for data-based quadratic stabilization, and the computation times taken by the solver to obtain a feasible solution $\mathcal{Q}$ are recorded for different cases of N or L (with the other being constant). These are shown in Figure 2.6, which indicate that the solver time increases with an increase in either N or L. This is as expected as the size of the decision variable $\mathcal{Q}$ increases with an increase in either N or L and hence requires more computation time to obtain a feasible solution.

These examples show that the maximal extended state cannot be utilized for any MIMO system of unknown complexity, and hence necessitate the usage of the reduced state (and the proposed state reduction methodology). Furthermore, the developed framework allows for an unknown system order and an unknown system lag, as the controller can be synthesized without prior knowledge about either of them. This is a significant strength of this methodology as it allows the framework to be employed for real-life applications. This conclusively shows that the extended state framework, together with the reduced state, can be used for controller synthesis.



(a) Solver time with $L = 8$ (b) Solver time with $N = 100$

Figure 2.6: Solver times (for Quadratic Stabilization)

2.7.4 Example - Data Informativity for Performance

Consider the system (2.41) with the discussed experiment settings. With $N = 100$, Theorem 2.7 and 2.8 are applied for obtaining the data-based $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms of the closed-loop system (2.33). It is observed that the data informativity plots are the same as in Figure 2.4b, so they have not been plotted again. The data-based norms are found as $\gamma_\infty = 1.598$ and $\gamma_2 = 2.562$. This shows that the extended state framework can also be used to establish data-based performance.

To check the time-domain performances, the data-based controllers are implemented on `Simulink` 2.3 with the system simulated with random initial conditions and subjected to a unit step signal at time instant $k = 2$ acting as the generalized input disturbance signal. Firstly, the input disturbance and output trajectories for a stabilizing controller are measured as shown in Figure 2.7. The trajectories are then obtained for $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control and are shown in Figures 2.8 and 2.9 respectively.

It can be observed that the system responses are dampened compared to the case of quadratic stabilization. This is due to the minimization of the energy gain between the generalized disturbance and the performance output, reflecting an improvement in the disturbance rejection properties of the system. However, considering that specifications on the desired time and frequency-domain responses are undefined, a direct comparison of the time-domain responses with any desired specifications cannot be made.

The considered examples establish the extended state framework to handle noise-free input-output data for control. Furthermore, the existing limitations on the extended state, namely multiple outputs, and unknown system lag and order, have been completely addressed with a working methodology validated through the simulation examples.

However, this method cannot be applied to real systems, as it does not account for measurement noise. Additionally, the time required to obtain a feasible controller is directly proportional to the number of samples N, making it computationally unattractive for applications with large datasets. These motivate an extension of the analysis to noisy input-output data to address these limitations.

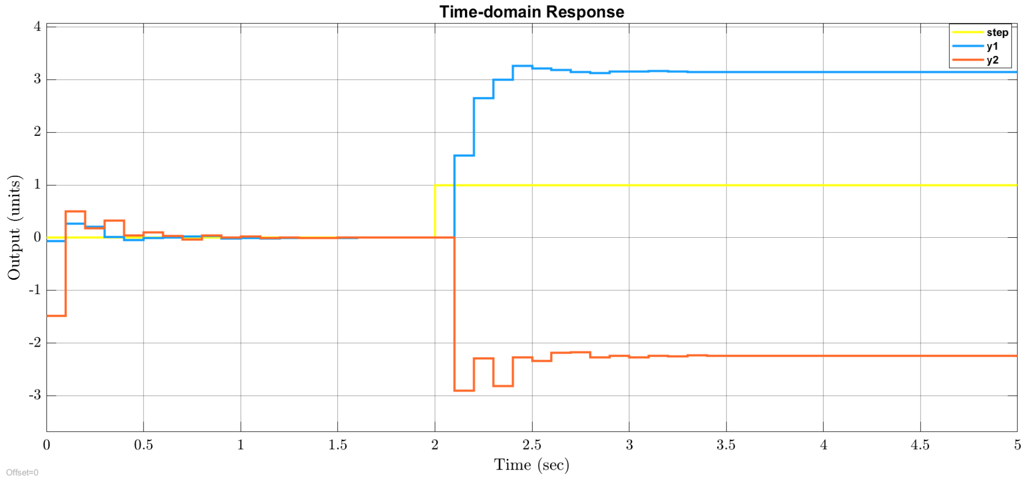


Figure 2.7: Time-domain Response for Stabilization

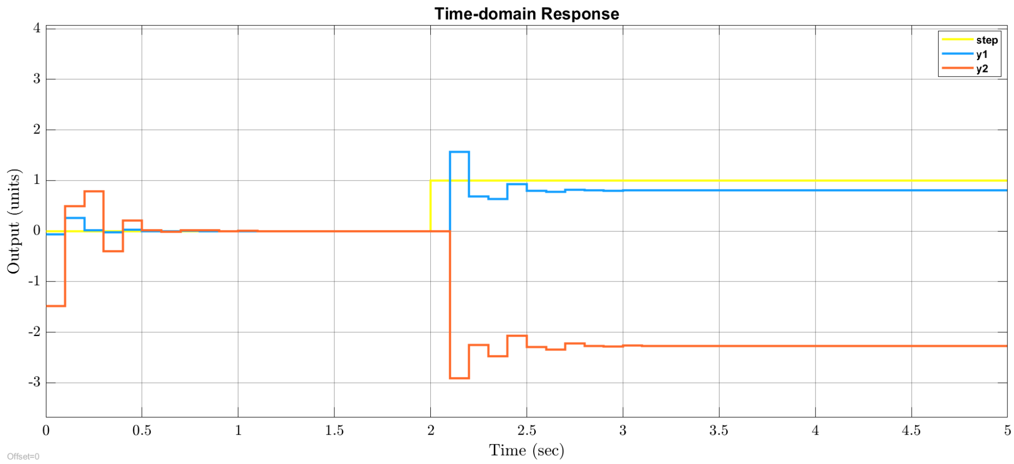


Figure 2.8: Time-domain Response for $\mathcal{H}_\infty$ control

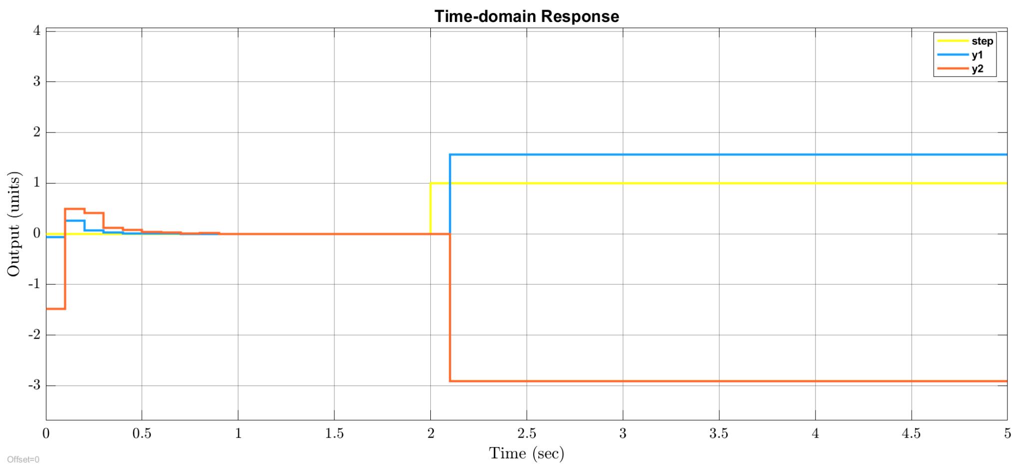


Figure 2.9: Time-domain Response for $\mathcal{H}_2$ control

Chapter 3

Data-Driven Control Using Noisy Data

In the previous chapter, it has been shown for the case of noise-free data that a maximally extended state cannot straightforwardly replace the actual system's state for the application of current state-of-the-art input-state data-driven control methods. A state reduction method based on a singular value decomposition has hence been proposed in order to use the input-state data-driven methods. This approach has been shown to be effective in the simulation examples in the absence of noise. This chapter focuses on the consideration of external noise entering the system through the measured output data.

An overview of existing results for input-state data-driven control in the presence of noise is provided. An assumption is considered on the noise, which enables the definition of an extension of the data-driven system representation for noise-contaminated data. This analysis is extended to the case of noisy input-output data-driven control by the introduction of an equivalent state reduction methodology. To effectively handle the measurement noise, a novel instrumental variable method is proposed in both open-loop and closed-loop settings. Utilizing these concepts, the extended state framework is expanded to consider the development of controller synthesis and performance techniques for noisy input-output data. The validity is showcased through simulations examples, and a proof-of-concept application on a real system.

3.1 Data-Driven control using Input-State Data

As in the previous chapter, the aim is to examine the existing procedures for data-based input-state control in the presence of noise and redefine them for the case of data-based input-output control using the (reduced) extended state. Hence, this section discusses the existing methodologies for input-state data-driven control (see [18]).

3.1.1 Model-based Input-State System Description

Consider the following system representation in the form of a state equation

$$x(k+1) = Ax(k) + Bu(k) + v(k), \tag{3.1}$$

with $u(k) \in \mathbb{R}^m$ denoting the input, $v(k) \in \mathbb{R}^p$ a random variable denoting the stochastic system noise and $x(k) \in \mathbb{R}^n$ denoting the state. Consider a measured input-state data trajectory $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^N$ of this system, which can be then used to define the system data equation

$$X_+ = AX_- + BU_- + V_-, \tag{3.2}$$

where the data matrices are given by

$$\begin{aligned}
X_- &= \mathcal{H}_1(x^d_{[0,N-1]}),\\
X_+ &= \mathcal{H}_1(x^d_{[1,N]}),\\
U_- &= \mathcal{H}_1(u^d_{[0,N-1]}),\\
V_- &= \mathcal{H}_1(v_{[0,N-1]}).
\end{aligned}$$

Note that the noise ($v(k)$), and consequently, the matrix of noise data samples (V_-), is not measured. The goal of this definition is to obtain a data-based representation for this noisy input-state system and define the data-informativity conditions that can be utilized for the development of the control methodology.

3.1.2 Assumption on the Noise

Since the considered data-generating system is affected by noise and thus noisy data is collected for control, an assumption is chosen on the energy bounds of the noise as

$$||v(k)||_2^2 \leq \epsilon \ \forall k, \text{ for some scalar } \epsilon > 0$$

This can be equivalently written as

$$V_- V_-^\top = \Sigma_{k=0}^N v(k) v(k)^\top \preceq \Phi_{11},$$

with $\Phi_{11} = (N+1)\epsilon I$. This assumption on the bounds of the noise samples can be generalized below.

Assumption 3.1. *([18]) The noise samples collected in V_- satisfy the bound*

$$\begin{bmatrix} I \\ V_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ V_-^\top \end{bmatrix} \succeq 0, \tag{3.3}$$

for known $\Phi_{11} = \Phi_{11}^\top = (N+1)\epsilon I$, $\epsilon > 0$, Φ_{12} and $\Phi_{22} = \Phi_{22}^\top \prec 0$.

An important point to be considered in Assumption 3.1 is that since Φ_{22} is negative definite, the noise matrix V_- will satisfy the bounds given by (3.3). Specifically, the assumption (3.3) can be equivalently written as

$$\begin{aligned}
&\Phi_{11} + V_-\Phi_{12}^\top + \Phi_{12}V_-^\top + V_-\Phi_{22}V_-^\top \succeq 0,\\
\implies &\Phi_{11} + V_-\Phi_{12}^\top + \Phi_{12}V_-^\top \succeq -V_-\Phi_{22}V_-^\top \succeq 0.
\end{aligned}$$

Based on this noise assumption, a data-based system representation can be formulated (as described by Equation (8) [18]).

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \succeq 0 \tag{3.4}$$

This data-based system representation can be used to develop the data-driven control methodologies for input-state data.

3.1.3 Control using Input-State Data

Initially consider the quadratic stabilization problem, i.e., the design of a stabilizing controller for the system described by (3.4). The notion of quadratic stabilization has been defined in the Definition 2.3. To utilize this notion, the Matrix-valued S-lemma can be used.

Lemma 3.1 (Matrix S-lemma [18]). *Suppose that there exists some matrix $\bar{Z} \in \mathbb{R}^{n\times i}$ and let $M, N \in \mathbb{R}^{(i+n)\times(i+n)}$ be symmetric matrices such that the generalized Slater's condition*

$$\begin{bmatrix} I \\ \bar{Z} \end{bmatrix}^\top N \begin{bmatrix} I \\ \bar{Z} \end{bmatrix} \succ 0, \tag{3.5}$$

holds true. Then the following statements are equivalent for some $\tilde{Z} \in \mathbb{R}^{n\times i}$.

