Automatica 183 (2026) 112545

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Controller synthesis from noisy-input noisy-output data[☆]

Lidong Li [a], Andrea Bisoffi [b,*], Claudio De Persis [a], Nima Monshizadeh [a]

[a] Engineering and Technology Institute, University of Groningen, 9747AG, The Netherlands
[b] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, 20133, Italy

ARTICLE INFO

Article history:
Received 18 February 2024
Received in revised form 17 March 2025
Accepted 17 July 2025

Keywords:
Linear time-invariant system
Data-based control
Analysis of systems with uncertainties
Robust control
Output feedback control
Dynamic controller

ABSTRACT

We consider the problem of synthesizing a dynamic output-feedback controller for a linear system, using solely input–output data corrupted by measurement noise. To handle input–output data, an auxiliary representation of the original system is utilized. By exploiting the structure of the auxiliary system, we design a controller that robustly stabilizes all possible systems consistent with data. Notably, we also provide a novel solution to extend the results to generic multi-input multi-output systems. The findings are illustrated by numerical examples.

© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Facilitated by recent technological advancements in sensor and measurement devices, along with the widespread availability of data, data-driven control has become an increasingly popular research avenue in recent years. Typically, data has been used to identify a mathematical model describing the dynamics of the system. Such a model can, in turn, be used for control and monitoring purposes. An alternative approach, often referred to as direct data-driven control, is to bypass the model and directly use data to synthesize the control algorithm (Campi, Lecchini, & Savaresi, 2002; Coulson, Lygeros, & Dörfler, 2019; De Persis & Tesi, 2020; Tanaskovic, Fagiano, Novara, & Morari, 2017).

One of the most classical problems in control theory is stabilizing a linear time-invariant system. In the realm of data-driven control, this raises the question of how to design stabilizing controllers directly from the data obtained from the system. In practice, such data come from multiple actuators and sensors and are not completely accurate. This leads to the problem of stabilizing a multi-input multi-output (MIMO) linear system using *noisy-input noisy-output* data. As detailed below, there have been numerous valuable efforts to tackle this challenging problem by approaching it from different angles. Nevertheless, to the best of our knowledge, a complete answer to this problem is still absent from the literature. The main motivation of this work is to fill this gap.

Related works

The central problem posed above has two key challenges: (i) the use of *input–output* data rather than state measurements and (ii) the presence of noise on both input and output channels. To cope with the first challenge, De Persis and Tesi (2020) proposes to work with an *auxiliary representation* of the system which is built using time-shifted inputs and outputs of the system (Goodwin & Sin, 2009). This auxiliary representation enables the designer to take advantage of data-driven control techniques that are applicable to input-state data; see, e.g., Al Makdah and Pasqualetti (2024), Berberich, Scherer, and Allgöwer (2023), Dai, De Persis, and Monshizadeh (2023), D'Amico, Bisoffi, and Farina (2023), de Jong, Breschi, Schoukens, and Formentin (2023), Jo and Shim (2022), Pan, Ou, and Faulwasser (2023), Steentjes, Lazar, and Van den Hof (2022). However, most of the aforementioned works efficiently address only single-input single-output or multiple-input single-output cases. This limitation primarily arises from the fact that the auxiliary representation of a minimal MIMO linear system is not necessarily reachable (Antsaklis & Michel, 2005, p. 566). To tackle the challenge posed by the absence of reachability, Alsalti, Lopez, and Müller (2025) and Sadamoto (2023) aim at extracting the reachable modes from the auxiliary representation. However, these developments are not readily applicable to the practically notable scenarios where noise is present on *both* input and output channels, bringing us to the second challenge of the problem.

[☆] This paper was not presented at any IFAC meeting. This work was supported by Project Digital Twin of Research Programme TTW Perspective which is (partly) financed by the Dutch Research Council (NWO) under Grant P18-03. L. Li is supported by the Chinese Scholarship Council. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Simone Formentin under the direction of Editor Alessandro Chiuso.

[*] Corresponding author.

E-mail addresses: l.li@rug.nl (L. Li), andrea.bisoffi@polimi.it (A. Bisoffi), c.de.persis@rug.nl (C. De Persis), n.monshizadeh@rug.nl (N. Monshizadeh).

https://doi.org/10.1016/j.automatica.2025.112545

0005-1098/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Besides the concern of unreachability of the auxiliary representation, the presence of noise in input–output data further complicates the task of control design. For the case of process noise, which acts as an additive disturbance to the system equations, we refer to Koch, Berberich, and Allgöwer (2022), Steentjes et al. (2022), van Waarde, Eising, Camlibel, and Trentelman (2024). Specifically, Steentjes et al. (2022) considers cross-covariance bounds on the process noise which are potentially less conservative than the norm bounds used in the literature, (van Waarde et al., 2024) uses a behavioral approach, and (Koch et al., 2022) verifies dissipativity properties. Unlike the aforementioned literature, the output and input measurement noise considered in this paper captures sensor and actuator inaccuracies. This type of noise was studied in classical system identification under the setting of "errors-in-variables" (Söderström, 2018) but is less explored in direct data-driven control. Measurement noise substantially increases the complexity since the system dynamics further intertwine the useful signal and detrimental noise in the collected data. The case of noisy state measurements is treated in De Persis and Tesi (2020), Miller, Dai, and Sznaier (2023a). In fact, by deriving a matrix elimination result and suitably leveraging Petersen's lemma (Petersen, 1987), one can establish a necessary and sufficient condition for robust stabilization of systems consistent with noisy-input noisy-*state* data (Bisoffi, Li, De Persis, & Monshizadeh, 2024). A treatment of the noisy-input noisy-*output* case has only been recently reported in Miller, Dai, and Sznaier (2023b) for single-input single-output systems via superstability, which is a stronger notion than asymptotic stability.

Contribution and proposed approach

As reviewed above, the main contribution of this work is to design a stabilizing controller by noisy-input noisy-output data acquired from a generic MIMO system.

Given a linear system of minimal order n, we distinguish between two cases. The first case is $n = p\ell$ where p is the number of outputs and ℓ is the observability index of the system. In this case, the auxiliary system can be reachable and, after separating the unknown blocks of the auxiliary representation from the known ones, we make use of Petersen's lemma in the context of data-driven control (Bisoffi, De Persis, & Tesi, 2022) to design a data-driven controller through a convenient linear matrix inequality. This controller stabilizes the auxiliary representation and, by a thorough investigation of the relationship between the input–output evolution of the original and the auxiliary system, we conclude that the controller stabilizes the original system as well; see Theorem 1.

The second case is $n \neq p\ell$, where the auxiliary system becomes unreachable and it turns out that the data-based conditions that are exploited in the first case do not hold. To address this challenge, we introduce a novel approach where we form an augmented system by a parallel interconnection of the original system with another linear system. By imposing suitable conditions on the added dynamics, we bring the setup back to the first case, thus extending the results to generic multi-input multi-output linear systems. In particular, we show that the data-driven controller designed for the augmented system together with the added linear system serve as a stabilizing dynamic output-feedback controller for the original system; see Theorem 2.

Outline

Section 2 introduces the auxiliary system and discusses its relation with the input–output evolution of the original system. We formulate the problem of interest in Section 3. Section 4 presents our main results for both cases of $n = p\ell$ and $n \neq p\ell$. The results are verified by numerical examples[1] in Section 5. Due to space limitations, the omitted proofs of auxiliary lemmas can be found in Li, Bisoffi, De Persis, and Monshizadeh (2024).

2. Preliminaries

2.1. Notation

For column vectors $v_1, \dots, v_n$, $(v_1, \dots, v_n)$ denotes the stacked vector $[v_1^\top \ \dots \ v_n^\top]^\top$. The identity matrix of size n and the zero matrix of size $m \times n$ are I_n and $0_{m\times n}$: the indices are dropped when no confusion arises. The largest and smallest eigenvalues of a symmetric matrix M are $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$. The largest and smallest singular values of a matrix M are $\sigma_{\max}(M)$ and $\sigma_{\min}(M)$. The set of eigenvalues of a matrix M is $\operatorname{eig}(M)$. The 2-norm of a vector v is $|v|$. The induced 2-norm of a matrix M is $\|M\|$ and is given by $\sigma_{\max}(M)$. For a symmetric matrix $\left[\begin{smallmatrix} M & N \\ N^\top & O \end{smallmatrix}\right]$, we may use the shorthand writing $\left[\begin{smallmatrix} M & N \\ \star & O \end{smallmatrix}\right]$. For a positive semidefinite matrix M, $M^{1/2}$ is the unique positive semidefinite square root of M. For a discrete-time signal $h\colon \mathbb{Z} \to \mathbb{R}^n$, $\{h(k)\}_{k=k_0}^{k_1}$ is the sequence of values $h(k_0), h(k_0+1), \dots, h(k_1)$, where $k_0 \le k_1$ and k_1 possibly ∞. For the signal $\{h(k)\}_{k=0}^{\infty}$, its $\mathcal{L}_\infty$ norm is $\|h\|_{\mathcal{L}_\infty} := \sup_{k\ge 0} |h(k)|$. Given an n-dimensional state–space model $x^+ = Ax + Bu$, its reachable subspace is $R(A, B) := \operatorname{range}[A^{n-1}B \ \cdots \ AB \ B]$.

2.2. The auxiliary system and its properties

Consider a generic discrete-time LTI system

$$x^+ = Ax + Bu \tag{1a}$$

$$y = Cx \tag{1b}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$. If the pair (A, C) is observable, there exist l such that

$$\operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{l-1} \end{bmatrix} = n;$$

the minimum l for which this rank condition holds is denoted by ℓ, which is called the observability index (of the pair (A, C)) (Kailath, 1980, p. 356–357). Then, if the pair (A, C) is observable, we can define the matrices

$$\mathcal{O}_\ell := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix} \in \mathbb{R}^{p\ell\times n} \tag{2a}$$

$$\mathcal{T}_\ell := \begin{bmatrix} 0_{p\times m} & 0 & \dots & 0 & 0 \\ CB & 0 & \dots & 0 & 0 \\ CAB & CB & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{\ell-2}B & CA^{\ell-3}B & \dots & CB & 0_{p\times m} \end{bmatrix} \in \mathbb{R}^{p\ell\times m\ell} \tag{2b}$$

$$\mathcal{R}_\ell := \begin{bmatrix} A^{\ell-1}B & \dots & AB & B \end{bmatrix} \in \mathbb{R}^{n\times m\ell}. \tag{2c}$$