I $\begin{bmatrix} I \\ \tilde{Z} \end{bmatrix}^\top M \begin{bmatrix} I \\ \tilde{Z} \end{bmatrix} \succeq 0$ *with* $\begin{bmatrix} I \\ \tilde{Z} \end{bmatrix}^\top N \begin{bmatrix} I \\ \tilde{Z} \end{bmatrix} \succeq 0.$

II $\begin{bmatrix} I \\ \tilde{Z} \end{bmatrix}^\top M \begin{bmatrix} I \\ \tilde{Z} \end{bmatrix} \succeq 0$ *with* $\begin{bmatrix} I \\ \tilde{Z} \end{bmatrix}^\top N \begin{bmatrix} I \\ \tilde{Z} \end{bmatrix} \succ 0.$

III There exists a scalar $\alpha \geq 0$ such that $M - \alpha N \succeq 0$.

Proof. See Theorem 9 from [18] for the proof. □

The Matrix S-lemma provides conditions under which the non-negativity of one quadratic matrix inequality (QMI) implies that of another one. Noticing that the data-based system representation (3.4) is in the form of a QMI, the Matrix S-lemma can be used for establishing the necessary conditions to ensure quadratic stability.

Theorem 3.1 (Stabilization for noisy Input-State data). *Consider the system* (3.1) *and let the data dictionary $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^N$ be a measured input-state trajectory of the system. Then, assuming that the generalized Slater's condition* (3.5) *holds, the data dictionary $\mathcal{D}_{NS}$ is said to be informative for quadratic stabilization using a static state-feedback controller $u = Kx$ if and only if there exists matrices $P = P^\top \succ 0 \in \mathbb{R}^{n\times n}$,*

$E \in \mathbb{R}^{m\times n}$, and scalars $\alpha \geq 0$ and $\beta > 0$ such that

$$\begin{bmatrix} P-\beta I & 0 & 0 & 0 \\ 0 & -P & -E^\top & 0 \\ 0 & -E & 0 & E \\ 0 & 0 & E^\top & P \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix}^\top \succeq 0, \tag{3.6}$$

holds. For such informative data, the stabilizing controller $u = Kx$ can be designed with $K = EP^{-1}$.

Proof. See Theorem 14 from [18] for the proof. □

Theorem (3.1) provides a necessary and sufficient condition for the design of stabilizing controllers from noise input-state data.

Similarly, the framework can also be extended to include performance specifications, specifically the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control problems based on Definitions 2.35 and 2.38 respectively. Consider a known performance output $z(k) \in \mathbb{R}^p$ such that $z(k) = C_z x(k) + D_z u(k)$. With the feedback law $u = Kx$, the closed-loop system becomes

$$\begin{aligned} x(k+1) &= (A+BK)x(k) + v(k), \\ z(k) &= (C+DK)x(k). \end{aligned}$$

The transfer matrix from $v(k)$ to $z(k)$ is given by

$$T_S(z) := (C+DK)(zI-(A+BK))^{-1},$$

with the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms given by $||T_S||_\infty$ and $||T_S||_2$ respectively. Then, the following results are established.

Theorem 3.2 ($\mathcal{H}_\infty$ control for Input-State Data)**.** *Consider the system (3.1) and let the data dictionary $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^N$ be a measured input-state trajectory of the system. Then, assuming that the generalized Slater's condition (3.5) holds, the data dictionary $\mathcal{D}_{NS}$ is said to be informative for $\mathcal{H}_\infty$ control (such that $||T_S||_\infty < \gamma$) with performance metric γ using a state-feedback controller of the form $u = Kx$ (2.30) if and only if there exist matrices $Y = Y^\top \succ 0$ and E, and scalars $\alpha \geq 0$ and $\beta > 0$ such that*

$$\begin{bmatrix} Y-\beta I & 0 & 0 & 0 & C_{Y,E}^\top \\ 0 & 0 & 0 & Y & 0 \\ 0 & 0 & 0 & E & 0 \\ 0 & Y & E^\top & Y-\frac{1}{\gamma^2}I & 0 \\ C_{Y,E} & 0 & 0 & 0 & I \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} (*)^\top \succeq 0,$$

$$Y - \frac{1}{\gamma^2}I \succ 0, \tag{3.7}$$

hold, with $C_{Y,E} = CY + DE$. For such informative data, the controller $u = Kx$ which achieves the performance γ can be designed as $K = EY^{-1}$.

Proof. See Theorem 20 from [18] for the proof. □

Theorem 3.3 ($\mathcal{H}_2$ control for Input-State Data)**.** *Consider the system* (3.1) *and let the data dictionary* $\mathcal{D}_{NS} = \{u_k^d, x_k^d\}_{k=0}^N$ *be a measured input-state trajectory of the system. Then, assuming that the generalized Slater's condition* (3.5) *holds, the data dictionary* $\mathcal{D}_{NS}$ *is said to be informative for* $\mathcal{H}_2$ *control (such that* $||T_S||_2 < \gamma$*) with performance metric* γ *using a state-feedback controller of the form* $u = Kx$ (2.30) *if and only if there exist matrices* $Y = Y^\top \succ 0$*,* $Z = Z^\top$ *and* E*, and scalars* $\alpha \geq 0$ *and* $\beta > 0$ *such that*

$$\begin{bmatrix} Y-\beta I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Y & 0 \\ 0 & 0 & 0 & E & 0 \\ 0 & Y & E^\top & Y & C_{Y,E}^\top \\ 0 & 0 & 0 & C_{Y,E} & I \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} (*)^\top \succeq 0,$$

$$\begin{bmatrix} Y & C_{Y,E}^\top \\ C_{Y,E} & I \end{bmatrix} \succ 0, \quad \begin{bmatrix} Z & I \\ I & Y \end{bmatrix} \succeq 0, \quad \mathrm{trace}(Z) < \gamma^2, \tag{3.8}$$

hold, with $C_{Y,E} = CY + DE$*. For such informative data, the controller* $u = Kx$ *which achieves the performance* γ *can be designed as* $K = EY^{-1}$*.*

Proof. See Theorem 17 from [18] for the proof. □

These results define the data-driven conditions for various performance specifications. Also, Theorems 3.1 - 3.3 outline the existing methodologies [18] for data-driven control using input-state data measured in the presence of noise. Based on these ideas, the extended state framework can be introduced to handle input-output data.

3.2 Extended State for Noisy Data

It has been well-established that the extended-state framework provides a working procedure for data-driven control using input-output data. This section discusses the input-output system representation and the corresponding extended-state space system representation.

3.2.1 Model-based System Representation

Consider a general representation of a DT-LTI MIMO system, which is supposed to be the data-generating system affected by noise and has the following minimal representation

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + H\eta(k), \\ y(k) &= Cx(k) + Du(k) + I\eta(k), \end{aligned} \tag{3.9}$$

with $y(k) \in \mathbb{R}^p$ denoting the output, $u(k) \in \mathbb{R}^m$ denoting the input, $\eta(k) \in \mathbb{R}^p$ is a random variable drawn from a Gaussian white noise process (representing the stochastic system noise) and $x(k) \in \mathbb{R}^n$ denoting the state. This is referred to as the innovation form [30], which is suitable to describe general LTI stochastic processes. The system (3.9) can be represented using an equivalent stochastic model structure. Consider a Box-Jenkins (BJ) model-structure representation

$$y(k) = \mathcal{G}u(k) + \mathcal{H}\eta(k) \tag{3.10}$$

where $\mathcal{G}$ represents the true system and $\mathcal{H}$ represents the (stable and inversely stable) noise model. To consider a noise process without any assumptions on the stochastic properties, the system is represented as

$$y(k) = \mathcal{G}u(k) + v(k) \tag{3.11}$$

such that $v(k) = \mathcal{H}\eta(k) \in \mathbb{R}^p$ represents an unknown external measurement noise. The input-output data is collected from the system (3.11) for the creation of the maximal extended state

3.2.2 Extended State System Representation

The collected input-output data dictionary $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N+L}$ can be used to construct a maximal extended (or extended) state $\chi(k) \in \mathbb{R}^{(m+p)L}$ as described by (2.11). Using the defined maximal extended state, the state space representation for the full system can be developed as

$$\begin{aligned} \chi(k+1) &= \mathcal{A}\chi(k) + \mathcal{B}u(k) + \mathcal{V}v(k), \\ y(k) &= \begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} \chi(k) + \mathsf{B}_0 u(k) + v(k) \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} \mathcal{A} = \Lambda_e + J_1 &= \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\bar{A}} & \multicolumn{2}{c}{\bar{B}} \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \hline I_{p(L-1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & I_{m(L-1)} & 0 \end{array}\right] \\ \mathcal{B} = B_e + J_2 &= \left[\begin{array}{c} B_0 \\ \hline 0 \\ 0 \\ 0 \end{array}\right] + \left[\begin{array}{c} 0 \\ \hline 0 \\ I_m \\ 0 \end{array}\right] \\ \mathcal{V} &= \left[\begin{array}{c} I_p \\ \hline 0 \\ 0 \\ 0 \end{array}\right] \end{aligned} \tag{3.13}$$

Notice that such a representation, specifically the partitioning of the state-space matrices $\mathcal{A}$ and $\mathcal{B}$, helps to separate them into the unknown parameter matrices Λ_e and B_e and binary matrices J_1 and J_2 of known dimensions and structure.

It has been established in the previous chapter that for employing the extended state framework for data-driven control, the maximal extended state has to be reduced to an appropriate lower dimension. However, the presence of noise motivates suitable modifications to consider the applicability of state reduction.

3.2.3 Data Informativity

It is known that for the case of noise-free input-output data, the extended state needs to satisfy the rank condition (2.17) to ensure equivalency in behavior, but the maximal extended state ($\chi(k)$) does not do so. This motivated the development of the reduced-order extended state ($\chi_r(k)$), which satisfies the condition as defined in (2.18) and represented the finite-horizon behavior $\mathfrak{B}|_L^{\mathcal{G}}$ after state reduction. However, this is not the case with noisy input-output data.

Proposition 3.1. *Consider the maximal extended state* $\chi(k)$ (2.11)*. In the presence of noise, following rank condition*

$$\operatorname{rank}\left(\begin{bmatrix}\mathcal{H}_1(u^d_{[L,N+L-1]})\\ \mathcal{H}_1(\chi^d_{[L,N+L-1]})\end{bmatrix}\right) = m(L+1) + pL, \tag{3.14}$$

holds if the noise satisfies the noise assumption (3.20)*.*

Proof. Consider the data matrices individually. By design, $\mathcal{H}_1(u^d_{[L,N+L-1]})$ has full row rank (i.e., rank $\left(\mathcal{H}_1(u^d_{[L,N+L-1]}\right) = m$). Considering the extended-state data matrix in the case of noise-free data, rank $(\mathcal{H}_1(\chi^d_{r[L,N+L-1]})) = mL + n$, and it has $pL - n$ linearly dependent rows. Specifically, amongst the pL output data rows, only n rows are linearly independent. This is by construction as

$$\begin{bmatrix}\mathcal{H}_{L+1}(u^d_{[0,N+L-1]})\\ \mathcal{H}_L(y^d_{[0,N+L-2]})\end{bmatrix} \xrightarrow[\text{Construction}]{\text{Extended State}} \begin{bmatrix}\mathcal{H}_1(u^d_{[L,N+L-1]})\\ \mathcal{H}_1(\chi^d_{[L,N+L-1]})\end{bmatrix}$$

However, in the presence of noise entering the output data rows, all pL output data rows become (numerically) linearly independent. Therefore, all the rows of the maximal extended state are linearly independent, leading to the condition, rank $(\mathcal{H}_1(\chi^d_{[L,N+L-1]})) = mL + pL$. □

This means that even though the data informativity rank condition (2.17) on the input-extended-state data does not hold as in the case of noise-free data, it is only due to a completely different reason of numerical independency. This means that even after knowing that the maximal extended state cannot be applied for control and necessitates state reduction, the proposed state reduction methodology in Theorem 2.4 cannot be directly applied.

It is important to understand that the extended state still contains all the required information for controller synthesis, but that information is masked due to the presence of noise. Furthermore, it is also known that the extended state contains unobservable data rows that have to be eliminated. This motivates the thought that the collected input-output data is informative for control using the extended state framework, but a suitable methodology has to be developed to reduce the state while accounting for the noise. If such a reduced state can be developed, then it can be hypothesized that the extended state framework can be applied for control using noisy input-output data.

3.2.4 State Reduction

Considering the state reduction in the case of noise-free data, the dimension of the reduced state is determined by the rank condition (2.18). This cannot be done in the case of noisy data as the data rows are linearly independent due to the presence of noise. Therefore, an alternative state reduction methodology is required which accounts for the noise, which is proposed in Algorithm 3.

Algorithm 3 State Reduction - SVD

Require: Maximal extended state data Hankel data matrix $\mathcal{H}_1(\chi^d)$, (approximate) reduced-order n_{sys}

1. Perform a singular value decomposition (SVD) $\mathcal{H}_1(\chi^d) = U_f S_f V_f^\top$.
2. Take $U_r = U_f(:, 1 : n_{red})$ and $S_r = S_f(1 : n_{red}, 1 : n_{red})$.
3. Create transformation matrix $\Theta = S_r^{\frac{-1}{2}} U_r \in \mathbb{R}^{n_{red}\times(m+p)L}$ to arrange data matrix $\mathcal{H}_1(\chi^d)$ such that $\mathcal{H}_1(\chi_r^d) = \Theta\mathcal{H}_1(\chi^d)$

return Reduced-order data matrix $\mathcal{H}_1(\chi_r^d)$, transformation matrix Θ

The main difference as compared to Algorithm 2 used for noise-free data is the selection of the (approximate) reduced-order n_{sys} which is not given by a definite condition, but selected based on the data as a cut-off of the singular values of $\mathcal{H}_1(\chi^d)$. The selection of the reduced-order n_{red} based on the data is explained in the following example.