Since $\operatorname{rank} \mathcal{O}_\ell = n$ by construction, $\mathcal{O}_\ell$ possesses a left inverse $\mathcal{O}_\ell^{\mathrm{L}}$ that satisfies $\mathcal{O}_\ell^{\mathrm{L}} \mathcal{O}_\ell = I_n$. Throughout the paper, we are also interested in the *auxiliary system*

$$\xi^+ = \mathbf{A}_\ell \xi + \mathbf{B}_\ell v := (\mathbf{F}_\ell + \mathbf{L}_\ell Z_\ell)\xi + \mathbf{B}_\ell v \tag{3}$$

where we define, for ℓ as above,

$$\begin{bmatrix} \mathbf{F}_\ell & \mathbf{L}_\ell & \mathbf{B}_\ell \end{bmatrix} := \tag{4a}$$

$$\left[\left[\begin{array}{ccccc|c} 0 & I_p & 0 & \cdots & 0 & \\ 0 & 0 & I_p & \cdots & 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0_{p\ell\times m\ell} \\ 0 & 0 & 0 & \cdots & I_p & \\ 0 & 0 & 0 & \cdots & 0 & \\ \hline & & & & & \begin{array}{ccccc} 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_m \\ 0 & 0 & 0 & \dots & 0 \end{array} \\ & & 0_{m\ell\times p\ell} & & & \end{array}\right] \quad \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ I_p \\ \hline 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}\right] \quad \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{array}\right] \right]$$

and, for (A, B, C) in (1) and $\mathcal{O}_\ell$, $\mathcal{T}_\ell$, $\mathcal{R}_\ell$ in (2),

$$Z_\ell := \begin{bmatrix} Z_{1\ell} & Z_{2\ell} \end{bmatrix} := \begin{bmatrix} CA^\ell \mathcal{O}_\ell^{\mathrm{L}} & C\mathcal{R}_\ell - CA^\ell \mathcal{O}_\ell^{\mathrm{L}} \mathcal{T}_\ell \end{bmatrix}. \tag{4b}$$

---

[1] The code of the numerical examples is available at https://github.com/BG5CPU/DataDrivenNoisyIO

Note that from (3) and (4a),

$$\mathbf{A}_\ell = \left[\begin{array}{c|c} \begin{matrix} 0 & I_p & 0 & \dots & 0 \\ 0 & 0 & I_p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_p \end{matrix} & \begin{matrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \\ Z_{1\ell} & Z_{2\ell} \\ \hline 0 & \begin{matrix} 0 & I_m & 0 & \dots & 0 \\ 0 & 0 & I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_m \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \end{array}\right]$$

$$= \left[\begin{array}{c|c} \begin{matrix} 0 & I_p & 0 & \dots & 0 \\ 0 & 0 & I_p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_p \end{matrix} & \begin{matrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \\ CA^\ell \mathcal{O}_\ell^{\mathrm{L}} & C\mathcal{R}_\ell - CA^\ell \mathcal{O}_\ell^{\mathrm{L}} \mathcal{T}_\ell \\ \hline 0 & \begin{matrix} 0 & I_m & 0 & \dots & 0 \\ 0 & 0 & I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_m \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \end{array}\right]. \tag{5}$$

In Appendix A, we report some straightforward results for the linear system in (1) and the auxiliary linear system in (3), which are invoked in the proofs of the subsequent Lemmas 1, 2, 3. Define the matrix H_ℓ as

$$H_\ell := \left[\begin{array}{c|c} \mathcal{O}_\ell & \mathcal{T}_\ell \\ \hline 0 & I_{m\ell} \end{array}\right] \in \mathbb{R}^{(p\ell+m\ell)\times(n+m\ell)}. \tag{6}$$

We have the next result, which was claimed within Sadamoto (2023, Proof of Lemma 4), but not proven therein; the reachable subspace is defined in Section 2.1.

Lemma 1. *For (A, C) observable and (A, B) reachable, let $R(\mathbf{A}_\ell, \mathbf{B}_\ell)$ be the reachable subspace of the pair $(\mathbf{A}_\ell, \mathbf{B}_\ell)$. Then, $\operatorname{range} H_\ell = R(\mathbf{A}_\ell, \mathbf{B}_\ell)$.*

Proof. See Li et al. (2024, Proof of Lemma 1). □

The next result is a relevant consequence of Lemma 1.

Lemma 2. *For (A, C) observable and (A, B) reachable, the pair $(\mathbf{A}_\ell, \mathbf{B}_\ell)$ is reachable if and only if $p\ell = n$.*

Proof. See Li et al. (2024, Proof of Lemma 2). □

Finally, we show next how the input–output evolution of solutions to (1) can be captured by solutions to (3).

Lemma 3. *Given the matrices (A, B, C) of (1), let (A, C) be observable and $\mathbf{A}_\ell$, $\mathbf{B}_\ell$, Z_ℓ be as in (3)–(4). For each $\hat{x}$ and sequence $\{u(k)\}_{k=0}^{\infty}$, there exists $\hat{\xi}$ such that:*

- *the solution $x(\cdot)$ to (1) with initial condition $x(0) = \hat{x}$ and input $\{u(k)\}_{k=0}^{\infty}$,*
- *the corresponding output response $y(\cdot) = Cx(\cdot)$, and*
- *the solution $\xi(\cdot)$ to (3) with initial condition $\xi(\ell) = \hat{\xi}$ and input $\{v(k)\}_{k=\ell}^{\infty} = \{u(k)\}_{k=\ell}^{\infty}$*

satisfy

$$\left[\begin{array}{c} y(k-\ell) \\ \vdots \\ y(k-1) \\ \hline u(k-\ell) \\ \vdots \\ u(k-1) \end{array}\right] = \xi(k) \quad \forall k \geq \ell \tag{7a}$$

$$y(k) = Z_\ell \xi(k) \quad \forall k \geq \ell. \tag{7b}$$

Proof. See Li et al. (2024, Proof of Lemma 3). □

We note that Lemma 3 compares the solution to (1) for a generic initial condition $\hat{x}$ with the solution to (3) with $\xi(\ell) = \hat{\xi}$ for a suitable $\hat{\xi}$ from $k = \ell$ onward.

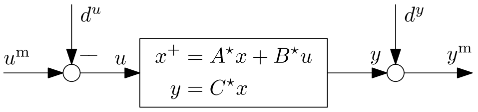

Fig. 1. Scheme of the data collection experiment.

3. Problem formulation

Consider the discrete-time linear-time-invariant system

$$x^+ = A^\star x + B^\star u \tag{8a}$$

$$y = C^\star x \tag{8b}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^m$ and output $y \in \mathbb{R}^p$. The matrices $A^\star$, $B^\star$, $C^\star$ are unknown to us and, instead of their knowledge, we rely on collecting noisy-input noisy-output measurements to design a controller for (8), as we explain below. Our prior knowledge on (8) is summarized in the next assumption.

Assumption 1. The pair $(A^\star, C^\star)$ is observable, and the observability index $\ell^\star$ of $(A^\star, C^\star)$ is known.

Knowing the observability index $\ell^\star$ allows us to construct some of the quantities used in the sequel. Based on Assumption 1, we can define the unknown matrices $\mathcal{O}_{\ell^\star}$, $\mathcal{T}_{\ell^\star}$, $\mathcal{R}_{\ell^\star}$ as in (2a), (2b), (2c), where, since $(A^\star, C^\star)$ is observable, $\operatorname{rank} \mathcal{O}_{\ell^\star} = n$ and $\mathcal{O}_{\ell^\star}$ possesses a left inverse $\mathcal{O}_{\ell^\star}^{\mathrm{L}}$ that satisfies $\mathcal{O}_{\ell^\star}^{\mathrm{L}} \mathcal{O}_{\ell^\star} = I_n$.

Input–output data are collected by performing an experiment on (8). Consider

$$u^{\mathrm{m}} := u + d^u \text{ and } y^{\mathrm{m}} := y + d^y \tag{9}$$

where the measured input u^{m} differs from the actual input u to (8) by unknown noise d^u, and the measured output y^{m} differs from the actual output y of (8) by unknown noise d^y. The data collection experiment is depicted in Fig. 1 and is as follows: for $k = 0, \dots, T$, apply the signal $u^{\mathrm{m}}(k)$; along with noise $d^u(k)$, this results in (unknown) input $u(k) = u^{\mathrm{m}}(k) - d^u(k)$, (unknown) state $x(k)$ and (unknown) output $y(k)$ for some initial condition $x(0)$; measure the signal $y^{\mathrm{m}}(k) = y(k) + d^y(k)$. The available data, on which our control design is based, are then $\{u^{\mathrm{m}}(k), y^{\mathrm{m}}(k)\}_{k=0}^{T}$ and are gathered from

$$x^+ = A^\star x + B^\star(u^{\mathrm{m}} - d^u) \tag{10a}$$

$$y^{\mathrm{m}} = C^\star x + d^y. \tag{10b}$$

For the sequel we define

$$\Psi_1 := \left[\begin{array}{cccc} y^{\mathrm{m}}(1) & y^{\mathrm{m}}(2) & \dots & y^{\mathrm{m}}(T-\ell^\star+1) \\ \vdots & \vdots & & \vdots \\ y^{\mathrm{m}}(\ell^\star) & y^{\mathrm{m}}(\ell^\star+1) & \dots & y^{\mathrm{m}}(T) \\ \hline u^{\mathrm{m}}(1) & u^{\mathrm{m}}(2) & \dots & u^{\mathrm{m}}(T-\ell^\star+1) \\ \vdots & \vdots & & \vdots \\ u^{\mathrm{m}}(\ell^\star) & u^{\mathrm{m}}(\ell^\star+1) & \dots & u^{\mathrm{m}}(T) \end{array}\right] \tag{11a}$$

$$\Psi_0 := \left[\begin{array}{cccc} y^{\mathrm{m}}(0) & y^{\mathrm{m}}(1) & \dots & y^{\mathrm{m}}(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ y^{\mathrm{m}}(\ell^\star-1) & y^{\mathrm{m}}(\ell^\star) & \dots & y^{\mathrm{m}}(T-1) \\ \hline u^{\mathrm{m}}(0) & u^{\mathrm{m}}(1) & \dots & u^{\mathrm{m}}(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ u^{\mathrm{m}}(\ell^\star-1) & u^{\mathrm{m}}(\ell^\star) & \dots & u^{\mathrm{m}}(T-1) \end{array}\right] \tag{11b}$$

$$U_1 := \left[\, u^{\mathrm{m}}(\ell^\star) \; u^{\mathrm{m}}(\ell^\star+1) \; \dots \; u^{\mathrm{m}}(T) \,\right] \tag{11c}$$

based on the available data $\{u^{\mathrm{m}}(k), y^{\mathrm{m}}(k)\}_{k=0}^{T}$ and also the unknown

$$\Delta_{10} := \left[\begin{array}{cccc} d^y(\ell^\star) & d^y(\ell^\star+1) & \dots & d^y(T) \\ \hline d^y(0) & d^y(1) & \dots & d^y(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^y(\ell^\star-1) & d^y(\ell^\star) & \dots & d^y(T-1) \\ \hline d^u(0) & d^u(1) & \dots & d^u(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^u(\ell^\star-1) & d^u(\ell^\star) & \dots & d^u(T-1) \end{array}\right]. \tag{11d}$$

For $\mathbf{F}_{\ell^\star}$, $\mathbf{L}_{\ell^\star}$ and $\mathbf{B}_{\ell^\star}$ as in (4a), define

$$\begin{aligned} Z_{\ell^\star} &:= \begin{bmatrix} Z_{1\ell^\star} & Z_{2\ell^\star} \end{bmatrix} \\ &:= \begin{bmatrix} C^\star A^{\star\ell^\star} \mathcal{O}^{\mathrm{L}}_{\ell^\star} & C^\star \mathcal{R}_{\ell^\star} - C^\star A^{\star\ell^\star} \mathcal{O}^{\mathrm{L}}_{\ell^\star} \mathcal{T}_{\ell^\star} \end{bmatrix} \end{aligned} \tag{12}$$

$$\mathbf{A}_{\ell^\star} := \mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} \begin{bmatrix} Z_{1\ell^\star} & Z_{2\ell^\star} \end{bmatrix} \tag{13}$$

$$= \left[\begin{array}{c|c} \begin{matrix} 0 & I_p & 0 & \dots & 0 \\ 0 & 0 & I_p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_p \end{matrix} & \begin{matrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \\ Z_{1\ell^\star} & Z_{2\ell^\star} \\ \hline 0 & \begin{matrix} 0 & I_m & 0 & \dots & 0 \\ 0 & 0 & I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_m \\ 0 & 0 & 0 & \dots & 0 \end{matrix} \end{array}\right]$$

$$\mathbf{B}^{\mathrm{d}}_{\ell^\star} := \mathbf{L}_{\ell^\star} \begin{bmatrix} I_p & -Z_{1\ell^\star} & -Z_{2\ell^\star} \end{bmatrix} \tag{14}$$

$$= \left[\begin{array}{ccc} 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ I_p & -Z_{1\ell^\star} & -Z_{2\ell^\star} \\ \hline 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{array}\right]$$

where $Z_{\ell^\star}$ is as in (4b) and $\mathbf{A}_{\ell^\star}$ is as in (5). We summarize the relations satisfied by data in the next result.