Example 3.1. Consider a randomly generated and unknown DT LTI system of the form (3.9) consisting of multiple inputs and outputs, and acting as the data-generating system. The system is defined with complexity $(n, m, \ell) = (3, 2, 2)$ and $p = 2$. The noisy input-output data set $\mathcal{D}_N = \{u_k^d, y_k^d\}_{k=0}^{N+L}$ with $N = 1000$ is collected with a Signal-to-Noise ratio (SNR) of 30.98 dB. This data dictionary is used to define the input-output Hankel data matrix $\mathcal{H}_L$ of depth $L = 6$, such that

$$\mathcal{H}_L = \begin{bmatrix} \mathcal{H}_L(u^d) \\ \mathcal{H}_L(y^d) \end{bmatrix}.$$

The singular values (σ) of $\mathcal{H}_L$ are computed and are plotted on a logarithmic scale in Figure 3.1. It is clearly visible that there is a sharp decrease at the 15^{th} singular value, with the magnitude of the 16^{th} singular value being quite low. This leads to the estimation of the reduced-order as $n_{red} = 15$. When being compared with the desired condition using system knowledge (i.e., knowing that n =3), the desired order is obtained as $n_{des} = mL + n = 2 \cdot 6 + 3 = 15$, which matches the approximate order estimated from data. This implies that the proposed algorithm can be used to reduce the maximal extended state to an approximate reduced-order which is informative based on results discussed in the previous chapter. □

Of course, a suitable methodology still needs to be considered for utilizing the reduced state affected by the noise. This is discussed in the following sections, and the validity of the designed reduced state towards control applications is discussed via simulation

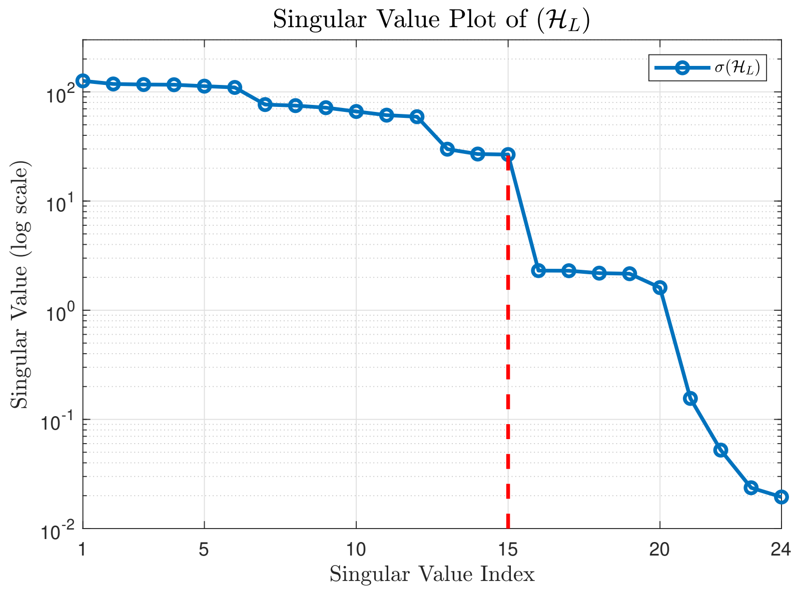


Figure 3.1: Singular Value Plot

examples. In applications with significant noise (characterized by a low signal-to-noise ratio (SNR)), it is possible that the reduced-order n_{red} might be inaccurately estimated and might require a trial-and-error methodology for successful implementation. However, this is not a drawback or a disadvantage but rather a fundamental characteristic while dealing with noisy data.

3.3 Extended State Reduction

The methodology for reducing the maximal extended state has been discussed along with the data informativity analysis. Therefore, the state reduction can be used for obtaining a data-based system representation which is necessary for data-driven control

3.3.1 Reduced State Space Representation

The maximal extended state $\chi_r(k)$ can be reduced based on Algorithm 3 to obtain the reduced state $\chi_r(k)$, defined by the relation $\chi_r = \Theta\chi$. Hence, a reduced-order state space representation can be formulated as

$$\begin{aligned} \chi_r(k+1) &= \mathcal{A}_r\chi_r(k) + \mathcal{B}_r u(k) + \mathcal{V}_r v(k), \\ y(k) &= \mathcal{C}_r\chi_r(k) + \mathtt{B}_0 u(k) + v(k), \end{aligned} \tag{3.15}$$

with the relations

$$\begin{aligned}
\mathcal{A}_r &= \Theta\mathcal{A}\Theta^\dagger = \Theta\Lambda_e\Theta^\dagger + \Theta J_1\Theta^\dagger = \Lambda_{er} + J_{1r},\\
\mathcal{B}_r &= \Theta\mathcal{B} = \Theta B_e + \Theta J_2 = B_{er} + J_{2r},\\
\mathcal{V}_r &= \Theta\mathcal{V},\\
\mathcal{C}_r &= \begin{bmatrix}\bar{A} & \bar{B}\end{bmatrix}\Theta^\dagger.
\end{aligned} \tag{3.16}$$

Let $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state trajectory of the reduced system (3.15) based on the measured noisy input-output data of the original data generating system (3.9). For simplicity, the collected data can be structured into the respective Hankel matrices with the following short-hand notations.

$$\begin{aligned}
U_- &= \mathcal{H}_1(u^d_{[L,N+L-1]}),\\
Y_- &= \mathcal{H}_1(y^d_{[L,N+L-1]}),\\
Z_- &= \mathcal{H}_1(\chi^d_{r[L,N+L-1]}),\\
Z_+ &= \mathcal{H}_1(\chi^d_{r[L+1,N+L]}),\\
V_- &= \mathcal{H}_1(v_{[L,N+L-1]}).
\end{aligned} \tag{3.17}$$

Note that the current input-output data sample is measured to define Z_+. Also, the matrix of noise data samples (V_-) is not measured. This means that (3.15) can be written in the form of data equations as

$$Z_+ = \mathcal{A}_r Z_- + \mathcal{B}_r U_- + \mathcal{V}_r V_-, \tag{3.18a}$$
$$Y_- = \mathcal{C}_r Z_- + \mathtt{B}_0 U_- + V_-. \tag{3.18b}$$

Since the input-output data is measured in the presence of noise, an assumption on the noise bounds is considered. This is motivated by the noise bounds discussed in the case of input-state data (subsection 3.1.2).

3.3.2 Cross-Covariance Noise Bonds

The noise assumption has been proposed in [19] to provide a general non-conservative noise characterization in terms of cross-covariance bounds. Consider the sample cross-covariance of the noise $v(k)$ with respect to a variable $r(k) \in \mathbb{R}^M$, given by

$$\frac{1}{\sqrt{(N+L+1)}}\Sigma_{k=0}^{N+L} v(k)r(k)^\top = \frac{1}{\sqrt{(N+L+1)}} V_- R_-^\top.$$

The variable $r(k)$ is selected as an instrumental variable, based on the prior knowledge of the noise assumed as

$$V_- R_-^\top R_- V_-^\top \preceq (N+L+1) H_u I_p. \tag{3.19}$$

for a scalar $H_u > 0$.

This noise assumption can be generalized in the form of a QMI as follows.

Assumption 3.2 (Noise Assumption[19]). *The noise samples collected in V_- satisfy the bound*

$$\begin{bmatrix} I \\ R_- {V_-}^\top \end{bmatrix}^\top \begin{bmatrix} Q_{11} & Q_{12} \\ {Q_{12}}^\top & Q_{22} \end{bmatrix} \begin{bmatrix} I \\ R_- {V_-}^\top \end{bmatrix} \succeq 0, \tag{3.20}$$

for known matrices $Q_{11} = {Q_{11}}^\top$, Q_{12} and $Q_{22} = {Q_{22}}^\top \prec 0$.

Similar to Assumption 3.1, since Q_{22} is negative definite, the noise matrix V_- will satisfy the bound given by (3.20). Furthermore, considering the special case $Q_{11} = (N + L + 1)H_u I_p \succ 0$, $Q_{12} = 0$ and $Q_{22} = -I \prec 0$, (3.20) can be reduced to the bounds given in (3.19).

Assumption 3.2 allows for the development of a methodology for data-driven control based on the collected noisy input-output data using the extended state framework. With the aim being the application of the Matrix S-lemma (Lemma 3.1), a data-based system parameterization in the form of a QMI has to be developed.

3.3.3 Data-driven System Representation

A data-based system representation has to be developed to ensure that the (unknown) system parameter matrices are compatible with the data generated from the system. Utilizing the partitioning (3.13) along with the relations (3.16), the set of all unknown system parameter matrices $(\Lambda_{er},\ B_{er})$ that are compatible with the data (U_-, Y_-) is defined as

$$\Sigma_{(U,Y)} := \{(\Lambda_{er}, B_{er}) \mid \exists\ V_- \text{ such that (3.18b) and (3.20) hold}\}.$$

Therefore, the data informativity of the constructed input-reduced-state trajectory $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ is shown in the following lemma.

Lemma 3.2 (Noisy Data-based System Representation [19]). *Let a matrix $\mathcal{Q}_{er}$ be defined as*

$$\mathcal{Q}_{er} := \begin{bmatrix} \mathcal{V}_r Q_{11} \mathcal{V}_r^\top & \mathcal{V}_r Q_{12} \\ Q_{12}^\top \mathcal{V}_r^\top & Q_{22} \end{bmatrix}, \tag{3.21}$$

such that

$$\begin{bmatrix} I \\ \Lambda_{er}^\top \\ B_{er}^\top \end{bmatrix}^\top \begin{bmatrix} I & \mathcal{V}_r Y_- R_- \\ 0 & -Z_- R_- \\ 0 & -U_- R_- \end{bmatrix} \mathcal{Q}_{er} \begin{bmatrix} I & \mathcal{V}_r Y_- R_- \\ 0 & -Z_- R_- \\ 0 & -U_- R_- \end{bmatrix}^\top \begin{bmatrix} I \\ \Lambda_{er}^\top \\ B_{er}^\top \end{bmatrix} \succeq 0, \tag{3.22}$$

holds. Then, the set of all pairs of (Λ_{er}, B_{er}) (3.16) are compatible with the constructed data dictionary $\mathcal{D}_{Nr}$.

Proof. Consider the Assumption 3.2, characterizing the sample cross-covariance bounds on the noise. The equation (3.20) is firstly pre-multiplied by $\mathcal{V}_r$ and post-multiplied by

$\mathcal{V}_r^\top$, to introduce the model-based terms Λ_{er} and B_{er} at a later step

$$\begin{aligned}
&\mathcal{V}_r \begin{bmatrix} I \\ R_- V_-{}^\top \end{bmatrix}^\top \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}{}^\top & Q_{22} \end{bmatrix} \begin{bmatrix} I \\ R_- V_-{}^\top \end{bmatrix} \mathcal{V}_r^\top \succeq 0 \\
\implies & \begin{bmatrix} \mathcal{V}_r Q_{11} + \mathcal{V}_r V_- R_-^\top Q_{12}^\top & \mathcal{V}_r Q_{12} + \mathcal{V}_r V_- R_-^\top Q_{22}^\top \end{bmatrix} \begin{bmatrix} I \\ R_- V_-{}^\top \end{bmatrix} \mathcal{V}_r^\top \succeq 0 \\
\implies & \begin{bmatrix} I \\ R_- (\mathcal{V}_r V_-)^\top \end{bmatrix}^\top \begin{bmatrix} \mathcal{V}_r Q_{11} \mathcal{V}_r^\top & \mathcal{V}_r Q_{12} \\ Q_{12}{}^\top \mathcal{V}_r^\top & Q_{22} \end{bmatrix} \begin{bmatrix} I \\ R_- (\mathcal{V}_r V_-)^\top \end{bmatrix} \succeq 0
\end{aligned} \tag{3.23}$$

Consider the output data equation (3.18b), which can be rewritten as a noise data equation

$$\mathcal{V}_r V_- = \mathcal{V}_r Y_- - \mathcal{V}_r \mathcal{C}_r Z_- - \mathcal{V}_r \mathtt{B}_0 U_-. \tag{3.24}$$

Notice that $\Lambda_{er} = \mathcal{V}_r \mathcal{C}_r$ and $B_{er} = \mathcal{V}_r \mathtt{B}_0$, which are the unknown model-based system terms The noise data equation (3.24) is substituted back in (3.23)

$$\begin{bmatrix} I \\ R_- (\mathcal{V}_r Y_- - \Lambda_{er} Z_- - B_{er} U_-)^\top \end{bmatrix}^\top \begin{bmatrix} \mathcal{V}_r Q_{11} \mathcal{V}_r^\top & \mathcal{V}_r Q_{12} \\ Q_{12}{}^\top \mathcal{V}_r^\top & Q_{22} \end{bmatrix} (*) \succeq 0$$

which can be rearranged to obtain the required representation (3.22). $\square$

This is a data-based parameterization of the set $\Sigma_{(U_-, Y_-)}$ utilizing the output data equation (3.18b), which means that all the systems that can be characterized by the collected input-output data (U_-, Y_-) can be represented by (3.22). As discussed in [19], the state data equation (3.18a) can also be used to obtain a data-based representation, but it is numerically sensitive due to repeated input-output data while considering the data matrices Z_- and Z_+.