Lemma 4. *Under Assumption 1, data satisfy for* $k = \ell^\star, \dots, T$,

$$\begin{aligned} y^{\mathrm{m}}(k) = {} & Z_{1\ell^\star} \begin{bmatrix} y^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ y^{\mathrm{m}}(k-1) \end{bmatrix} + Z_{2\ell^\star} \begin{bmatrix} u^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ u^{\mathrm{m}}(k-1) \end{bmatrix} \\ & - Z_{1\ell^\star} \begin{bmatrix} d^y(k-\ell^\star) \\ \vdots \\ d^y(k-1) \end{bmatrix} - Z_{2\ell^\star} \begin{bmatrix} d^u(k-\ell^\star) \\ \vdots \\ d^u(k-1) \end{bmatrix} + d^y(k) \end{aligned} \tag{15}$$

$$\begin{aligned} \left[\begin{array}{c} y^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ y^{\mathrm{m}}(k) \\ \hline u^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ u^{\mathrm{m}}(k) \end{array}\right] = {} & (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star}[Z_{1\ell^\star} \; Z_{2\ell^\star}]) \left[\begin{array}{c} y^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ y^{\mathrm{m}}(k-1) \\ \hline u^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ u^{\mathrm{m}}(k-1) \end{array}\right] \\ & + \mathbf{B}_{\ell^\star} u^{\mathrm{m}}(k) + \mathbf{L}_{\ell^\star}[I_p \;\; -Z_{1\ell^\star} \;\; -Z_{2\ell^\star}] \left[\begin{array}{c} d^y(k) \\ \hline d^y(k-\ell^\star) \\ \vdots \\ d^y(k-1) \\ \hline d^u(k-\ell^\star) \\ \vdots \\ d^u(k-1) \end{array}\right] \end{aligned} \tag{16}$$

$$\begin{aligned} \Psi_1 &= (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z_{\ell^\star}) \Psi_0 + \mathbf{B}_{\ell^\star} U_1 + \mathbf{L}_{\ell^\star}[I_p \;\; - Z_{\ell^\star}] \Delta_{10} \\ &= \mathbf{A}_{\ell^\star} \Psi_0 + \mathbf{B}_{\ell^\star} U_1 + \mathbf{B}^{\mathrm{d}}_{\ell^\star} \Delta_{10}. \end{aligned} \tag{17}$$

Proof. See Appendix B. □

In (17), $\mathbf{B}_{\ell^\star}$ as in (4a) is known whereas $\mathbf{A}_{\ell^\star}$ in (13) and $\mathbf{B}^{\mathrm{d}}_{\ell^\star}$ in (14) contain the unknown $Z_{\ell^\star}$. Without any prior knowledge on Δ_{10}, (17) would not provide sufficient information to design a stabilizing controller for (8). Thus, we introduce prior knowledge in the form of an energy bound on Δ_{10}. For some $\Theta = \Theta^\top \succeq 0$, define

$$\mathcal{D} := \{\Delta \in \mathbb{R}^{(p+p\ell^\star+m\ell^\star)\times(T-\ell^\star+1)} : \Delta\Delta^\top \preceq \Theta\} \tag{18}$$

and, for the sequel, partition Θ as

$$\Theta =: \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{12}^\top & \Theta_{22} \end{bmatrix} \tag{19}$$

with $\Theta_{11} = \Theta_{11}^\top \in \mathbb{R}^{p\times p}$ and $\Theta_{22} = \Theta_{22}^\top \in \mathbb{R}^{(p\ell^\star+m\ell^\star)\times(p\ell^\star+m\ell^\star)}$. We assume then that the noise sequence acting during data collection belongs to $\mathcal{D}$, i.e., $\Delta_{10} \in \mathcal{D}$. The interpretation of (18) as an energy bound is justified in the next remark.

Remark 1. The permutation matrix $\Pi := \begin{bmatrix} 0 & I_{\ell^\star p} & 0 \\ I_p & 0 & 0 \\ 0 & 0 & I_{\ell^\star m} \end{bmatrix}$ ensures that

$$\Pi\Delta_{10} = \left[\begin{array}{cccc} d^y(0) & d^y(1) & \dots & d^y(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^y(\ell^\star) & d^y(\ell^\star+1) & \dots & d^y(T) \\ \hline d^u(0) & d^u(1) & \dots & d^u(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^u(\ell^\star-1) & d^u(\ell^\star) & \dots & d^u(T-1) \end{array}\right].$$

Then, $\Delta_{10} \in \mathcal{D}$ if and only if $\Pi\Delta_{10}\Delta_{10}^\top\Pi^\top \preceq \Pi\Theta\Pi^\top$, i.e.,

$$\sum_{j=0}^{T-\ell^\star} \left[\begin{array}{c} d^y(j) \\ \vdots \\ d^y(j+\ell^\star) \\ \hline d^u(j) \\ \vdots \\ d^u(j+\ell^\star-1) \end{array}\right] \left[\begin{array}{c} d^y(j) \\ \vdots \\ d^y(j+\ell^\star) \\ \hline d^u(j) \\ \vdots \\ d^u(j+\ell^\star-1) \end{array}\right]^\top \preceq \Pi\Theta\Pi^\top.$$

This condition highlights that Θ, modulo permutations, expresses an energy bound on the sequence of vectors $\big(d^y(0), \dots, d^y(\ell^\star), d^u(0), \dots, d^u(\ell^\star - 1)\big)$, …, $\big(d^y(T - \ell^\star), \dots, d^y(T), d^u(T - \ell^\star), \dots, d^u(T - 1)\big)$. ■

A typical way to obtain the energy bound given by Θ is presented in the next remark.

Remark 2. Suppose we know that for some $\bar{d}^y \geq 0$ and $\bar{d}^u \geq 0$,

$$\|d^y\|_{\mathcal{L}_\infty} \leq \bar{d}^y \text{ and } \|d^u\|_{\mathcal{L}_\infty} \leq \bar{d}^u.$$

Then, for each $k = \ell^\star - 1, \dots, T - 1$,

$$\begin{aligned} |\delta(k)|^2 &:= \left\| \begin{bmatrix} d^y(k+1) \\ d^y(k-\ell^\star+1) \\ \vdots \\ d^y(k) \\ d^u(k-\ell^\star+1) \\ \vdots \\ d^u(k) \end{bmatrix} \right\|^2 \\ &= \sum_{j=k-\ell^\star+1}^{k+1} |d^y(j)|^2 + \sum_{j=k-\ell^\star+1}^{k} |d^u(j)|^2 \leq (\ell^\star + 1)(\bar{d}^y)^2 + \ell^\star(\bar{d}^u)^2 \end{aligned}$$

and, from (11d),

$$\begin{aligned} \Delta_{10}\Delta_{10}^\top &= \sum_{k=\ell^\star-1}^{T-1} \delta(k)\delta(k)^\top \preceq \sum_{k=\ell^\star-1}^{T-1} |\delta(k)|^2 I \\ &\preceq (T - \ell^\star + 1)\big((\ell^\star + 1)(\bar{d}^y)^2 + \ell^\star(\bar{d}^u)^2\big)I. \end{aligned}$$

In this way, we can take Θ as $(T - \ell^\star + 1)\big((\ell^\star + 1)(\bar{d}^y)^2 + \ell^\star(\bar{d}^u)^2\big)I$. ■

Under this setting, where the actual parameters $Z_{\ell^\star}$ arising from (8) are unknown, we introduce the set of parameters Z consistent with data in (17) and with the noise bound in (18) as

$$\begin{aligned} \mathcal{C} := \Big\{ Z \in \mathbb{R}^{p\times(p\ell^\star+m\ell^\star)} \colon & \Psi_1 = (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z)\Psi_0 \\ & + \mathbf{B}_{\ell^\star} U_1 + \mathbf{L}_{\ell^\star}[I_p \;\; - Z]\Delta, \quad \Delta \in \mathcal{D} \Big\}. \end{aligned} \tag{20}$$

Particularly, (17) and $\Delta_{10} \in \mathcal{D}$ imply $Z_{\ell^\star} \in \mathcal{C}$. It can be shown that for any $Z \in \mathcal{C}$, there exist state/input/output matrices (A, B, C) consistent with data in (11) and noise bound in (18) that, substituted into (4b), return $Z_\ell = Z$. Hence, neglecting the underlying structure of (4b) in (20) does not introduce any conservatism. In the sequel, we will need to work with the set (20), instead of with the unknown $Z_{\ell^\star}$. Moreover, we make the next assumption on data.

Assumption 2. $\Psi_0\Psi_0^\top \succ \Theta_{22}$.

The interpretation and implications of this assumption are discussed in Section 4.4.

As in the data collection experiment, we rely on current and past outputs and inputs to control (8) and render the origin asymptotically stable. In other words, we use the feedback law

$$u(k)=\mathbf{K}\left[\begin{array}{c} y(k-\ell^\star)\\ \vdots\\ y(k-1)\\ \hline u(k-\ell^\star)\\ \vdots\\ u(k-1)\end{array}\right]\quad \forall k\geq \ell^\star \tag{21}$$

for some matrix $\mathbf{K}$ to be designed. More precisely, such feedback law corresponds to a dynamic controller

$$\chi^+=\mathbf{F}_{\ell^\star}\chi+\mathbf{L}_{\ell^\star}y+\mathbf{B}_{\ell^\star}u \tag{22a}$$

$$u=\mathbf{K}\chi \tag{22b}$$

where the matrices $\mathbf{F}_{\ell^\star}$, $\mathbf{L}_{\ell^\star}$, $\mathbf{B}_{\ell^\star}$ are completely known, see (4a). Indeed, (22a) yields

$$\chi_1^+=\chi_2,\ldots,\ \chi_{\ell^\star-1}^+=\chi_{\ell^\star},\ \chi_{\ell^\star}^+=y,$$
$$\chi_{\ell^\star+1}^+=\chi_{\ell^\star+2},\ldots,\ \chi_{\ell^\star+\ell^\star-1}^+=\chi_{\ell^\star+\ell^\star},\ \chi_{\ell^\star+\ell^\star}^+=u$$

and, when fed with input sequences $\{u(k)\}_{k=0}^\infty$ and $\{y(k)\}_{k=0}^\infty$, (22a) ensures that the solution $\chi(\cdot)$ satisfies

$$\left[\begin{array}{c} \chi_1(k)\\ \vdots\\ \chi_{\ell^\star}(k)\\ \hline \chi_{\ell^\star+1}(k)\\ \vdots\\ \chi_{\ell^\star+\ell^\star}(k)\end{array}\right]=\left[\begin{array}{c} y(k-\ell^\star)\\ \vdots\\ y(k-1)\\ \hline u(k-\ell^\star)\\ \vdots\\ u(k-1)\end{array}\right]\quad \forall k\geq \ell^\star. \tag{23}$$

This shows that (22a) creates the stack of the past $\ell^\star$ values of output and input that are needed in (21).

With all the ingredients in place, we can give our problem statement.

Problem 1. With collected data $\{u^{\mathrm{m}}(k),y^{\mathrm{m}}(k)\}_{k=0}^T$ and under Assumptions 1–2, design a matrix $\mathbf{K}$ for the dynamic controller in (22) such that the feedback interconnection of (8) and (22) ensures that $(x,\chi)=0$ is globally asymptotically stable.

As will be discussed later, the assumptions made in Problem 1 impose a constraint on the number of states, as discussed in Section 4.4. While we conclude this section by explaining our route to the results solving Problem 1, we reconsider this problem in Section 4.5 to demonstrate how these results are extended to a generic MIMO system.

In the sequel we consider the auxiliary system

$$\xi^+=\mathbf{A}_{\ell^\star}\xi+\mathbf{B}_{\ell^\star}v=(\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z_{\ell^\star})\xi+\mathbf{B}_{\ell^\star}v, \tag{24}$$

see (3). We do so because this system can capture the input–output evolution of solutions to (8) in the sense of Lemma 3, namely: under the assumptions of Lemma 3 (suitable initial condition $\xi(\ell^\star)$ and same input sequence $\{v(k)\}_{k=\ell^\star}^\infty=\{u(k)\}_{k=\ell^\star}^\infty$) we have

$$\xi(k)=\left[\begin{array}{c} y(k-\ell^\star)\\ \vdots\\ y(k-1)\\ \hline u(k-\ell^\star)\\ \vdots\\ u(k-1)\end{array}\right]\quad \forall k\geq \ell^\star. \tag{25}$$

By (23) and (25), $u(k)=\mathbf{K}\chi(k)=\mathbf{K}\xi(k)$ for all $k\geq\ell^\star$. If we set $v(k)=u(k)=\mathbf{K}\xi(k)$ for all $k\geq\ell^\star$, this corresponds, loosely speaking, to interconnecting the auxiliary system (24) with the feedback $v=u=\mathbf{K}\xi$ as

$$\xi^+=\mathbf{A}_{\ell^\star}\xi+\mathbf{B}_{\ell^\star}\mathbf{K}\xi=(\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z_{\ell^\star}+\mathbf{B}_{\ell^\star}\mathbf{K})\xi \tag{26}$$

and motivates the relevance of (24) for Problem 1. In principle, if $Z_{\ell^\star}$ were known in (26), we would like to render $\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z_{\ell^\star}+\mathbf{B}_{\ell^\star}\mathbf{K}$ Schur. In lieu of knowing $Z_{\ell^\star}$, we exploit the information available from data and embedded in the set $\mathcal{C}$, and find $\mathbf{K}$ such that $\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z+\mathbf{B}_{\ell^\star}\mathbf{K}$ is Schur for all $Z\in\mathcal{C}$, certified by a common Lyapunov function as in quadratic stabilization (Barmish, 1985). This is equivalent to the robust control problem

$$\text{find } \mathbf{K}, P=P^\top\succ 0 \tag{27a}$$

$$\text{s.t. } (\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z+\mathbf{B}_{\ell^\star}\mathbf{K})P(\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z+\mathbf{B}_{\ell^\star}\mathbf{K})^\top -P\prec 0 \qquad \forall Z\in\mathcal{C}. \tag{27b}$$

We show in Section 4.1 (see Lemma 5 below) that, under Assumption 2, we can rewrite the set $\mathcal{C}$ in a form instrumental to render $\mathbf{F}_{\ell^\star}+\mathbf{L}_{\ell^\star}Z+\mathbf{B}_{\ell^\star}\mathbf{K}$ Schur for all $Z\in\mathcal{C}$ by a common Lyapunov function. Moreover, the set $\mathcal{C}$ is bounded, which is beneficial in solving (27) (as opposed to an unbounded set $\mathcal{C}$). Section 4.2 aims at finding a semidefinite program equivalent to (27) in terms of feasibility. Section 4.3 shows rigorously how to transfer the stabilization of (24), by a $\mathbf{K}$ that satisfies (27), to the stabilization of (8), as required by Problem 1. Moreover, Section 4.4 discusses the implications of Assumption 2 and shows, as a key result (see Lemma 9 below), that observability of $(A^\star,C^\star)$ and Assumption 2 imply $p\ell^\star=n$. Based on this discussion, we conclude that our assumptions imply $p\ell^\star=n$; so, we first give a solution for the case $p\ell^\star=n$; in Section 4.5 we revisit Problem 1 and show how to extend the results to the case $p\ell^\star\neq n$.

4. Results

4.1. Reformulation of set $\mathcal{C}$ of parameters consistent with data

The set $\mathcal{C}$ in (20) can be equivalently rewritten as

$$\mathcal{C}=\{Z:\Psi_1-\mathbf{F}_{\ell^\star}\Psi_0-\mathbf{B}_{\ell^\star}U_1=\mathbf{L}_{\ell^\star}Z\Psi_0 + \mathbf{L}_{\ell^\star}[I_p\ \ -Z]\Delta,\ \ \Delta\Delta^\top\preceq\Theta\}.$$

From (17), we observe that

$$\Psi_1-\mathbf{F}_{\ell^\star}\Psi_0-\mathbf{B}_{\ell^\star}U_1=\mathbf{L}_{\ell^\star}Z_{\ell^\star}\Psi_0+\mathbf{L}_{\ell^\star}[I_p\ \ -Z_{\ell^\star}]\Delta_{10}$$

and, hence, all block rows of $\Psi_1-\mathbf{F}_{\ell^\star}\Psi_0-\mathbf{B}_{\ell^\star}U_1$ are zero except for block row $\ell^\star$. Block row $\ell^\star$ can be obtained as

$$\mathbf{L}_{\ell^\star}^\top(\Psi_1-\mathbf{F}_{\ell^\star}\Psi_0-\mathbf{B}_{\ell^\star}U_1)=\mathbf{L}_{\ell^\star}^\top\Psi_1$$

by the definition of $\mathbf{F}_{\ell^\star}$ and $\mathbf{B}_{\ell^\star}$ as in (4a). Thanks to these observations, $\mathcal{C}$ is equivalently rewritten as

$$\mathcal{C}=\{Z:\mathbf{L}_{\ell^\star}^\top\Psi_1-Z\Psi_0=[I_p\ \ -Z]\Delta,\ \ \Delta\Delta^\top\preceq\Theta\}$$

where we use $\mathbf{L}_{\ell^\star}^\top\mathbf{L}_{\ell^\star}=I_p$. To further arrange the set $\mathcal{C}$ in a convenient form, we use the next key result from (Bisoffi et al., 2024).

Proposition 1 (*Bisoffi et al. (2024, Prop. 1)*). *Consider the matrices* $E\in\mathbb{R}^{n_1\times n_2}$, $F\in\mathbb{R}^{n_1\times n_3}$, $G\in\mathbb{R}^{n_3\times n_3}$ *with* $G=G^\top\succeq 0$. *Then,*

$$EE^\top\preceq FGF^\top \tag{28a}$$

if and only if there exists $D\in\mathbb{R}^{n_3\times n_2}$ *such that*

$$E=FD,\quad DD^\top\preceq G. \tag{28b}$$

This proposition can also be proven using (van Waarde, Camlibel, Eising, & Trentelman, 2023, Lemma A.1) or Douglas' lemma (Douglas, 1966). Since going from (28b) to (28a) dispenses with matrix D, Proposition 1 can be interpreted as a matrix elimination result. By Proposition 1,

$$\mathcal{C}=\Big\{Z:(\mathbf{L}_{\ell^\star}^\top\Psi_1-Z\Psi_0)(\mathbf{L}_{\ell^\star}^\top\Psi_1-Z\Psi_0)^\top \preceq[I_p\ \ -Z]\Theta[I_p\ \ -Z]^\top\overset{(19)}{=}[I_p\ \ -Z]\begin{bmatrix}\Theta_{11} & \Theta_{12}\\ \Theta_{12}^\top & \Theta_{22}\end{bmatrix}[I_p\ \ -Z]^\top\Big\}.$$

By expanding the products and defining $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, we have

$$\mathcal{C}=\big\{Z:Z\mathscr{A}Z^\top+Z\mathscr{B}^\top+\mathscr{B}Z^\top+\mathscr{C}\preceq 0\big\}, \tag{29a}$$

$$\mathscr{A} := \Psi_0\Psi_0^\top - \Theta_{22}, \mathscr{B} := -\mathbf{L}_{\ell\star}^\top \Psi_1\Psi_0^\top + \Theta_{12}, \tag{29b}$$

$$\mathscr{C} := \mathbf{L}_{\ell\star}^\top \Psi_1\Psi_1^\top \mathbf{L}_{\ell\star} - \Theta_{11}. \tag{29c}$$

Assumption 2 amounts to asking $\mathscr{A} \succ 0$, hence, $\mathscr{A}^{-1}$ exists. We can then define

$$\mathscr{Z} := -\mathscr{B}\mathscr{A}^{-1}, \quad \mathscr{Q} := \mathscr{B}\mathscr{A}^{-1}\mathscr{B}^\top - \mathscr{C}. \tag{30}$$

Assumption 2 guarantees the next result.

Lemma 5. *Under Assumption 2, i.e., $\mathscr{A} \succ 0$, we have that:*

$$\mathcal{C} = \left\{Z : (Z - \mathscr{Z})\mathscr{A}(Z - \mathscr{Z})^\top \preceq \mathscr{Q}\right\}; \tag{31}$$

$$\mathscr{Q} \succeq 0\,;$$

$$\mathcal{C} = \left\{\mathscr{Z} + \mathscr{Q}^{1/2}\Upsilon\mathscr{A}^{-1/2} : \Upsilon\Upsilon^\top \preceq I_p\right\}; \tag{32}$$

and the set $\mathcal{C}$ is bounded.

Proof. See Appendix C. □

4.2. Stabilization of the auxiliary system

After Problem 1, we motivated the relevance of the auxiliary system (24) in the context of our setting. Since the actual $Z_{\ell\star}$ is unknown, we would like to find $\mathbf{K}$ rendering $\mathbf{F}_{\ell\star} + \mathbf{L}_{\ell\star}Z + \mathbf{B}_{\ell\star}\mathbf{K}$ Schur for all $Z \in \mathcal{C}$, i.e., solve (27). By the approach in Bisoffi et al. (2022) and the use of Petersen's lemma (Petersen, 1987), feasibility of (27) can be equivalently reformulated as in the next result.