To complete the discussion about the data-based system representation, it is paramount to explore the term R_-, which is the data matrix of the instrument variable $r(k)$, which is essential for handling the measurement noise.

3.4 Noise Handling using Instrumental Variables

An *instrumental variable* (IV) $r(k) \in \mathbb{R}^M$ is defined in system identification [31] as a variable used for the identification and estimation of discrete-time stochastic models. The main advantage of using an instrumental variable is the reduction of the asymptotic bias in the estimate (as the number of samples increase) while ensuring a small algorithmic complexity. Existing guidelines [30, 32] mention that the instrumental variable estimator will provide a consistent estimate of the system parameters under the following two conditions.

- The selected variable $r(k)$ is highly correlated with the input $u(k)$ and output $y(k)$ samples.

- The selected variable $r(k)$ is uncorrelated with the noise $v(k)$ samples.

This suggests the choice of (filtered or delayed versions) of the input and/or output sequence for data collection and IV generation.

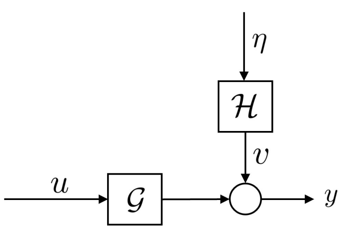
Ideally, the instruments are chosen to ensure a high correlation with the noise-free system input and output variables. However, real system measurements are always affected by noise. Furthermore, unstable systems, or systems that are unsafe to operate, require a stabilizing controller to collect data. Hence, extending it to practical IVs to be used in real systems or in a closed-loop setting requires modifications [33].

Initially, an open-loop data collection setting is considered. A *double experiment*, i.e., two back-to-back experiments are performed with the same conditions and the same input signal (u_1^d) but with different noise realizations, is conducted. The first set of collected IO data (u_1^d, y_1^d) is used to generate the instrumental variable, while the second dataset (u_1^d, y_2^d) is used for control (by creating the reduced-state data matrix). In the case of a closed-loop setting, the input disturbance signal (d_1^d) is kept the same, and the same double experiment is performed to collect two datasets, the first (d_1^d, u_1^d, y_1^d) for IV generation and the second (d_1^d, u_2^d, y_2^d) for data-driven control.

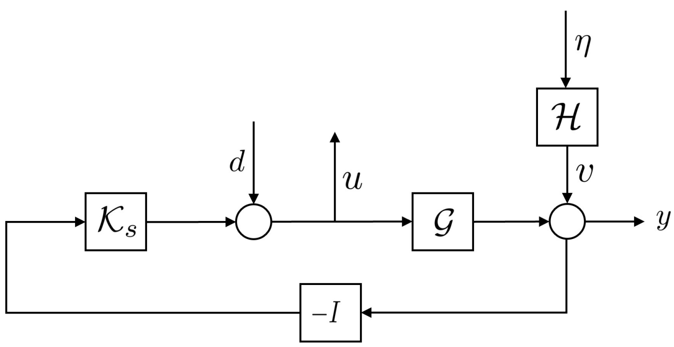
Consider the Box-Jenkins model structure (3.10) used for data generation. The primary challenge here is that the output $y(k)$ is affected by both the input $u(k)$ and the noise $v(k)$ samples and the input itself can be correlated with the noise samples. This is solved by the IV method using the double experiment, which will ensure a high correlation of the IV with the input-output data samples but reduce the correlation with the noise samples, therefore leading to a reduction in the bias in the controller design with an increase in the number of samples. Hence, the instrument can be formally defined as

$$r(k) = \mathrm{col}\begin{bmatrix} u_1^d(k) & \cdots & u_1^d(k - \frac{M-1}{2}) & y_1^d(k) & \cdots & y_1^d(k - \frac{M-1}{2}) \end{bmatrix}, \tag{3.25}$$

with M being the length of the instrument. The open-loop and closed-loop control architecture of the data-generating system used to obtain the instrument as well as the data dictionary $\mathcal{D}_{N_r}$ is shown in Figure 3.2.



(a) Open loop setting



(b) Closed loop setting

Figure 3.2: Data Generation for IV and Control

Lemma 3.2 provides an additional guideline on the selection of the length of the instrumental variable $r(k)$, i.e., the matrix $R_- \left[Z_-^\top U_-^\top\right]$ is required to have full column rank, and it will do so only if the number of instrumental variables M satisfies $M \geq m(\ell+1)+n$. For real systems, the value of lag ℓ and the order n is unknown. However, considering that this is just a lower limit on the length of the instrument, a practical IV can be taken as (3.25) with the length $M \geq n_{red}$ without facing any numerical errors.

Remark 3.1. Even though the noise samples are different, since they are drawn from the noise process, they can have a degree of correlation. This can be addressed in future work by introducing another instrument that allows reduced correlation.

Now, a data-based system representation is obtained along with a complete definition of the instrumental variable. This knowledge of the system can be used to formulate a static-state feedback controller of the form $u = \mathcal{K}_r \chi_r$ (2.30), which can stabilize the reduced-order system (3.15). By the argument of the reduced-state being informative with respect to the original system (3.9), an equivalent stabilizing dynamic output-feedback controller can be developed.

3.5 Synthesis for Stabilization

Motivated by the input-state methodology, the aim is to utilize the Matrix S-lemma to establish a relationship between two QMIs, i.e., the QMI for data-based system representation (3.22) and a QMI establishing quadratic stability. For obtaining the latter, consider the notion of quadratic stabilization (Definition 2.3), and the corresponding stability condition (2.10). The relation (3.16) can be used to rewrite it as follows.

$$\begin{aligned} &P - (\mathcal{A} + \mathcal{B}\mathcal{K})P(\mathcal{A} + \mathcal{B}\mathcal{K})^\top \succ 0, \\ \implies &P - (\Lambda_{er} J_{1r} + B_{er}\mathcal{K} + J_{2r}\mathcal{K})P(\Lambda_{er} J_{1r} + B_{er}\mathcal{K} + J_{2r}\mathcal{K})^\top \succ 0. \end{aligned} \tag{3.26}$$

Since the aim is to get a QMI with the form

$$\begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix}^\top \Pi_{stab} \begin{bmatrix} I \\ \Lambda_e^\top \\ B_e^\top \end{bmatrix} \succ 0, \tag{3.27}$$

the terms of (3.26) can be rearranged to obtain (3.27) with the matrix Π_{stab} given by (3.28). Therefore, the problem of the existence of $\mathcal{K}$ and $P \succ 0$ so that (3.27) holds true can be now solved using the Matrix S-lemma, which leads to the following theorem about data-based quadratic stabilization of the system (3.15).

Theorem 3.4 (Stabilization for Noisy data [19])**.** *Consider the system* (2.23) *and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (3.9)*. Then the data $\mathcal{D}_{Nr}$ is said to be informative for quadratic stabilization by a static reduced-state feedback controller* (2.30) *if there exist matrices $P = P^\top \succ 0$ and E, and scalars $\alpha \geq 0$ and $\beta > 0$ such that* (3.29) *holds true. For such informative data, the controller $u = \mathcal{K}_r \chi_r$ can be designed with its coefficients given by $\mathcal{K}_r = EP^{-1}$.*

Proof. Applying the Matrix S-lemma to (3.22) and (3.27) results in

$$\Pi_{stab} - \alpha\Lambda \succeq \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix},$$

$$\Pi_{stab} = \begin{bmatrix} P - (J_{1r} + J_{2r}\mathcal{K})P(J_{1r} + J_{2r}\mathcal{K})^\top & -(J_{1r} + J_{2r}\mathcal{K})P & (J_{1r} + J_{2r}\mathcal{K})P\mathcal{K}^\top \\ -P(J_{1r} + J_{2r}\mathcal{K})^\top & -P & -P\mathcal{K}^\top \\ -\mathcal{K}P(J_{1r} + J_{2r}\mathcal{K})^\top & -\mathcal{K}P & -\mathcal{K}P\mathcal{K}^\top \end{bmatrix} \tag{3.28}$$

$$\begin{bmatrix} P - \beta I & -J_{1r}P - J_{2r}E & 0 & J_{1r}P + J_{2r}E \\ -PJ_{1r}^\top - E^\top J_{2r}^\top & -P & -E^\top & 0 \\ 0 & -E & 0 & E \\ PJ_{1r}^\top + E^\top J_{2r}^\top & 0 & E^\top & P \end{bmatrix} - \alpha \begin{bmatrix} I & \mathcal{V}_r Y_- R_-^\top \\ 0 & Z_- R_-^\top \\ 0 & -U_- R_-^\top \\ 0 & 0 \end{bmatrix} \mathcal{Q}_{er} (*)^\top \succeq 0 \tag{3.29}$$

---

where

$$\Lambda = \begin{bmatrix} I & \mathcal{V}_r Y_- \\ 0 & -Z_- \\ 0 & -U_- \end{bmatrix} \mathcal{Q}_{er} \begin{bmatrix} I & \mathcal{V}_r Y_- \\ 0 & -Z_- \\ 0 & -U_- \end{bmatrix}^\top$$

Hence, by standard techniques of Schur complement, the above equation can be equivalently written as (3.29) with a symmetric $P \succ 0$, E, $\alpha \geq 0$, $\beta > 0$ and $\mathcal{K}_r := EP^{-1}$. □

Notice that the QMIs have been converted to an LMI which is affine with respect to the decision variables P, E, α, and β. The solution (value of the decision variables) can be obtained by defining an SDP and solving a feasibility problem.

Remark 3.2. It can be noticed that while the dimensions of the decision variables P and E depend on the depth (L) of the Hankel matrices, they are defined independently of the number of data samples being considered (N). This means that this method allows for a large dataset to be considered for controller synthesis and, hence, is a computationally attractive method.

Similar to the previous discussions about data-driven control, the framework can also be extended to design output-feedback controllers to ensure performance, as discussed in the following section.

3.6 Synthesis for Performance

It is known that a performance control problem is defined as finding a stabilizing reduced-state feedback controller (2.30) that achieves an $\mathcal{H}_\infty$/ $\mathcal{H}_2$ performance bound from the measured input-output data.

3.6.1 Closed-loop System

Consider a known performance output $z(k)$ such that $z(k) = C_z\chi(k) + D_z u(k)$. For any pair $(\mathcal{A}_r, \mathcal{B}_r)$, this results in a closed-loop system

$$\begin{aligned} \chi_r(k+1) &= (\mathcal{A}_r + \mathcal{B}_r\mathcal{K}_r)\chi_r(k) + \mathcal{V}_r v(k) \\ z(k) &= (C_z + D_z\mathcal{K}_r)\chi_r(k) \end{aligned} \tag{3.30}$$

Hence, the transfer matrix from $v(k)$ to $z(k)$ is given by,

$$T(z) = (C_z + D_z\mathcal{K}_r)(zI - \mathcal{A}_r - \mathcal{B}_r\mathcal{K}_r)^{-1}\mathcal{V}_r, \tag{3.31}$$

for which the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms are defined as $||T||_\infty$ and $||T||_2$ respectively.

Remark 3.3. It can be seen here that the closed-loop system is effectively defined from the noise $v(k)$ to the reduced extended state $\chi_r(k)$. By choosing C_z and D_z appropriately (as $I_{p\times n_{sys}}$ and $0_{p\times m}$ respectively), the loop transfer is defined from the noise $v(k)$ to the lagged output $y(k-1)$. However, the preferred definition is to establish the system output as the performance output. The main limitation here is that the output has to be directly related to the extended state. Specifically, defining the performance output as $y(k)$ would require unknown (data-based) performance matrices C_z and D_z, and this cannot be effectively handled by this methodology. The reason for this is that the data-based matrix C_z is present in a second LMI constraint, where it should instead act as a known matrix. Furthermore, notice that there is no feedthrough term in the performance output. This is due to the existing limitations of the methodology, as discussed in Appendix A. Hence, it remains a future work to introduce the required modifications.

3.6.2 $\mathcal{H}_\infty$ control using Noisy Data

$\mathcal{H}_\infty$ control refers to ensuring a desired $\mathcal{H}_\infty$ norm performance guarantee for the closed-loop system (3.30) under the worst-case signal behaviors. Similar to the previous case of quadratic stabilization, the Matrix S-lemma can be used to define the necessary informativity conditions for $\mathcal{H}_\infty$ control based on Definition 2.35.