Lemma 6. *Under Assumption 2, feasibility of (27) is equivalent to feasibility of*

$$\text{find} \qquad \mathbf{Y}, P = P^\top \succ 0 \tag{33a}$$

$$\text{s.t.} \qquad \begin{bmatrix} -P - \mathbf{L}_{\ell\star}\mathscr{C}\mathbf{L}_{\ell\star}^\top & \mathbf{F}_{\ell\star}P + \mathbf{B}_{\ell\star}\mathbf{Y} & \mathbf{L}_{\ell\star}\mathscr{B} \\ P\mathbf{F}_{\ell\star}^\top + \mathbf{Y}^\top\mathbf{B}_{\ell\star}^\top & -P & -P \\ \mathscr{B}^\top\mathbf{L}_{\ell\star}^\top & -P & -\mathscr{A} \end{bmatrix} \prec 0. \tag{33b}$$

If (33) is feasible, a $\mathbf{K}$ satisfying (27) is $\mathbf{K} = \mathbf{Y}P^{-1}$.

Proof. See Appendix D. □

The matrix inequality (33b) is equivalent, from (D.2) and by $\mathscr{A} \succ 0$ from Assumption 2 and Schur complement, to the linear matrix inequality (in $\mathbf{Y}$ and P)

$$0 \succ \begin{bmatrix} -P + \mathbf{L}_{\ell\star}\mathscr{Q}\mathbf{L}_{\ell\star}^\top & \mathbf{F}_{\ell\star}P + \mathbf{L}_{\ell\star}\mathscr{Z}P + \mathbf{B}_{\ell\star}\mathbf{Y} & 0 \\ \star & -P & P \\ \star & \star & -\mathscr{A} \end{bmatrix}.$$

4.3. From stabilization of the auxiliary system to stabilization of the actual system

To transfer the result in Lemma 6 on the stabilization of (24) to the result on the stabilization of (8), as required by Problem 1, we need to investigate the relation between their respective solutions.

We consider the generic system in (1) with (A, C) observable so that the matrices $\mathcal{O}_\ell$, $\mathcal{T}_\ell$, $\mathcal{R}_\ell$ can be defined as in (2), with $\operatorname{rank}\mathcal{O}_\ell = n$. We consider also the system in (3) with matrices $\mathbf{F}_\ell$, $\mathbf{L}_\ell$, $\mathbf{B}_\ell$, Z_ℓ selected, as in (4), from the matrices (A, B, C) of (1). System (1) is the counterpart of (8), and (3) is the counterpart of (24). We use (1) and (3) because the considerations in this section are general.

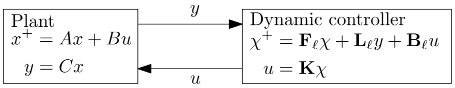

Fig. 2. Plant and dynamic controller.

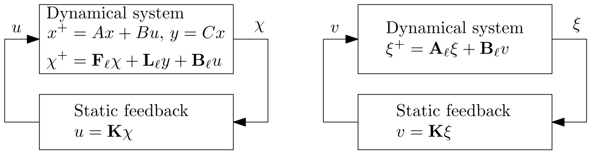

Fig. 3. (Left) Plant in (1) and dynamic controller in (34). (Right) Auxiliary system in (3) and static controller.

Suppose that there exists a control law $v = \mathbf{K}\xi$ that asymptotically stabilizes (3), i.e., that makes the matrix $\mathbf{A}_\ell + \mathbf{B}_\ell\mathbf{K} = \mathbf{F}_\ell + \mathbf{L}_\ell Z_\ell + \mathbf{B}_\ell\mathbf{K}$ Schur. Our goal for this section is to show that the same $\mathbf{K}$ can asymptotically stabilize (1) if the dynamic controller

$$\chi^+ = \mathbf{F}_\ell\chi + \mathbf{L}_\ell y + \mathbf{B}_\ell u \tag{34a}$$

$$u = \mathbf{K}\chi \tag{34b}$$

is put in feedback with (1), as depicted in Fig. 2. The scheme in Fig. 2 can be equivalently rearranged as on the left of Fig. 3, which leads us to consider the combination of (1) and (34a):

$$x^+ = Ax + Bu, \quad y = Cx, \tag{35a}$$

$$\chi^+ = \mathbf{F}_\ell\chi + \mathbf{L}_\ell y + \mathbf{B}_\ell u. \tag{35b}$$

We have the relation for the input–output evolution of (3) and (35) as in the next result.

Lemma 7. *Let (A, C) be observable. For each $\hat{x}$, $\hat{\chi}$ and sequence $\{u(k)\}_{k=0}^\infty$, there exists $\hat{\xi}$ such that:*

- *the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot) \end{smallmatrix}\right]$ to (35) with initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0) \end{smallmatrix}\right] = \left[\begin{smallmatrix} \hat{x} \\ \hat{\chi} \end{smallmatrix}\right]$ and with input $\{u(k)\}_{k=0}^\infty$,*
- *the corresponding output response $y(\cdot) = Cx(\cdot)$, and*
- *the solution $\xi(\cdot)$ to (3) with initial condition $\xi(\ell) = \hat{\xi}$ and input $\{v(k)\}_{k=\ell}^\infty = \{u(k)\}_{k=\ell}^\infty$*

satisfy

$$\xi(k) = \begin{bmatrix} y(k-\ell) \\ \vdots \\ y(k-1) \\ \hline u(k-\ell) \\ \vdots \\ u(k-1) \end{bmatrix} = \chi(k) \quad \forall k \geq \ell.$$

Proof. See Appendix E. □

With Lemma 7 in place, we can turn to the core result of this section, which is the next one.

Lemma 8. *Let (A, C) be observable. If $\mathbf{K}$ makes the matrix $\mathbf{A}_\ell + \mathbf{B}_\ell\mathbf{K} = \mathbf{F}_\ell + \mathbf{L}_\ell Z_\ell + \mathbf{B}_\ell\mathbf{K}$ Schur, then $(x, \chi) = 0$ is globally asymptotically stable for*

$$x^+ = Ax + Bu, \quad y = Cx \tag{36a}$$

$$\chi^+ = \mathbf{F}_\ell\chi + \mathbf{L}_\ell y + \mathbf{B}_\ell u \tag{36b}$$

$$u = \mathbf{K}\chi. \tag{36c}$$

Proof. See Appendix F. □

By standard properties of linear systems, an immediate consequence of Lemma 8 is that the matrix $\left[\begin{smallmatrix} A & B\mathbf{K} \\ \mathbf{L}_\ell C & \mathbf{F}_\ell + \mathbf{B}_\ell\mathbf{K} \end{smallmatrix}\right]$, which corresponds to the closed-loop system in (36), is Schur. The next result is obtained immediately from Lemma 6, Lemma 7, Lemma 8.

Theorem 1. *With collected data* $\{u^{\mathrm{m}}(k), y^{\mathrm{m}}(k)\}_{k=0}^{T}$ *and under Assumptions 1–2, suppose (33) is feasible and* $\mathbf{K}$ *is a controller returned from (33). Then,* $(x, \chi) = 0$ *is globally asymptotically stable for the feedback interconnection of the unknown plant*

$$x^+ = A^\star x + B^\star u, \;\; y = C^\star x \tag{37a}$$

and the controller

$$\chi^+ = \mathbf{F}_{\ell^\star}\chi + \mathbf{L}_{\ell^\star} y + \mathbf{B}_{\ell^\star} u \tag{37b}$$

$$u = \mathbf{K}\chi. \tag{37c}$$

Moreover, for each $k \geq \ell^\star$,

$$\chi(k) = \big(y(k-\ell^\star), \ldots, y(k-1), u(k-\ell^\star), \ldots, u(k-1)\big).$$

Proof. Assumption 2 and Lemma 6 ensure that a $\mathbf{K}$ returned from (33) makes $\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z + \mathbf{B}_{\ell^\star}\mathbf{K}$ Schur for all $Z \in \mathcal{C}$, and in particular for $Z = Z_{\ell^\star}$, corresponding to $(A^\star, B^\star, C^\star)$. Since $\mathbf{A}_{\ell^\star} + \mathbf{B}_{\ell^\star}\mathbf{K}$ is Schur, Assumption 1 and Lemma 8 ensure that $(x, \chi) = 0$ is globally asymptotically stable for (37). Moreover, Assumption 1 and Lemma 7 ensure that $\chi(k) = \big(y(k-\ell^\star), \ldots, y(k-1), u(k-\ell^\star), \ldots, u(k-1)\big)$ for each $k \geq \ell^\star$. $\square$

Theorem 1 provides our solution to Problem 1, since (37b) and (37c) are equivalent to (22a) and (22b).

4.4. Interpretation and implications of Assumption 2

To discuss Assumption 2, we introduce

$$N_0 := \left[\begin{array}{cccc} d^y(0) & d^y(1) & \ldots & d^y(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^y(\ell^\star-1) & d^y(\ell^\star) & \ldots & d^y(T-1) \\ \hline d^u(0) & d^u(1) & \ldots & d^u(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^u(\ell^\star-1) & d^u(\ell^\star) & \ldots & d^u(T-1) \end{array}\right] \tag{38a}$$

$$S_0 := \left[\begin{array}{cccc} y(0) & y(1) & \ldots & y(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ y(\ell^\star-1) & y(\ell^\star) & \ldots & y(T-1) \\ \hline u(0) & u(1) & \ldots & u(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ u(\ell^\star-1) & u(\ell^\star) & \ldots & u(T-1) \end{array}\right] \tag{38b}$$

where S_0 corresponds to the "clean" data we would collect from (8) if input and output noise d^u and d^y were absent. From (9) and (11), we have that

$$\Psi_0 = S_0 + N_0. \tag{39}$$

The next result shows that observability of $(A^\star, C^\star)$ and Assumption 2 imply $p\ell^\star = n$.

Lemma 9. *Let* $(A^\star, C^\star)$ *be observable. Then,*

$$\text{Assumption 2} \implies \operatorname{rank} S_0 = (p+m)\ell^\star$$
$$\implies \operatorname{rank} \mathcal{O}_{\ell^\star} = p\ell^\star \implies p\ell^\star = n.$$

Proof. See Appendix G. $\square$

By Lemma 9, if Assumption 2 holds, then we know that we are in the case $p\ell^\star = n$, without the need to know n. Moreover, a sufficient condition for Assumption 2 is provided next, which is of signal-to-noise-ratio type as it involves a ratio between eigenvalues of the matrix $S_0 S_0^\top$ of "clean" data and the matrix Θ_{22} of the noise bound.

Lemma 10. *If* $\lambda_{\min}(S_0 S_0^\top) > 4\lambda_{\max}(\Theta_{22})$, *then Assumption 2 holds.*

Proof. See Appendix H. $\square$

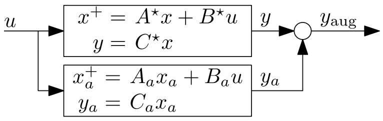

Fig. 4. Parallel connection of original and artificial systems.

4.5. Extension to the case $p\ell^\star > n$

Observability of $(A^\star, C^\star)$ in Assumption 1 implies that $p\ell^\star \geq n$. As shown in Section 4.3, under Assumption 2, the controller proposed in Theorem 1 stabilizes the actual system in (8). By Lemma 9, this assumption also implies that $p\ell^\star = n$. However, if Assumption 2 does not hold, the previous treatment (e.g., Lemma 6) cannot be applied. In case Assumption 2 is violated, the signal-to-noise ratio of the collected data is insufficient (cf. Lemma 10) or we are in the case $p\ell^\star > n$. While the former case suggests recollecting data of potentially better quality to meet the signal-to-noise requirement, the latter case can be addressed by suitably extending the previous scheme. To this end, we impose the next assumption related to the state dimension.

Assumption 3. The state dimension n in (8) is known.

The purpose of this assumption is to directly verify the condition $p\ell^\star > n$ and quantify the mismatch $p\ell^\star - n$. We propose to augment the system (8) with additional dynamics of order $p\ell^\star - n$ as

$$\begin{bmatrix} x^+ \\ x_a^+ \end{bmatrix} = \underbrace{\begin{bmatrix} A^\star & 0 \\ 0 & A_a \end{bmatrix}}_{=:A_{\mathrm{aug}}} \begin{bmatrix} x \\ x_a \end{bmatrix} + \underbrace{\begin{bmatrix} B^\star \\ B_a \end{bmatrix}}_{=:B_{\mathrm{aug}}} u \tag{40a}$$

$$y_{\mathrm{aug}} = \underbrace{\begin{bmatrix} C^\star & C_a \end{bmatrix}}_{=:C_{\mathrm{aug}}} \begin{bmatrix} x \\ x_a \end{bmatrix} \tag{40b}$$

with artificial state/input/output matrices

$$A_a \in \mathbb{R}^{(p\ell^\star-n)\times(p\ell^\star-n)},\, B_a \in \mathbb{R}^{(p\ell^\star-n)\times m},\, C_a \in \mathbb{R}^{p\times(p\ell^\star-n)}.$$

The augmented system can be viewed as the parallel connection of the original system $(A^\star, B^\star, C^\star)$ and the introduced artificial system (A_a, B_a, C_a), as in Fig. 4.

The idea pursued here is to apply the controller synthesis procedure discussed before to the augmented system (40), which is of dimension $n_{\mathrm{aug}} := p\ell^\star$. So, consider

$$\mathcal{O}_{\ell^\star}^{\mathrm{aug}} := \begin{bmatrix} C^\star & C_a \\ C^\star A^\star & C_a A_a \\ \vdots & \vdots \\ C^\star {A^\star}^{\ell^\star-1} & C_a A_a^{\ell^\star-1} \end{bmatrix} \in \mathbb{R}^{n_{\mathrm{aug}}\times n_{\mathrm{aug}}}, \tag{41}$$

i.e., the $\ell^\star$-step observability matrix of the augmented system (40), and make the next assumption on $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$.

Assumption 4. The selected artificial state and output matrices A_a and C_a are such that the matrix $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ is nonsingular.

We note that if $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ is nonsingular, $\mathcal{O}_{\ell^\star}$ (corresponding to the first n columns of $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$) must be full column rank and, thus, Assumption 4 subsumes Assumption 1. In the sequel we impose a data-dependent condition that serves as a sufficient condition for Assumption 4, see Lemma 11.

Remark 3. The artificial (A_a, C_a) should be selected such that the augmentation in (41), see Fig. 4, does not spoil the observability of the original $(A^\star, C^\star)$, and the observability index remains equal to $\ell^\star$. To guarantee the first condition, it is sufficient to select an observable pair (A_a, C_a) with $\mathrm{eig}(A^\star) \cap \mathrm{eig}(A_a) = \emptyset$. Such a selection can be done if a strict lower bound on the smallest modulus of the eigenvalues of $A^\star$ is available. As for the second requirement, namely, the observability index remaining equal to $\ell^\star$, we note that $\mathcal{O}^{\mathrm{aug}}_{\ell^\star}$ is a square submatrix of the observability matrix of the interconnected system in Fig. 4. Our numerical investigation shows that arbitrarily selecting observable (A_a, C_a) renders this submatrix nonsingular. A more rigorous treatment of this selection remains an open question. ■

As mentioned before, we introduce the artificial system to bring the setup back to the case where the state dimension is equal to $p\ell^\star$. An immediate obstacle is that data collection does not exactly match the previous case in Fig. 1. This obstacle is discussed and addressed next. Due to the presence of noise, we can collect input–output data from the augmented system (40) through the noisy augmented dynamics

$$\begin{bmatrix} x^+ \\ \overset{\mathrm{m}}{x}{}_a^{+} \end{bmatrix} = \begin{bmatrix} A^\star & 0 \\ 0 & A_a \end{bmatrix} \begin{bmatrix} x \\ \overset{\mathrm{m}}{x}_a \end{bmatrix} + \begin{bmatrix} B^\star \\ B_a \end{bmatrix} \overset{\mathrm{m}}{u} - \begin{bmatrix} B^\star \\ 0 \end{bmatrix} d^u \tag{42a}$$

$$\overset{\mathrm{m}}{y}_{\mathrm{aug}} = \begin{bmatrix} C^\star & C_a \end{bmatrix} \begin{bmatrix} x \\ \overset{\mathrm{m}}{x}_a \end{bmatrix} + d^y \overset{(10)}{=} \overset{\mathrm{m}}{y} + C_a \overset{\mathrm{m}}{x}_a \tag{42b}$$

where the superscript "m" denotes the quantities that can be measured. The input noise d^u affects only the dynamics of the actual system and not the artificial one, and (42a) can be rewritten as

$$\begin{bmatrix} x^+ \\ \overset{\mathrm{m}}{x}{}_a^{+} \end{bmatrix} = \begin{bmatrix} A^\star & 0 \\ 0 & A_a \end{bmatrix} \begin{bmatrix} x \\ \overset{\mathrm{m}}{x}_a \end{bmatrix} + \begin{bmatrix} B^\star \\ B_a \end{bmatrix} (\overset{\mathrm{m}}{u} - d^u) + \begin{bmatrix} 0 \\ B_a \end{bmatrix} d^u.$$

Note that neglecting the last term on the right hand side of the previous equation would result in the same setup as in Fig. 1, where the actual system $(A^\star, B^\star, C^\star)$ is replaced by the augmented one $(A_{\mathrm{aug}}, B_{\mathrm{aug}}, C_{\mathrm{aug}})$. While that term cannot be neglected, it can be pushed to the output channel as

$$\begin{bmatrix} x^+ \\ x_a^+ \end{bmatrix} = \begin{bmatrix} A^\star & 0 \\ 0 & A_a \end{bmatrix} \begin{bmatrix} x \\ x_a \end{bmatrix} + \begin{bmatrix} B^\star \\ B_a \end{bmatrix} (\overset{\mathrm{m}}{u} - d^u) \tag{43a}$$

$$\hat{\overset{\mathrm{m}}{y}}_{\mathrm{aug}} := \begin{bmatrix} C^\star & C_a \end{bmatrix} \begin{bmatrix} x \\ x_a \end{bmatrix} + d^y + d^{y_a} \tag{43b}$$

where, for each $k \geq 0$,

$$d^{y_a}(k) := \sum_{j=0}^{k-1} C_a A_a^{k-j-1} B_a d^u(j). \tag{43c}$$

This leads to different state variables $\overset{\mathrm{m}}{x}_a$ and x_a in (42a) and (43a), respectively. Still, the input–output behavior of (42) and (43) is identical because, starting from the same initial condition $(x(0), \overset{\mathrm{m}}{x}_a(0)) = (x(0), x_a(0))$ and applying the same input/noise sequences $\{\overset{\mathrm{m}}{u}(k)\}_{k=0}^\infty$, $\{d^u(k)\}_{k=0}^\infty$, $\{d^y(k)\}_{k=0}^\infty$, we have

$$\{\overset{\mathrm{m}}{y}_{\mathrm{aug}}(k)\}_{k=0}^\infty = \{\hat{\overset{\mathrm{m}}{y}}_{\mathrm{aug}}(k)\}_{k=0}^\infty, \tag{44}$$

see also Fig. 5.

In summary, bearing in mind (43) and (44), we can proceed to work with the augmented system with state $x_{\mathrm{aug}} := (x, x_a)$ and $(A_{\mathrm{aug}}, B_{\mathrm{aug}}, C_{\mathrm{aug}})$ defined in (40):

$$x_{\mathrm{aug}}^+ = A_{\mathrm{aug}} x_{\mathrm{aug}} + B_{\mathrm{aug}} (\overset{\mathrm{m}}{u} - d^u) \tag{45a}$$

$$\overset{\mathrm{m}}{y}_{\mathrm{aug}} = C_{\mathrm{aug}} x_{\mathrm{aug}} + d^y + d^{y_a} =: C_{\mathrm{aug}} x_{\mathrm{aug}} + d^{y_{\mathrm{aug}}}, \tag{45b}$$

which is the counterpart of (10). For the measured data $\{\overset{\mathrm{m}}{u}(k), \overset{\mathrm{m}}{y}_{\mathrm{aug}}(k)\}_{k=0}^T$, we construct the data matrices

$$\Psi_1^{\mathrm{aug}} := \left[\begin{array}{cccc} \overset{\mathrm{m}}{y}_{\mathrm{aug}}(1) & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(2) & \ldots & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(T-\ell^\star+1) \\ \vdots & \vdots & & \vdots \\ \overset{\mathrm{m}}{y}_{\mathrm{aug}}(\ell^\star) & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(\ell^\star+1) & \ldots & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(T) \\ \hline \overset{\mathrm{m}}{u}(1) & \overset{\mathrm{m}}{u}(2) & \ldots & \overset{\mathrm{m}}{u}(T-\ell^\star+1) \\ \vdots & \vdots & & \vdots \\ \overset{\mathrm{m}}{u}(\ell^\star) & \overset{\mathrm{m}}{u}(\ell^\star+1) & \ldots & \overset{\mathrm{m}}{u}(T) \end{array}\right] \tag{46a}$$

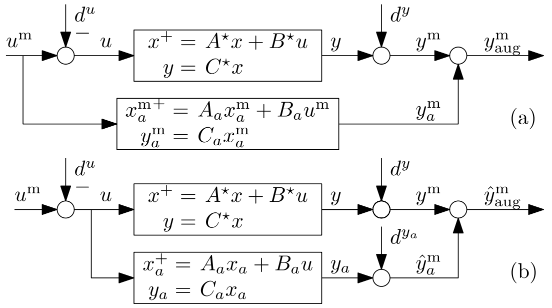

Fig. 5. Data collection on the augmented system. Subfigures (a) and (b) correspond respectively to (42) and (43).