Theorem 3.5 (Data-based $\mathcal{H}_\infty$ control [19])**.** *Consider the system* (3.30) *and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (3.9)*. Then the data $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_\infty$ control* ($||T||_\infty < \gamma$) *with performance γ using a static reduced-state feedback controller* (2.30) *if there exist matrices $P = P^\top \succ 0$ and E, and scalars $\alpha \geq 0$ and $\beta > 0$, such that* (3.32) *and* (3.33) *hold true with $F = (C_z P + D_z E)$. For such informative data, the controller $u = \mathcal{K}_r\chi_r$, which achieves the performance γ, can be designed with its coefficients given by $\mathcal{K}_r = EP^{-1}$.*

Proof. The proof follows from (3.16) and Theorem 2 from [19]. □

It can be observed that the condition (3.32) is affine with respect to P, E, α, and β, but not with respect to γ. Hence, a bisection-based search is performed to obtain a feasible value of γ.

$$\begin{bmatrix} P-\beta I-\gamma\mathcal{V}_r\mathcal{V}_r^\top & 0 & 0 & J_1P+J_2E & 0 \\ 0 & 0 & 0 & P & 0 \\ 0 & 0 & 0 & E & 0 \\ P^\top J_1^\top + L^\top J_2^\top & P & E^\top & P & F^\top \\ 0 & 0 & 0 & F & \gamma I \end{bmatrix} - \alpha \begin{bmatrix} I & \mathcal{V}_rY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix} Q_{er}(*)^\top \succeq 0 \tag{3.32}$$

$$\begin{bmatrix} P & F^\top \\ F & \gamma I \end{bmatrix} \succ 0 \tag{3.33}$$

$$\begin{bmatrix} P-\beta I & 0 & 0 & J_{1r}P+J_{2r}E & 0 \\ 0 & 0 & 0 & P & 0 \\ 0 & 0 & 0 & E & 0 \\ P^\top J_{1r}^\top + E^\top J_{2r}^\top & P & E^\top & P & F^\top \\ 0 & 0 & 0 & F & I \end{bmatrix} - \alpha \begin{bmatrix} I & \mathcal{V}_rY_-R_-^\top \\ 0 & -Z_-R_-^\top \\ 0 & -U_-R_-^\top \\ 0 & 0 \\ 0 & 0 \end{bmatrix} Q_{er}(*)^\top \succeq 0 \tag{3.35}$$

3.6.3 $\mathcal{H}_2$ control using Noisy Data

$\mathcal{H}_2$ control refers to ensuring a desired $\mathcal{H}_2$ norm performance guarantee for the closed-loop system under an average signal behavior. Definition 2.38 is used to determine a QMI for $\mathcal{H}_2$ control analysis, and then the Matrix S-lemma is used to define the informativity conditions as follows.

Theorem 3.6 (Data-based $\mathcal{H}_2$ control [19]). *Consider the system* (3.30) *and let the data dictionary $\mathcal{D}_{Nr} = \{u_k^d, \chi_{rk}^d\}_{k=L}^{N+L}$ be a constructed input-reduced-state data set of the system based on the measured input-output data of the original data generating system* (3.9). *Then the data $\mathcal{D}_{Nr}$ is said to be informative for $\mathcal{H}_2$ control ($||T||_2 < \gamma$) with performance γ using a static reduced-state feedback controller* (2.30) *if there exist matrices $P = P^\top \succ 0$, $Z = Z^\top$ and E, and scalars $\alpha \geq 0$ and $\beta > 0$, such that* (3.35) *and*

$$\operatorname{trace}(Z) < \gamma^2, \quad \begin{bmatrix} P & F^\top \\ F & I \end{bmatrix} \succ 0 \quad \text{and} \quad \begin{bmatrix} Z & \mathcal{V}_r^\top \\ \mathcal{V}_r & P \end{bmatrix} \succeq 0, \tag{3.34}$$

hold true with $F = (C_zP + D_ZE)$. For such informative data, the controller $u = \mathcal{K}_r\chi_r$, which achieves the performance γ, can be designed with its coefficients given by $\mathcal{K}_r = EP^{-1}$.

Proof. The proof follows from (3.16) and Theorem 3 from [19]. □

It can be observed that the conditions (3.35) and (3.34) are affine with respect to the decision P, Z, E, α, and β, which helps in finding the solution of the decision variables by solving an appropriately defined SDP for minimizing the value of γ.

These results establish the reduced-state-based methodologies for stabilization and performance while considering noisy input-output data. Their implementation and validity are examined by considering simulation examples as discussed in the next section.

3.7 Simulation Examples

This section discusses the validation of the extended state framework developed while considering noisy input-output data. The specifics of data generation and collection have been well established for the case of noise-free data. To introduce noise in the system, and specifically measurement noise, the data-generating system of the form (3.10) is considered. The experiment settings described earlier (in 2.7.1) are considered, which are briefly re-iterated for completeness, along with the introduction of noise.

3.7.1 Experiment Setting

The noisy data-generating system with a Box-Jenkins model structure consists of the true system $\mathcal{G}_D$ (3.36) defined by the complexity $(n, m, \ell) = (3, 2, 2)$ with $p = 2$, and a (stable and inversely stable) noise model $\mathcal{H}_D$ (3.37) which is randomly generated with desired dimensions and a fixed seed. The full system is depicted in (3.38).

$$\begin{aligned} x(k+1)_{\mathcal{G}_D} &= \begin{bmatrix} 0.850 & -0.038 & -0.380 \\ 0.735 & 0.815 & 1.594 \\ -0.664 & 0.697 & -0.064 \end{bmatrix} x(k)_{\mathcal{G}_D} + \begin{bmatrix} 1.431 & 0.705 \\ 1.620 & -1.129 \\ 0.913 & 0.369 \end{bmatrix} u(k), \\ y(k)_{\mathcal{G}_D} &= \begin{bmatrix} 0.645 & 1.118 & -1.169 \\ -1.465 & -0.005 & 0.586 \end{bmatrix} x(k)_{\mathcal{G}_D}. \end{aligned} \tag{3.36}$$

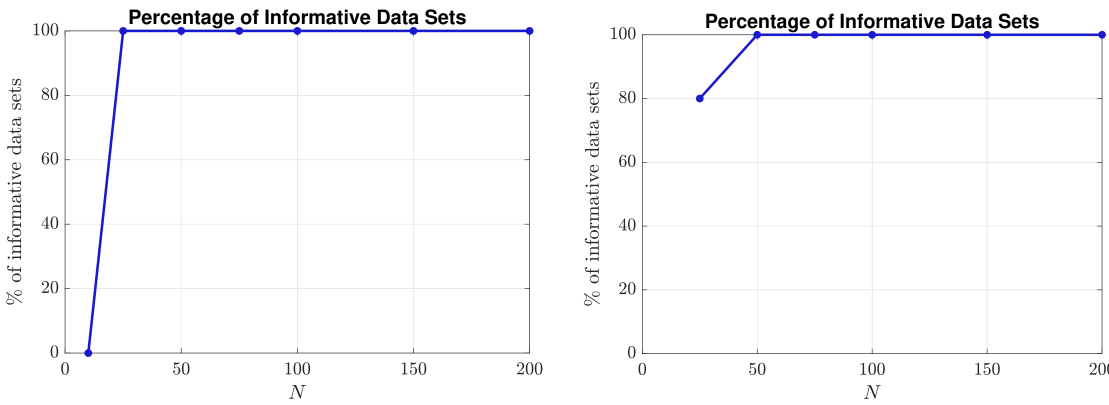
$$\begin{aligned} x(k+1)_{\mathcal{H}_D} &= \begin{bmatrix} 0.154 & 0.149 & -0.039 \\ 0.004 & 0.099 & 0.152 \\ 0.126 & 0.045 & 0.034 \end{bmatrix} x(k)_{\mathcal{H}_D} + \begin{bmatrix} 0.088 & 0.004 \\ 0.685 & 0.512 \\ 0.953 & 0.812 \end{bmatrix} \eta(k), \\ v(k) &= \begin{bmatrix} 0.612 & 0.291 & 0.714 \\ 0.721 & 0.917 & 0.542 \end{bmatrix} x(k)_{\mathcal{H}_D}. \end{aligned} \tag{3.37}$$

$$y(k) = \mathcal{G}_D u(k) + \mathcal{H}_D \eta(k) = y(k)_{\mathcal{G}_D} + v(k). \tag{3.38}$$

The number of data samples N has varying values taken as $N \in [10, 25, 50, 75, 100, 150, 200]$. For each value of N, 10 Monte Carlo simulation runs are performed in the closed-loop setting. The input disturbance $(d(k))$ consists of samples drawn such that $d_k^d \sim \mathcal{N}(0, 1)$. The input noise $\eta(k)$ is randomly drawn so that the signal-to-noise ratio (SNR) is around 31.1 dB, along with $H_u = 10$ to ensure that the noise assumptions are satisfied. The defined experiment settings are considered for the following examples.

3.7.2 Example - Data Informativity for Quadratic Stabilization

Consider the system (3.38) with the discussed experiment settings. Theorem 3.4 for quadratic stabilization is implemented, along with Algorithm 3 for state reduction with the selection of an appropriate reduced order from the data, as discussed in Example 3.1. The data informativity percentage is shown in Figure 3.3 for varying values of the depth L, as $L = 4$ and $L = 9$ with the reduced order considered as $n_{red} = 11$ and $n_{red} = 21$ respectively. For the second case, it is not possible to consider $N = 10$ due to the minimum requirement on the number of data samples, so it has been neglected.



(a) $L = \ell + 2 = 4$

(b) $L = \ell + 7 = 9$

Figure 3.3: Stabilization for MIMO systems

The main observation is that the proposed state reduction methodology is effective in reducing the dimension of the maximal extended state, enabling it to be utilized for control. Furthermore, the considered instrumental variable accounts for the measurement noise and allows the synthesis of stabilizing controllers.

It can be seen that for $N \geq 50$ all the data is always informative for control. This matches with the intuition from classical system identification that a larger dataset helps to capture the system dynamics and, hence, can be correlated to direct data-driven control by the argument that a larger dataset helps in reducing the effect of noise on data and helps in controller synthesis.

For the particular case of $N = 200$ and $L = 4$, the controller is synthesized offline and is implemented on `Simulink` as shown in Figure 3.4. The system is simulated with random initial conditions, and is subjected to a unit step signal at time instant $k = 2$. When subjected to noise, the resulting input and output trajectories are shown in Figure 3.5, validating the closed-loop stability of the system $\mathcal{G}_D$. When simulated in the absence of noise, the obtained response is shown in Figure 3.6, which shows that the synthesis process is also applicable to noise-free settings.

Remark 3.4. It is observed that increasing H_u reduces data informativity at a particular value of N, which can be attributed to the fact that H_u denotes the upper bound on the sample cross-covariance of the noise with the instrumental variable. A higher value of the cross-covariance implies a higher effect of noise on the data and hence reduces the utilization of data for control.

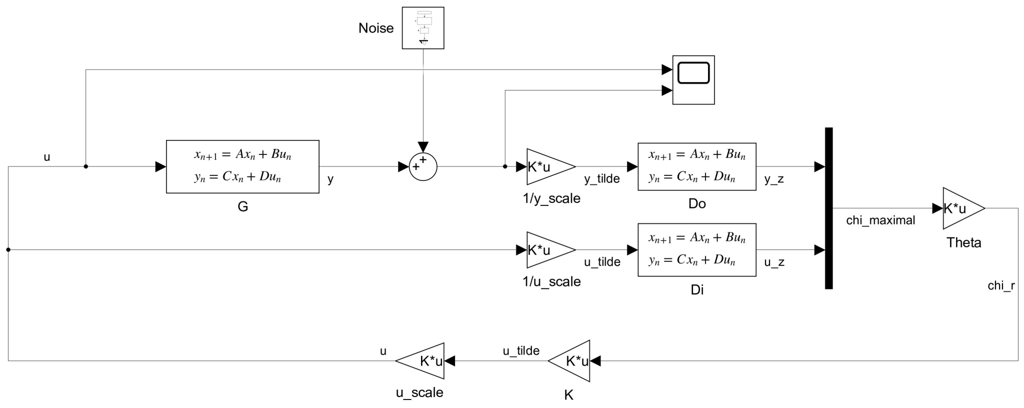


Figure 3.4: Controller Implementation - MIMO System

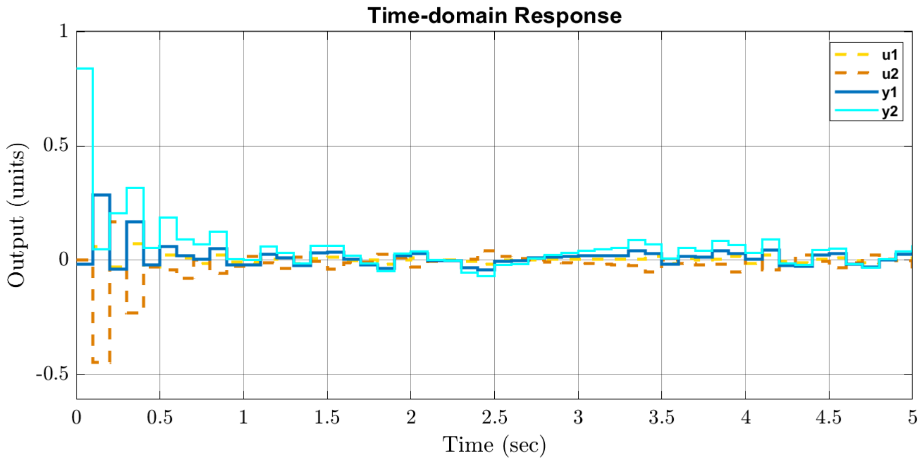


Figure 3.5: Time-domain Response for a MIMO System

3.7.3 Example - Solver Time

Consider the system (3.38) with the discussed experiment settings. Consider that L is varied while keeping N constant at $N = 200$. The computation times taken by the solver to obtain a feasible solution on the application of Theorem 3.4 measured for different cases of L are shown in Figure 3.7. This shows that the solver time increases with an increase in the value of L. This is expected as the dimensions (and hence the computation time) of the decision variables is dependent on the finite-horizon L.