$$\Psi_0^{\mathrm{aug}} := \left[\begin{array}{cccc} \overset{\mathrm{m}}{y}_{\mathrm{aug}}(0) & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(1) & \ldots & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ \overset{\mathrm{m}}{y}_{\mathrm{aug}}(\ell^\star-1) & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(\ell^\star) & \ldots & \overset{\mathrm{m}}{y}_{\mathrm{aug}}(T-1) \\ \hline \overset{\mathrm{m}}{u}(0) & \overset{\mathrm{m}}{u}(1) & \ldots & \overset{\mathrm{m}}{u}(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ \overset{\mathrm{m}}{u}(\ell^\star-1) & \overset{\mathrm{m}}{u}(\ell^\star) & \ldots & \overset{\mathrm{m}}{u}(T-1) \end{array}\right] \tag{46b}$$

$$U_1^{\mathrm{aug}} := \left[\, \overset{\mathrm{m}}{u}(\ell^\star)\ \overset{\mathrm{m}}{u}(\ell^\star+1)\ \ldots\ \overset{\mathrm{m}}{u}(T) \,\right] \tag{46c}$$

$$\Delta_{10}^{\mathrm{aug}} := \left[\begin{array}{cccc} d^{y_{\mathrm{aug}}}(\ell^\star) & d^{y_{\mathrm{aug}}}(\ell^\star+1) & \ldots & d^{y_{\mathrm{aug}}}(T) \\ \hline d^{y_{\mathrm{aug}}}(0) & d^{y_{\mathrm{aug}}}(1) & \ldots & d^{y_{\mathrm{aug}}}(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^{y_{\mathrm{aug}}}(\ell^\star-1) & d^{y_{\mathrm{aug}}}(\ell^\star) & \ldots & d^{y_{\mathrm{aug}}}(T-1) \\ \hline d^u(0) & d^u(1) & \ldots & d^u(T-\ell^\star) \\ \vdots & \vdots & & \vdots \\ d^u(\ell^\star-1) & d^u(\ell^\star-1) & \ldots & d^u(T-1) \end{array}\right]. \tag{46d}$$

These matrices are the counterparts of those in (11) and satisfy the counterpart of (17), namely,

$$\Psi_1^{\mathrm{aug}} = (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z_{\ell^\star}^{\mathrm{aug}}) \Psi_0^{\mathrm{aug}} + \mathbf{B}_{\ell^\star} U_1^{\mathrm{aug}} + \mathbf{L}_{\ell^\star} \begin{bmatrix} I_p & -Z_{\ell^\star}^{\mathrm{aug}} \end{bmatrix} \Delta_{10}^{\mathrm{aug}} \tag{47}$$

where $\mathbf{F}_{\ell^\star}$, $\mathbf{L}_{\ell^\star}$ and $\mathbf{B}_{\ell^\star}$ are as in (4a), with $\ell = \ell^\star$, and

$$\mathcal{T}_{\ell^\star}^{\mathrm{aug}} := \begin{bmatrix} 0_{p\times m} & 0 & \ldots & 0 & 0 \\ C_{\mathrm{aug}} B_{\mathrm{aug}} & 0 & \ldots & 0 & 0 \\ C_{\mathrm{aug}} A_{\mathrm{aug}} B_{\mathrm{aug}} & C_{\mathrm{aug}} B_{\mathrm{aug}} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{\mathrm{aug}} A_{\mathrm{aug}}^{\ell^\star-2} B_{\mathrm{aug}} & C_{\mathrm{aug}} A_{\mathrm{aug}}^{\ell^\star-3} B_{\mathrm{aug}} & \ldots & C_{\mathrm{aug}} B_{\mathrm{aug}} & 0_{p\times m} \end{bmatrix}$$

$$\mathcal{R}_{\ell^\star}^{\mathrm{aug}} := \begin{bmatrix} A_{\mathrm{aug}}^{\ell^\star-1} B_{\mathrm{aug}} & \ldots & A_{\mathrm{aug}} B_{\mathrm{aug}} & B_{\mathrm{aug}} \end{bmatrix}$$

$$Z_{\ell^\star}^{\mathrm{aug}} := \left[\, C_{\mathrm{aug}} A_{\mathrm{aug}}^{\ell^\star} (\mathcal{O}_{\ell^\star}^{\mathrm{aug}})^{-1} \quad C_{\mathrm{aug}} \mathcal{R}_{\ell^\star}^{\mathrm{aug}} - C_{\mathrm{aug}} A_{\mathrm{aug}}^{\ell^\star} (\mathcal{O}_{\ell^\star}^{\mathrm{aug}})^{-1} \mathcal{T}_{\ell^\star}^{\mathrm{aug}} \,\right] \tag{48}$$

with $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ in (41), which is invertible by Assumption 4. We stress that the matrix $Z_{\ell^\star}^{\mathrm{aug}}$ still contains the actual matrices $A^\star$, $B^\star$ and $C^\star$, so it is unknown to us.

As we did in Section 3, we assume that the noise sequence $\Delta_{10}^{\mathrm{aug}}$ in (46d) satisfies an energy bound, i.e., $\Delta_{10}^{\mathrm{aug}} \in \mathcal{D}_{\mathrm{aug}}$ with $\mathcal{D}_{\mathrm{aug}}$ defined, for some $\Theta^{\mathrm{aug}} = \Theta^{\mathrm{aug}\top} \succeq 0$, as

$$\mathcal{D}_{\mathrm{aug}} := \{\Delta \in \mathbb{R}^{(p+p\ell^\star+m\ell^\star)\times(T-\ell^\star+1)} : \Delta\Delta^\top \preceq \Theta^{\mathrm{aug}}\}. \tag{49}$$

We partition Θ^{aug} as

$$\Theta^{\mathrm{aug}} =: \begin{bmatrix} \Theta_{11}^{\mathrm{aug}} & \Theta_{12}^{\mathrm{aug}} \\ \Theta_{12}^{\mathrm{aug}\top} & \Theta_{22}^{\mathrm{aug}} \end{bmatrix} \tag{50}$$

with $\Theta_{11}^{\mathrm{aug}} = \Theta_{11}^{\mathrm{aug}\top} \in \mathbb{R}^{p\times p}$ and $\Theta_{22}^{\mathrm{aug}} = \Theta_{22}^{\mathrm{aug}\top} \in \mathbb{R}^{(p\ell^\star+m\ell^\star)\times(p\ell^\star+m\ell^\star)}$. Following Remark 2, we present a way to obtain Θ^{aug} in the next remark.

Remark 4. Suppose we know that for some $\bar{d}^y \geq 0$ and $\bar{d}^u \geq 0$,

$$\|d^y\|_{\mathcal{L}_\infty} \leq \bar{d}^y \text{ and } \|d^u\|_{\mathcal{L}_\infty} \leq \bar{d}^u. \tag{51}$$

For $\Delta_{10}^{\mathrm{aug}}$ in (46d), we have $d^{y_{\mathrm{aug}}} = d^{y} + d^{y_a}$, see (45b), with d^{y_a} as in (43c). If we can find $\bar{d}^{y_a}$ such that

$$\|d^{y_a}\|_{\mathcal{L}_\infty} \leq \bar{d}^{y_a}, \tag{52}$$

the very same steps in Remark 2 would allow us to take Θ^{aug} as $(T - \ell^\star + 1)\big((\ell^\star + 1)(\bar{d}^{y} + \bar{d}^{y_a})^2 + \ell^\star(\bar{d}^{u})^2\big)I$. By choosing any Schur stable matrix A_a, the artificial system has a finite peak-to-peak gain ($\mathcal{L}_\infty$ induced-norm) (Antsaklis & Michel, 2005, §6.10, Thm. 10.17) and, thus, from (43c),

$$\|d^{y_a}\|_{\mathcal{L}_\infty} \leq \eta \|d^{u}\|_{\mathcal{L}_\infty} \tag{53}$$

for some $\eta > 0$; $\bar{d}^{y_a}$ in (52) can then be chosen as $\eta\bar{d}^{u}$. ■

We now impose the next condition, which is the counterpart of Assumption 2.

Assumption 5. $\Psi_0^{\mathrm{aug}}\Psi_0^{\mathrm{aug}\top} \succ \Theta_{22}^{\mathrm{aug}}$.

The next lemma contains some relevant observations on the assumptions we have made so far and on the observability index of the augmented system.

Lemma 11. *If $\Psi_0^{\mathrm{aug}}\Psi_0^{\mathrm{aug}\top} \succ \Theta_{22}^{\mathrm{aug}}$ as per Assumption 5, then $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ is nonsingular as per Assumption 4, and $(A_{\mathrm{aug}}, C_{\mathrm{aug}})$ is observable with observability index equal to $\ell^\star$.*

Proof. See Appendix I. □

We discuss the consequences of Lemma 11. First, the *data-dependent* condition in Assumption 5 is a sufficient condition for Assumption 4; thus Assumption 4 is not invoked in the subsequent results. Second, the fact that the observability index of $(A_{\mathrm{aug}}, C_{\mathrm{aug}})$ is equal to $\ell^\star$ allows us to fall back to the setting of Sections 4.1–4.3 for the *augmented* system.

Since the matrix $Z_{\ell^\star}^{\mathrm{aug}}$ in (48) is unknown, we introduce the set of matrices consistent with data in (47) and with the noise bound in (49), namely,

$$\begin{aligned}\mathcal{C}_{\mathrm{aug}} := \Big\{ & Z \in \mathbb{R}^{p\times(p\ell^\star + m\ell^\star)} : \Psi_1^{\mathrm{aug}} = (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z)\Psi_0^{\mathrm{aug}} \\ & + \mathbf{B}_{\ell^\star} U_1^{\mathrm{aug}} + \mathbf{L}_{\ell^\star}[I_p \ \ -Z]\Delta, \ \ \Delta \in \mathcal{D}_{\mathrm{aug}} \Big\}. \end{aligned} \tag{54}$$

It is worth noting that $Z_{\ell^\star}^{\mathrm{aug}} \in \mathcal{C}_{\mathrm{aug}}$ because $\Delta_{10}^{\mathrm{aug}} \in \mathcal{D}_{\mathrm{aug}}$. By the same steps leading from (20) to (29), the set $\mathcal{C}_{\mathrm{aug}}$ can be equivalently rewritten as

$$\mathcal{C}_{\mathrm{aug}} = \big\{ Z : Z\mathscr{A}_{\mathrm{aug}}Z^\top + Z\mathscr{B}_{\mathrm{aug}}{}^\top + \mathscr{B}_{\mathrm{aug}}Z^\top + \mathscr{C}_{\mathrm{aug}} \preceq 0 \big\}, \tag{55a}$$

$$\mathscr{A}_{\mathrm{aug}} := \Psi_0^{\mathrm{aug}}\Psi_0^{\mathrm{aug}\top} - \Theta_{22}^{\mathrm{aug}}, \tag{55b}$$

$$\mathscr{B}_{\mathrm{aug}} := -\mathbf{L}_{\ell^\star}^\top \Psi_1^{\mathrm{aug}}\Psi_0^{\mathrm{aug}\top} + \Theta_{12}^{\mathrm{aug}}, \tag{55c}$$

$$\mathscr{C}_{\mathrm{aug}} := \mathbf{L}_{\ell^\star}^\top \Psi_1^{\mathrm{aug}}\Psi_1^{\mathrm{aug}\top}\mathbf{L}_{\ell^\star} - \Theta_{11}^{\mathrm{aug}}. \tag{55d}$$

We would like to find $\mathbf{K}$ such that $\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star}Z + \mathbf{B}_{\ell^\star}\mathbf{K}$ is Schur for all $Z \in \mathcal{C}_{\mathrm{aug}}$, certified by a common Lyapunov function, i.e., to solve the robust stabilization problem

$$\text{find } \ \mathbf{K}, P = P^\top \succ 0 \tag{56a}$$

$$\begin{aligned}\text{s.t. } \ & (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star}Z + \mathbf{B}_{\ell^\star}\mathbf{K})P(\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star}Z + \mathbf{B}_{\ell^\star}\mathbf{K})^\top \\ & -P \prec 0 \qquad \forall Z \in \mathcal{C}_{\mathrm{aug}}. \end{aligned} \tag{56b}$$

The problem (56) is the counterpart of (27) and is solved next.