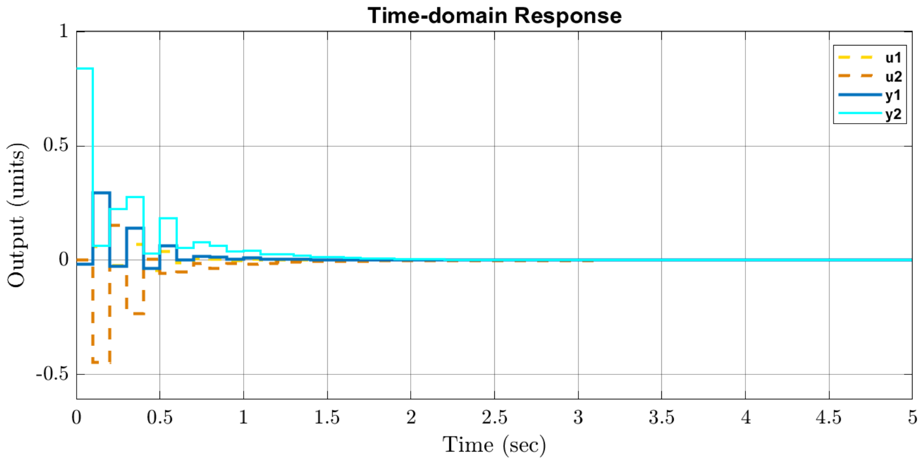


Figure 3.6: Time-domain Response for a MIMO System

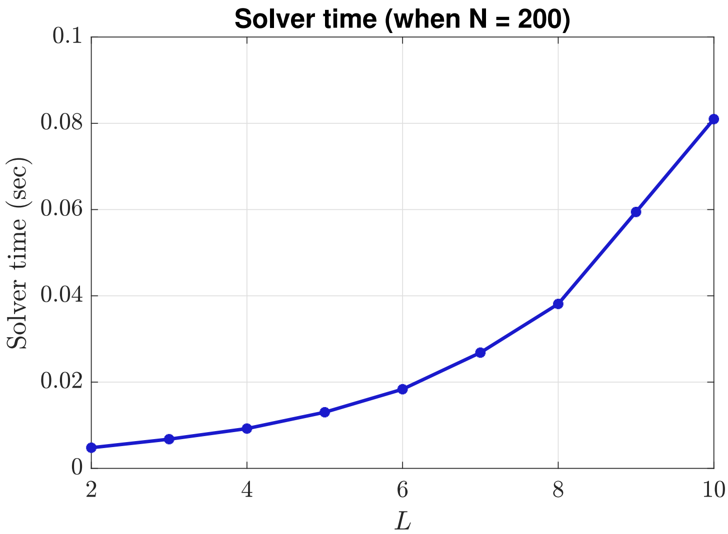


Figure 3.7: Solver time with $N = 200$

These examples show that the reduced-order extended state can be utilized for controller synthesis even in the presence of noise. The state reduction is performed according to the defined algorithm and results in a stabilizing controller, whose behavior is verified through its implementation as a dynamic output-feedback controller. Next, two examples are considered to study and validate the performance methodology.

3.7.4 Example - Data Informativity for $\mathcal{H}_\infty$ Control

Consider the system (3.38) with the discussed experiment settings. Having examined stabilization in the presence of noise, the same data-generating system is used for analyzing $\mathcal{H}_\infty$ performance with known performance matrices $C_z = I_{p\times n_{sys}}$ and $D_z = 0_{p,m}$. On the application of Theorem 3.5 with $H_u = 25$, the percentage of data informative for

$\mathcal{H}_\infty$ control is obtained and displayed in Figure 3.8.

Furthermore, the performance metric γ is found and is indicated by the light blue shaded region in Figure 3.9, which represents the data-based $\mathcal{H}_\infty$ norm, i.e., the value of γ obtained after the bisection search. Also, for $L = 3$, the designed controller is implemented on the closed-loop system, and the model-based $\mathcal{H}_\infty$ norm is calculated, as shown by the dark blue plot. It is seen that the data-based norm γ indeed acts as the upper bound on $||T(z)||_\infty$.

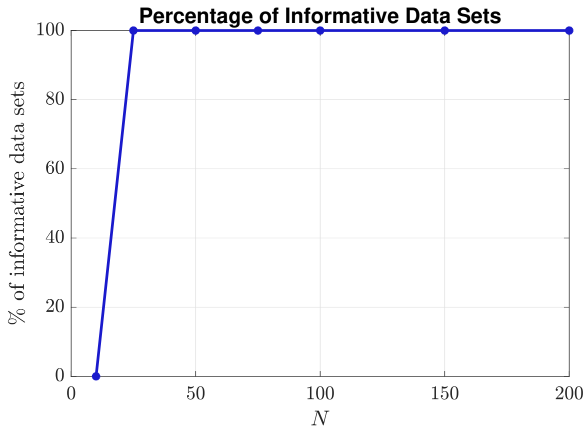
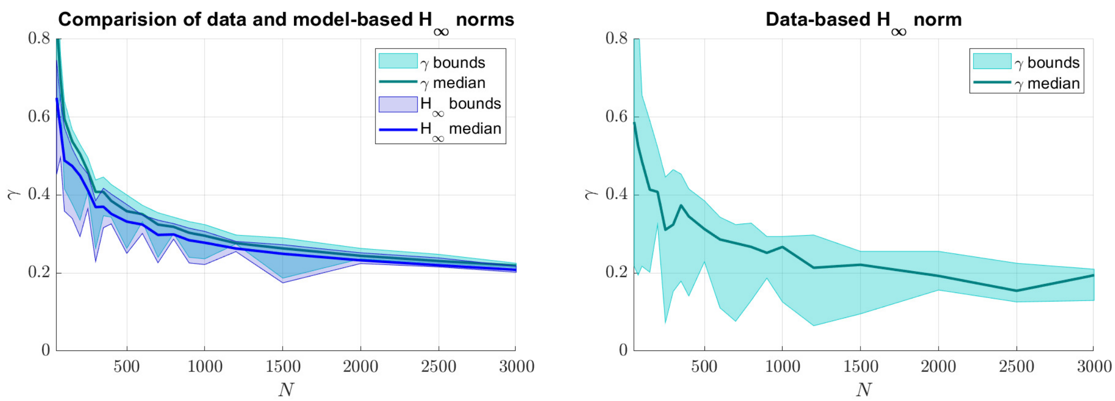


Figure 3.8: Data informativity of $\mathcal{H}_\infty$ control



(a) $L = \ell + 1 = 3$

(b) $L = \ell + 4 = 6$

Figure 3.9: $\mathcal{H}_\infty$ control of MIMO systems

The bisection search itself is performed with an initial stopping condition of 0.05, which resulted in numerical inaccuracies. On improving the stopping condition to 0.005, the displayed result is obtained with better accuracy of the data-based norm. However, the average runtime for each iteration increases from 3.2 seconds to 5 seconds due to the increased expectation of numerical precision.

It can be observed that there is a decrease in the value of γ with an increasing value of N. This is due to a larger number of data points increasing the accuracy of the methodology, i.e., more data points help in better capturing the system dynamics and minimizing the worst case energy gain.

3.7.5 Example - Data Informativity for $\mathcal{H}_2$ Control

Consider the system (3.38) with the discussed experiment settings. Consider the system $\mathcal{G}_D$ (2.41) with the previous settings as in $\mathcal{H}_\infty$ control. Taking $H_u = 25$, the SDP is solved for finding the decision variables, and hence, the controller coefficients. On the application of Theorem 3.6, the percentage of data informative for $\mathcal{H}_2$ control is obtained and displayed in Figure 3.10.

As previously discussed, the light blue plot shows the data-based $\mathcal{H}_2$ norm of the system. Also, for $L = 3$, the designed controller is implemented on the closed-loop system, and the model-based $\mathcal{H}_2$ norm can be calculated, as shown by the dark blue plot. It is seen that the data-based norm γ indeed acts as the upper bound on $||T(z)||_2$. Furthermore, it can be noticed that as the depth increases, there is a decrease in the magnitude of the norm. This can be attributed to the fact that increasing depth captures more system information and thus results in the minimization of the system norm. The computation time is measured to be around 0.6 seconds for $L = 3$ and has a similar trend as compared with the case of quadratic stabilization.

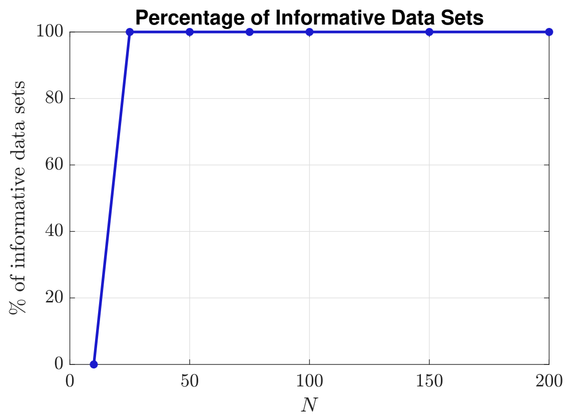
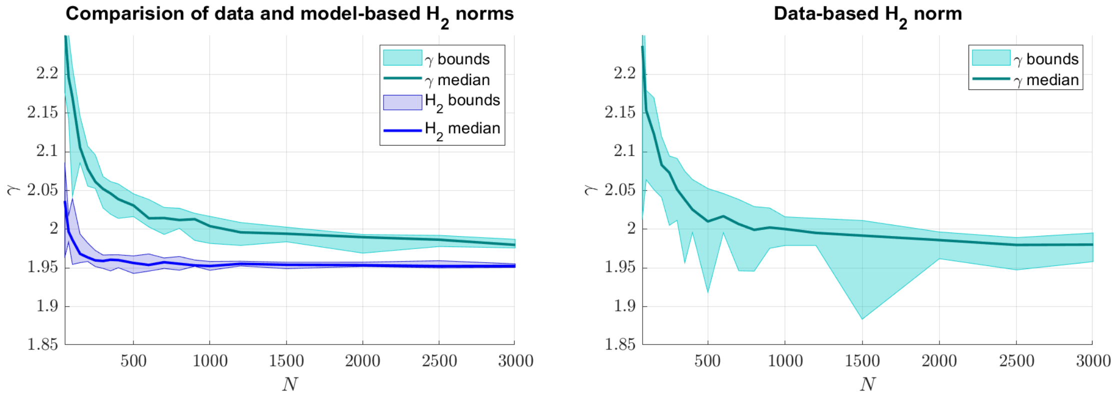


Figure 3.10: Data informativity of $\mathcal{H}_2$ control

These examples show the successful implementation of the extended state framework towards $\mathcal{H}_\infty$ and $\mathcal{H}_2$ control. This firmly establishes the applicability and validity of the extended state framework for noisy data. Also, the state reduction method is particularly applicable to performance. Therefore, the extended state framework has been developed to ensure stability as well as performance in the case of an unweighted generalized plant and in the presence of external measurement noise.



(a) $L = \ell + 1 = 3$ (b) $L = \ell + 4 = 6$

Figure 3.11: $\mathcal{H}_2$ control of MIMO systems

3.8 Implementation on a Real System

The extended state framework is versatile enough to account for noise, which motivates its application for data-driven control of a real system. The system is called PATO setup. PATO stands for 'Postacademisch Technisch Onderwijs' [34] and is used as a representation of a benchmark mass-spring-damper (MSD) system. It consists of two cylindrical masses connected to each other by a stiff rod, as shown in Figure 3.12. One of the masses is connected to a motor, which provides the input to the system, while both masses are connected to encoders, which are used to collect the output measurements. The system is monitored by a software called dSPACE, whose task is to scale the input signal to be provided to the motor and to collect the measured output from the encoders. The control architecture is designed in `Simulink`, which is integrated with the dSPACE environment to monitor and control the setup. Hence, the controller implementation and modifications can be done in `Simulink`, while the actual implementation is performed in the background.

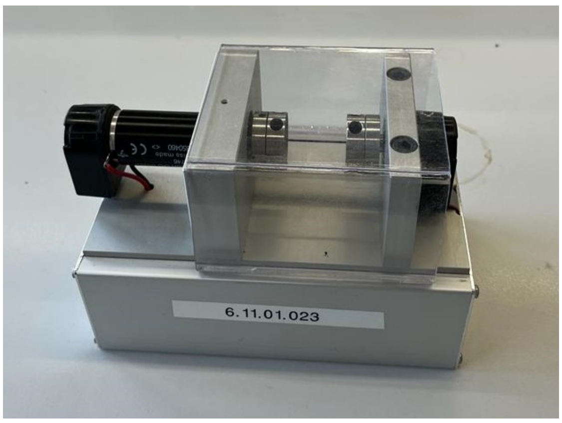


Figure 3.12: PATO Setup

The sampling frequency of the system is 2kHz. The input signal provided in `Simulink` is considered with appropriate magnitude such that it is then multiplied by a conversion factor of 1.37 in dSPACE to convert it to Volts to be provided as the input to the motor. The input to the system has a saturation range of ± 1.83 Volts, which is enforced in `Simulink`. Furthermore, to protect the system, a safety mechanism is implemented, which cuts off the supply to the motor when the angular velocity exceeds 400 rad/sec, which ensures the safe operation of the system. Since it is a real system, it is assumed that the measured output signals are contaminated with measurement noise by the encoders.

The system is assumed to be an LTI system so that the developed extended state framework can be utilized for control. Specifically, the following objective is defined.

Control Objective: Stabilize the angular velocity and ensure disturbance rejection.

The control architecture of the system is shown in Figure 3.13. A stabilizing controller is present, which allows for closed-loop data collection. The input disturbance (d) consists of samples drawn from a Gaussian white noise disturbance such that $d(k) \sim \mathcal{N}(0, 0.01)$. This signal is also measured, along with the input (u_k^d) and output (y_k^d) signals. The double experiment design to construct an IV requires that the same disturbance signal be measured. Therefore, a delay of 20 seconds is introduced in the `Simulink` model after it is compiled and built, which allows for enough time for dSPACE to start monitoring the system and recording the data. This means that the input disturbance to the system starts from the same data value in all experiments.

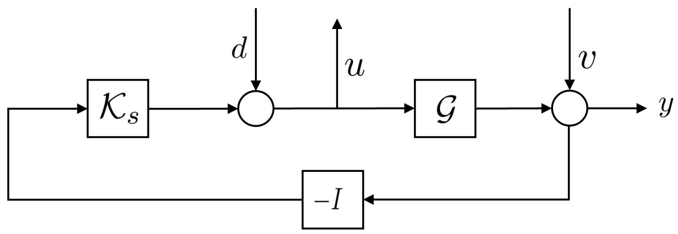
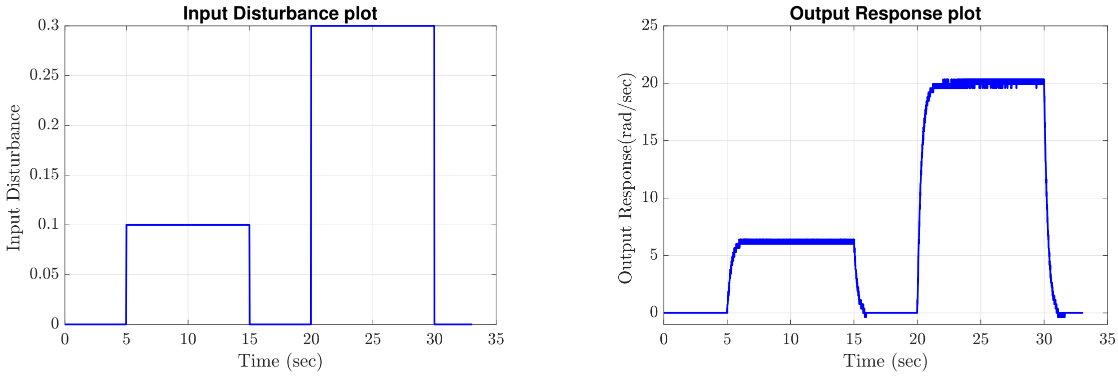


Figure 3.13: System Architecture

The system has an input signal with $m = 1$ and the angular velocity of the co-located mass is measured ($p = 1$). Two data sets are obtained after measurement for approximately 20 seconds. This results in datasets with each signal having 80000 data points. Initial 10000 points are discarded to account for the transient response of the system. Having obtained the data, the first dataset is used to create the instrumental variable, while the second dataset is structured into the desired input and output Hankel matrices. The stabilization procedure as discussed in Theorem 3.4 is followed. It is known that the setup can be represented by a parametric model of order 4. Based on this prior knowledge, the (upper-bound) on the lag of the system is taken $\ell^u = 4$. Therefore, the depth of the data matrices is considered to be $L = 4$.

It is initially observed that numerical errors are encountered while finding a feasible solution for the semi-definite program. To correct this, the measured data is scaled accordingly, with care taken to ensure that an unscaling operation is performed while implementing the controller. This ensures a reduced numerical sensitivity and results in a feasible SDP and, hence, the synthesis of a stabilizing controller.

To test the obtained controller, the existing model-based stabilizing controller is disconnected, and the static reduced-state feedback controller is implemented as a dynamic output-feedback controller, with the implementation similar to the one discussed in the examples. The disturbances are considered as two rectangular pulses of varying magnitudes and of 5 seconds duration. Both the input disturbance and the corresponding output response are measured and are shown in Figure 3.14.



(a) Input Disturbances

(b) Output Response

Figure 3.14: Response for Controller Implementation

It can be seen that the controller is able to stabilize the angular velocity of the real system even in the presence of external unknown noise and disturbances. This shows the effectiveness of the developed extended state framework and validates its effectiveness for real-life applications.

Chapter 4

Conclusions and Future Work

In this thesis, an extended state framework for direct data-driven control utilizing measured input-output data samples has been developed. This chapter summarizes the findings and addresses the raised research questions. Future recommendations are also presented here.

4.1 Conclusions

The research questions regarding the development of the framework have been raised in Chapter 1.3. These questions are addressed in this section, along with their proposed solutions and developments.

Research Question 1: How can a non-minimal extended state based on collected input-output data be constructed so as to render input-state data-driven control techniques feasible with minimal prior knowledge of the system?

The maximal extended is known to have rank deficiency according to the Fundamental Lemma. This is addressed by proposing a state-reduction methodology, such that only the linearly independent data rows are considered for control. The introduction of state reduction allows the existing input-state data-driven control methodologies to be utilized for input-output control.

Furthermore, this thesis explores the existing limitations of the extended framework, which are the applicability of the methodology to the cases of multiple outputs, unknown system lag, and unknown system order. Through the proposed algorithm(s) for state reduction, these limitations are addressed completely. The framework currently allows for multiple outputs by performing state reduction to reduce the dimension. Similarly, unknown system lag and system order are accounted for by considering a larger depth and then reducing to the necessary dimensions, which is a significant contribution to the developed framework. A salient point of the framework is that the reduction can be made purely on the basis of knowledge from the data and minimum prior knowledge.

Research Question 2: How to ensure minimum stability and performance loss during direct data-driven control using an extended state while considering general noise conditions?

Real systems are always affected by measurement noise. Firstly, the developed framework provides the required algorithm for state reduction to ensure that informative data can be accurately selected and separated from the data points significantly affected by noise. Furthermore, the dimension of the reduced state is identified based on the data and it can be used for control without any other requirements.

The effect of noise itself is mitigated by the introduction of instrumental variables, which allow for noise characterization without any asymptotic bias. The particular contribution includes the consideration of an instrumental variable constructed from data using multiple experiments, and instrument generation methodologies for open and closed-loop settings. These completely address implementation scenarios as well.

Research Question 3: How to develop tractable conditions for data-driven control using an extended state to ensure performance optimization?

A novel methodology focusing on performance optimization for noise-free data-driven control has been developed. This considers a generalized input and output signal, which is a first attempt at performance for the particular case of noise-free input-output data. Furthermore, existing methodologies for noisy input-output data are modified to consider state reduction, which ensures their applicability to a general MIMO system with minimum prior knowledge.

The discussions show that the considered research questions have been addressed in detail with the proposed solutions forming a backbone of the developed extended state framework. These are the significant contributions of this work focused on the applicability of an extended state for direct data-driven control. Of course, research is just a continuous path toward newer developments. Hence, there are always ideas for improving the existing framework and adding newer elements, some of which are highlighted in the next section.

4.2 Future Work

The recommendations for future research are mentioned here.

1. Weighted Generalized Plant - The most pressing research area is to develop a methodology to handle a weighted generalized plant and enforce desired specifications through a mixed sensitivity approach. One possible suggestion for this extension is to integrate the model-based weighting filter elements (in a state space representation) with the existing data-based system representation. Therefore, controller synthesis will include the designed filter specifications.

2. Reference Tracking - To extend the utilization of the developed framework, a method to consider an external reference signal in the extended state has to be formulated. This extends the framework to be applicable for reference tracking.

3. Iterative Instrumental Variable method - The instrument can be generated iteratively through simulation of the data-based system representation to obtain a new data set. So, a new instrument can be considered that has reduced correlation with the original noise samples.

4. Performance Shaping - In the case of DDD control for noisy input-output data, the general closed-loop system with the performance output as the system output should be considered, which ensures compatibility with existing performance-shaping techniques.

Bibliography

[1] L. Ljung, "System identification," in *Signal analysis and prediction.* Springer, 1998, pp. 163–173.

[2] V. Krishnan and F. Pasqualetti, "On direct vs indirect data-driven predictive control," in *2021 60th IEEE Conference on Decision and Control (CDC).* IEEE, 2021, pp. 736–741.

[3] H. J. Van Waarde, J. Eising, M. K. Camlibel, and H. L. Trentelman, "The informativity approach: To data-driven analysis and control," *IEEE Control Systems Magazine*, vol. 43, no. 6, pp. 32–66, 2023.

[4] ——, "The informativity approach: To data-driven analysis and control," *IEEE Control Systems Magazine*, vol. 43, no. 6, pp. 32–66, 2023.

[5] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.

[6] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 883–897, 2022.

[7] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[8] I. Markovsky, J. C. Willems, P. Rapisarda, and B. L. De Moor, "Data driven simulation with applications to system identification," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 970–975, 2005.

[9] J. Berberich and F. Allgöwer, "A trajectory-based framework for data-driven system analysis and control," in *2020 European Control Conference (ECC).* IEEE, 2020, pp. 1365–1370.

[10] H. J. Van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 602–607, 2020.

[11] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.

[12] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.

[13] J. Berberich, A. Koch, C. W. Scherer, and F. Allgöwer, "Robust data-driven state-feedback design," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1532–1538.

[14] P. Verheijen, V. Breschi, and M. Lazar, "Handbook of linear data-driven predictive control: Theory, implementation and design," *Annual Reviews in Control*, vol. 56, p. 100914, 2023.

[15] K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[16] L. Li, A. Bisoffi, C. De Persis, and N. Monshizadeh, "Controller synthesis from noisy-input noisy-output data," *arXiv preprint arXiv:2402.02588*, 2024.

[17] M. Alsalti, V. G. Lopez, and M. A. Müller, "Notes on data-driven output-feedback control of linear mimo systems," *arXiv preprint arXiv:2311.17484*, 2023.

[18] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, "From noisy data to feedback controllers: Nonconservative design via a matrix s-lemma," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 162–175, 2020.

[19] T. R. Steentjes, M. Lazar, and P. M. Van den Hof, "On data-driven control: Informativity of noisy input-output data with cross-covariance bounds," *IEEE Control Systems Letters*, vol. 6, pp. 2192–2197, 2021.

[20] H. J. van Waarde, J. Eising, M. K. Camlibel, and H. L. Trentelman, "A behavioral approach to data-driven control with noisy input–output data," *IEEE Transactions on Automatic Control*, vol. 69, no. 2, pp. 813–827, 2023.

[21] H. J. Van Waarde, M. K. Camlibel, J. Eising, and H. L. Trentelman, "Quadratic matrix inequalities with applications to data-based control," *SIAM Journal on Control and Optimization*, vol. 61, no. 4, pp. 2251–2281, 2023.

[22] A. Bisoffi, C. De Persis, and P. Tesi, "Trade-offs in learning controllers from noisy data," *Systems & Control Letters*, vol. 154, p. 104985, 2021.

[23] J. C. Willems, "Paradigms and puzzles in the theory of dynamical systems," *IEEE Transactions on automatic control*, vol. 36, no. 3, pp. 259–294, 1991.

[24] L. M. Spin, C. Verhoek, W. M. H. Heemels, N. van de Wouw, and R. Tóth, "Unified behavioral data-driven performance analysis a generalized plant approach," in *2024 European Control Conference (ECC)*. IEEE, 2024, pp. 894–899.

[25] P. van den Hof, "Approximate realization of noisy linear systems: the hankel and page matrix approach," 1983.

[26] I. Markovsky, "The most powerful unfalsified model for data with missing values," *Systems & Control Letters*, vol. 95, pp. 53–61, 2016.

[27] M. J, "Extract linearly independent subset of matrix columns (https://www.mathworks.com/matlabcentral/fileexchange/77437-extract-linearly-independent-subset-of-matrix-columns), matlab central file exchange., retrieved october 6, 2024."

[28] C. Scherer and S. Weiland, "Linear matrix inequalities in control," *Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands*, vol. 3, no. 2, 2000.

[29] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 10.1.9*, 2024. [Online]. Available: http://docs.mosek.com/latest/toolbox/index.html

[30] "Lecture notes: "system identification – data-driven modeling of dynamic systems", paul m.j. van den hof (version february 2020)."

[31] T. Söderström and P. Stoica, "Instrumental variable methods for system identification," *Circuits, Systems and Signal Processing*, vol. 21, no. 1, pp. 1–9, 2002.

[32] P. Young, "The instrumental variable method: a practical approach to identification and system parameter estimation," *IFAC Proceedings Volumes*, vol. 18, no. 5, pp. 1–15, 1985.

[33] M. Gilson and P. Van den Hof, "Instrumental variable methods for closed-loop system identification," *Automatica*, vol. 41, no. 2, pp. 241–249, 2005.

[34] G. Van Hattum, "A self-tuning approach toward maximal energy dissipation with application to active vibration isolation," 2021.

Appendix A

Complete Closed-Loop System for Noisy Data settings

Consider the closed-loop system with a direct relation between z(k) and v(k), such that

$$\begin{aligned} \chi_r(k+1) &= (\mathcal{A}_r + \mathcal{B}_r\mathcal{K})\chi_r(k) + \mathcal{V}_r v(k) \equiv A_K\chi_r(k) + B_K v(k), \\ z(k) &= (C_z + D_z\mathcal{K}_r)\chi_r(k) + D_K v(k) = C_K\chi_r(k) + D_K v(k). \end{aligned} \tag{A.1}$$

For such a system, considering $\mathcal{H}_\infty$ control, the condition (2.35) is modified as

$$\begin{bmatrix} X & 0 & A_K^\top X & C_K^\top \\ 0 & \gamma I & B_K^\top X & D_K^\top \\ XA_K & XB_K & X & 0 \\ C_K & D_K & 0 & \gamma I \end{bmatrix} \succ 0 \tag{A.2}$$

The aim is to develop the informativity LMI conditions to reduce the $\mathcal{H}_\infty$ norm γ of the closed-loop transfer $||T||_\infty$ where

$$T(z) = C_K(zI - A_K)^{-1}B_K + D_K. \tag{A.3}$$

The steps for obtaining a QMI for the application of the Matrix S-lemma (3.1) are examined.

Method 1 : Taking Schur complements of (A.2) while considering the $4^{th}, 1^{st}$ and 1^{st} diagonal elements in each step.
Considering the known definition

$$\begin{bmatrix} P & 0 & PA_K^\top & PC_K^\top \\ 0 & \gamma I & B_K^\top & D_K \\ A_K P & B_K & P & 0 \\ C_K P & D_K & 0 & \gamma I \end{bmatrix} \succ 0.$$

Take the Schur complement of the 4th diagonal element, with $\gamma I \succ 0$.

$$\implies \begin{bmatrix} P & 0 & PA_K^\top \\ 0 & \gamma I & B_K^\top \\ A_K P & B_K & P \end{bmatrix} - \begin{bmatrix} PC_K^\top \\ D_K \\ 0 \end{bmatrix} (\gamma I)^{-1} \begin{bmatrix} C_K P & D_K & 0 \end{bmatrix} \succ 0$$

$$\equiv \begin{bmatrix} P & 0 & PA_K^\top \\ 0 & \gamma I & B_K^\top \\ A_K P & B_K & P \end{bmatrix} - \begin{bmatrix} \gamma^{-1} PC_K^\top C_K P & \gamma^{-1} PC_K^\top D_K & 0 \\ \gamma^{-1} D_K C_K P & \gamma^{-1} D_K D_K^\top & 0 \\ 0 & 0 & 0 \end{bmatrix} \succ 0$$

$$\equiv \begin{bmatrix} P - \gamma^{-1} PC_K^\top C_K P & -\gamma^{-1} PC_K^\top D_K & PA_K^\top \\ -\gamma^{-1} D_K C_K P & \gamma I - \gamma^{-1} D_K D_K^\top & B_K^\top \\ A_K P & B_K & P \end{bmatrix} \succ 0.$$

Take the Schur complement of the 1^{st} diagonal element with the assumption

$$(S_1)^{-1} = P - \gamma^{-1} PC_K^\top C_K P \succ 0,$$

$$\implies \begin{bmatrix} \gamma I - \gamma^{-1} D_K D_K^\top & B_K^\top \\ B_K & P \end{bmatrix} - \begin{bmatrix} -\gamma^{-1} D_K C_K P \\ A_K P \end{bmatrix} S_1 \begin{bmatrix} -\gamma^{-1} PC_K^\top D_K & PA_K^\top \end{bmatrix} \succ 0,$$

$$\equiv \begin{bmatrix} \gamma I - \gamma^{-1} D_K D_K^\top & B_K^\top \\ B_K & P \end{bmatrix} - \begin{bmatrix} \gamma^{-2} D_K C_K P S_1 PC_K^\top D_K & -\gamma^{-1} D_K C_K P S_1 PA_K^\top \\ -\gamma^{-1} A_K P S_1 PC_K^\top D_K & A_K P S_1 PA_K^\top \end{bmatrix} \succ 0,$$

$$\equiv \begin{bmatrix} \gamma I - \gamma^{-1} D_K D_K^\top - \gamma^{-2} D_K C_K P S_1 PC_K^\top D_K & B_K^\top + \gamma^{-1} D_K C_K P S_1 PA_K^\top \\ B_K + \gamma^{-1} A_K P S_1 PC_K^\top D_K & P - A_K P S_1 PA_K^\top \end{bmatrix} \succ 0.$$

With the assumption

$$(S_2)^{-1} = \gamma I - \gamma^{-1} D_K D_K^\top - \gamma^{-2} D_K C_K P S_1 PC_K^\top D_K \succ 0,$$

$$\implies P - A_K P S_1 PA_K^\top - (B_K^\top + \gamma^{-1} D_K C_K P S_1 PA_K^\top) S_2 (B_K + \gamma^{-1} A_K P S_1 PC_K^\top D_K) \succ 0,$$

$$\equiv P - A_K P S_1 PA_K^\top - B_K^\top S_2 B_K - \gamma^{-1} B_K^\top S_2 A_K P S_1 PC_K^\top D_K - \gamma^{-1} D_K C_K P S_1 PA_K^\top S_2 B_K$$
$$- \gamma^{-2} D_K C_K P S_1 PA_K^\top S_2 A_K P S_1 PC_K^\top D_K \succ 0.$$

Here, the A_K term should be on the left-most side (and conversely, $A_K^\top$ to be on the rightmost side) for the creation of a QMI. This is not visible, specifically in the last three terms. Hence, this does not result in the desired structure.

Method 2 : Taking Schur complements of (A.2) while considering the 2×2 block diagonal element and then the 2^{nd} diagonal elements in each step.
Considering the known definition

$$\begin{bmatrix} P & 0 & PA_K^\top & PC_K^\top \\ 0 & \gamma I & B_K^\top & D_K \\ A_K P & B_K & P & 0 \\ C_K P & D_K & 0 & \gamma I \end{bmatrix} \succ 0.$$

Take the Schur complement of the upper 2×2 diagonal block element, resulting in the following matrix inequalities.

$$
\begin{aligned}
&\text{(a)} \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix} \succ 0, \\
&\text{(b)} \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix} - \begin{bmatrix} A_K P & B_K \\ C_K P & D_K \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix}^{-1} \begin{bmatrix} P A_K^\top & P C_K^\top \\ B_K^\top & D_K^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix} - \begin{bmatrix} A_K P & B_K \\ C_K P & D_K \end{bmatrix} \begin{bmatrix} P^{-1} & 0 \\ 0 & \gamma^{-1} I \end{bmatrix} \begin{bmatrix} P A_K^\top & P C_K^\top \\ B_K^\top & D_K^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix} - \begin{bmatrix} A_K & \gamma^{-1} B_K \\ C_K & \gamma^{-1} D_K \end{bmatrix} \begin{bmatrix} P A_K^\top & P C_K^\top \\ B_K^\top & D_K^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P & 0 \\ 0 & \gamma I \end{bmatrix} - \begin{bmatrix} A_K P A_K^\top + \gamma^{-1} B_K B_K^\top & A_K P C_K^\top + \gamma^{-1} B_K D_K^\top \\ C_K P A_K^\top + \gamma^{-1} D_K B_K^\top & C_K P C_K^\top + \gamma^{-1} D_K D_K^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P - A_K P A_K^\top - \gamma^{-1} B_K B_K^\top & -A_K P C_K^\top - \gamma^{-1} B_K D_K^\top \\ -C_K P A_K^\top - \gamma^{-1} D_K B_K^\top & \gamma I - C_K P C_K^\top - \gamma^{-1} D_K D_K^\top \end{bmatrix} \succ 0.
\end{aligned}
$$

Having simplified the inequalities, assume

$$S_1 = \gamma I - C_K P C_K^\top - \gamma^{-1} D_K D_K^\top \succ 0,$$

and take the Schur complement again as

$$
\begin{aligned}
&\textbf{(c)} \; (P - A_K P A_K^\top - \gamma^{-1} B_K B_K^\top) - (-A_K P C_K^\top - \gamma^{-1} B_K D_K^\top) S_1^{-1} (-C_K P A_K^\top - \gamma^{-1} D_K B_K^\top) \succ 0, \\
&\equiv \begin{bmatrix} P - A_K P A_K^\top - \gamma^{-1} B_K B_K^\top - A_K P C_K^\top S_1^{-1} C_K P A_K^\top - \gamma^{-1} A_K P C_K^\top S_1^{-1} D_K B_K^\top \cdots \\ \cdots - \gamma^{-1} B_K D_K^\top S_1^{-1} C_K P A_K^\top - \gamma^{-2} B_K D_K^\top S_1^{-1} D_K B_K^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P - A_K (P + P C_K^\top S_1^{-1} C_K P) A_K^\top - (\gamma^{-1} B_K B_K^\top + \gamma^{-2} B_K D_K^\top S_1^{-1} D_K B_K^\top) \cdots \\ \cdots - \gamma^{-1} A_K P C_K^\top S_1^{-1} D_K B_K^\top - \gamma^{-1} B_K D_K^\top S_1^{-1} C_K P A_K^\top \end{bmatrix}.
\end{aligned}
$$

Consider

$$
\begin{aligned}
&S_2 = P - (\gamma^{-1} B_K B_K^\top + \gamma^{-2} B_K D_K^\top S_1^{-1} D_K B_K^\top), \\
&S_3 = P + P C_K^\top S_1^{-1} C_K P, \\
&A_K = \mathcal{A}_r + \mathcal{B}_r \mathcal{K}_r = \Lambda_{er} + J_{1r} + \Lambda_{2r} \mathcal{K}_r + J_{2r} \mathcal{K}_r, \\
&C_K = C_z + D_z \mathcal{K}_r, \quad B_K = \mathcal{V}_r.
\end{aligned}
$$

This results in

$$
\begin{aligned}
&\equiv P - A_K S_3 A_K^\top - S_2 - \gamma^{-1} A_K P C_K^\top S_1^{-1} D_K C_K^\top - \gamma^{-1} B_K D_K^\top S_1^{-1} C_K P A_K^\top \succ 0, \\
&\equiv \begin{bmatrix} P - (\Lambda_{er} + J_{1r} + \Lambda_{2r} \mathcal{K}_r + J_{2r} \mathcal{K}_r) S_3 (\Lambda_{er} + J_{1r} + \Lambda_{2r} \mathcal{K}_r + J_{2r} \mathcal{K}_r)^\top \cdots \\ \cdots - \gamma^{-1} (\Lambda_{er} P + J_{1r} P + \Lambda_{2r} L + J_{2r} L) C_K^\top S_1^{-1} D_K B_K^\top \cdots \\ \cdots - \gamma^{-1} B_K D_K^\top S_1^{-1} C_K (\Lambda_{er} P + J_{1r} P + \Lambda_{2r} L + J_{2r} L)^\top \end{bmatrix} \succ 0, \\
&\equiv \begin{bmatrix} P - \Lambda_{er} S_3 \Lambda_{er}^\top - J_{1r} S_3 J_{1r} + \Lambda_{2r} \mathcal{K}_r + J_{2r} \mathcal{K}_r) S_3 (\Lambda_{er} + J_{1r} + \Lambda_{2r} \mathcal{K}_r + J_{2r} \mathcal{K}_r)^\top \cdots \\ \cdots - \gamma^{-1} (\Lambda_{er} P + J_{1r} P + \Lambda_{2r} L + J_{2r} L) C_K^\top S_1^{-1} D_K B_K^\top \cdots \\ \cdots - \gamma^{-1} B_K D_K^\top S_1^{-1} C_K (\Lambda_{er} P + J_{1r} P + \Lambda_{2r} L + J_{2r} L)^\top \end{bmatrix} \succ 0.
\end{aligned}
$$

It can be seen that a term $\mathcal{K}_r P$ is required, so that $E = \mathcal{K}_r P$, but that is not the case here. Hence, this method cannot be used. Similar research blocks are observed for other methods, and also for $\mathcal{H}_2$ control, leaving this as a future work.