Lemma 12. *Under Assumption 5, feasibility of (56) is equivalent to feasibility of*

$$\textit{find} \quad \mathbf{Y}, P = P^\top \succ 0 \tag{57a}$$

$$\textit{s.t.} \quad \begin{bmatrix} -P - \mathbf{L}_{\ell^\star}\mathscr{C}_{\mathrm{aug}}\mathbf{L}_{\ell^\star}^\top & \mathbf{F}_{\ell^\star}P + \mathbf{B}_{\ell^\star}\mathbf{Y} & \mathbf{L}_{\ell^\star}\mathscr{B}_{\mathrm{aug}} \\ P\mathbf{F}_{\ell^\star}^\top + \mathbf{Y}^\top\mathbf{B}_{\ell^\star}^\top & -P & -P \\ \mathscr{B}_{\mathrm{aug}}^\top\mathbf{L}_{\ell^\star}^\top & -P & -\mathscr{A}_{\mathrm{aug}} \end{bmatrix} \prec 0. \tag{57b}$$

If (57) is feasible, a $\mathbf{K}$ satisfying (56) is $\mathbf{K} = \mathbf{Y}P^{-1}$.

Proof. The proof is the same as the proof of Lemma 6 with $\mathscr{A}_{\mathrm{aug}}$, $\mathscr{B}_{\mathrm{aug}}$, $\mathscr{C}_{\mathrm{aug}}$ substituting $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$. □

This leads to the main result of this section.

Theorem 2. *With collected data $\{\overset{\mathrm{m}}{u}(k), \overset{\mathrm{m}}{y}_{\mathrm{aug}}(k)\}_{k=0}^{T}$ and under Assumption 3 and 5, suppose (57) is feasible and $\mathbf{K}$ is a controller returned from (57). Then, $(x, x_a, \chi) = 0$ is globally asymptotically stable for the feedback interconnection of the unknown plant*

$$x^+ = A^\star x + B^\star u, \ \ y = C^\star x \tag{58}$$

and the controller

$$\begin{bmatrix} x_a^+ \\ \chi^+ \end{bmatrix} = \begin{bmatrix} A_a & B_a\mathbf{K} \\ \mathbf{L}_{\ell^\star}C_a & \mathbf{F}_{\ell^\star} + \mathbf{B}_{\ell^\star}\mathbf{K} \end{bmatrix}\begin{bmatrix} x_a \\ \chi \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{L}_{\ell^\star} \end{bmatrix} y \tag{59a}$$

$$u = \mathbf{K}\chi. \tag{59b}$$

Proof. By Lemma 12, the reasoning in Theorem 1 shows that $(x, x_a, \chi) = 0$ is globally asymptotically stable for the feedback interconnection of the augmented system

$$\begin{bmatrix} x^+ \\ x_a^+ \end{bmatrix} = \begin{bmatrix} A^\star & 0 \\ 0 & A_a \end{bmatrix}\begin{bmatrix} x \\ x_a \end{bmatrix} + \begin{bmatrix} B^\star \\ B_a \end{bmatrix} u, \ \ y_{\mathrm{aug}} = C^\star x + C_a x_a$$

and the controller

$$\chi^+ = \mathbf{F}_{\ell^\star}\chi + \mathbf{L}_{\ell^\star}y_{\mathrm{aug}} + \mathbf{B}_{\ell^\star}u, \ \ u = \mathbf{K}\chi.$$

This interconnection is equivalent to (58), (59). □

5. Numerical examples

In this section we illustrate Theorems 1 and 2.

5.1. Example for the case $p\ell^\star = n$

We consider the unstable batch reactor process in Green and Limebeer (2012, §2.6), which can be discretized by the zero-order-hold method (with sample period of 0.2 s) to obtain

$$\left[\begin{array}{c|c} A^\star & B^\star \\ \hline C^\star & 0 \end{array}\right] = \left[\begin{array}{cccc|cc} 1.427 & 0.039 & 0.854 & -0.622 & 0.034 & -0.305 \\ -0.096 & 0.455 & -0.034 & 0.109 & 0.787 & 0.008 \\ 0.115 & 0.538 & 0.384 & 0.529 & 0.571 & -0.380 \\ -0.012 & 0.537 & 0.122 & 0.777 & 0.570 & -0.050 \\ \hline 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right]$$

where $n = 4$, $m = 2$, $p = 2$ and $\ell^\star = 2$. The values of $(A^\star, B^\star, C^\star)$ are used only for data generation, but not for controller design. According to (12),

$$Z_{\ell^\star} = \begin{bmatrix} -0.374 & -0.714 & 1.870 & 1.870 & -1.311 & 0.317 & 0.035 & -0.634 \\ -0.016 & -0.289 & -0.037 & 1.173 & -0.524 & 0.007 & 0.787 & 0.008 \end{bmatrix}.$$

Let $\omega \sim U[a, b]$ denote a random variable ω uniformly distributed in the interval $[a, b]$. We apply inputs $\overset{\mathrm{m}}{u}_1, \overset{\mathrm{m}}{u}_2 \sim U[-20, 20]$, input noise $d_1^u, d_2^u \sim U[-0.01, 0.01]$ and output noise $d_1^y, d_2^y \sim U[-0.01, 0.01]$ to collect 40 data points: 10 experiments are performed each consisting of 4 data points, since the system is unstable. We store the data points into Ψ_0, $\Psi_1 \in \mathbb{R}^{8\times 20}$, $U_1 \in \mathbb{R}^{2\times 20}$ as in (11). Bearing in mind (18), the upper bounds on the noise components (i.e., 0.01 for input noise, and 0.01 for output noise), and the size of the noise data matrix Δ_{10} (i.e., 10×20), we have $\varTheta = 0.02 I_{10}$, from Remark 2. Assumption 2 is satisfied as $\lambda_{\min}(\varPsi_0 \varPsi_0^\top - \varTheta_{22}) = 0.220 > 0$. By Lemma 6, solving (33) via CVX (Grant & Boyd, 2008, 2014) yields

$$\mathbf{K}=\begin{bmatrix} -0.005 & 3.585 & 0.134 & -9.281 & 6.428 & -0.036 & -0.841 & 0.003 \\ -1.079 & -3.117 & 5.431 & 7.980 & -5.665 & 0.957 & 0.240 & -1.448 \end{bmatrix}.$$

The eigenvalues of the closed-loop system (37) are

$$\operatorname{eig}\left(\begin{bmatrix} A^\star & B^\star\mathbf{K} \\ \mathbf{L}_{\ell^\star} C^\star & \mathbf{F}_{\ell^\star}+\mathbf{B}_{\ell^\star}\mathbf{K} \end{bmatrix}\right)=\{0.467,\ -0.048,\ 0,\ 0,\ 0.040 \pm i0.526,\ 0.288 \pm i0.345,\ -0.161 \pm i0.084,\ 0,\ 0\}$$

and are all in the open unit disk. The closed-loop system (37) is asymptotically stable as per Theorem 1.

5.2. Example for the case $p\ell^\star > n$

Consider the marginally stable discrete-time system given by

$$\left[\begin{array}{c|c} A^\star & B^\star \\ \hline C^\star & 0 \end{array}\right] = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right].$$

In this case, $n = 3$, $m = 2$, $p = 2$, $\ell^\star = 2$, and $p\ell^\star = 4 > 3 = n$, thus, we need to use the results in Section 4.5, namely, Lemma 12 and Theorem 2. We choose an artificial system with state dimension $4 - 3 = 1$ and

$$\left[\begin{array}{c|c} A_a & B_a \\ \hline C_a & 0 \end{array}\right] = \left[\begin{array}{c|cc} 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right].$$

The augmented system is then

$$\left[\begin{array}{c|c} A_{\text{aug}} & B_{\text{aug}} \\ \hline C_{\text{aug}} & 0 \end{array}\right] = \left[\begin{array}{cccc|cc} 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right].$$

According to (48), $Z_{\ell^\star}^{\text{aug}} = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & -1 & 3 & 1 \\ 1 & -1 & 0 & 1 & -2 & -2 & 2 & 2 \end{bmatrix}$.

We apply inputs $\overset{\text{m}}{u}_1, \overset{\text{m}}{u}_2 \sim U[-2, 2]$, input noise $d_1^u, d_2^u \sim U[-0.01, 0.01]$ and output noise $d_1^y, d_2^y \sim U[-0.01, 0.01]$ to collect 32 data points. The data collection scheme is the one in Fig. 5(a). We store the data points into $\varPsi_0^{\text{aug}}, \varPsi_1^{\text{aug}} \in \mathbb{R}^{8\times 30}$, $U_1^{\text{aug}} \in \mathbb{R}^{2\times 30}$ as in (46). By $A_a = 0$, the quantity η in (53) can be computed, see (43c), as $\|C_a B_a\| = 2$, which yields $\bar{d}^{y_a} = \sqrt{2}/50$ in (52). Bearing in mind (49), the upper bounds on the noise components (i.e., 0.01 for input noise, and 0.01 for output noise), $\bar{d}^{y_a} = \sqrt{2}/50$, and the size of the noise data matrix $\varDelta_{10}^{\text{aug}}$ (i.e., 10×30), we have $\varTheta^{\text{aug}} = 0.174 I_{10}$, from Remark 4. It can be verified that $\lambda_{\min}(\varPsi_0^{\text{aug}} \varPsi_0^{\text{aug}\top} - \varTheta_{22}^{\text{aug}}) = 2.960 > 0$, i.e., data satisfy Assumption 5. Solving (57) in Lemma 12 yields

$$\mathbf{K} = \begin{bmatrix} -0.138 & 0.144 & 0.076 & -0.218 & 0.371 & 0.358 & -0.255 & -0.159 \\ -0.096 & 0.097 & -0.127 & 0.135 & 0.073 & 0.069 & 0.253 & -0.081 \end{bmatrix}.$$

The eigenvalues of the closed-loop system (58), (59) are

$$\operatorname{eig}\left(\begin{bmatrix} A^\star & 0 & B^\star\mathbf{K} \\ 0 & A_a & B_a\mathbf{K} \\ \mathbf{L}_{\ell^\star} C^\star & \mathbf{L}_{\ell^\star} C_a & \mathbf{F}_{\ell^\star}+\mathbf{B}_{\ell^\star}\mathbf{K} \end{bmatrix}\right)=\{0.837,\ 0.358,\ -0.148,\ -0.005,\ 0,\ 0,\ 0,\ 0,\ -0.004 \pm i0.766,\ -0.185 \pm i0.040\}$$

and are all in the open unit disk. The closed-loop system is asymptotically stable as per Theorem 2.

Finally, we examine the effect of measurement noise *during* the execution of the controller. In this case, the closed-loop system (58), (59) modifies to

$$x^+ = A^\star x + B^\star(\overset{\text{m}}{u} - d^u),\ \overset{\text{m}}{y} = C^\star x + d^y$$

$$\begin{bmatrix} x_a^+ \\ \chi^+ \end{bmatrix} = \begin{bmatrix} A_a & B_a\mathbf{K} \\ \mathbf{L}_{\ell^\star} C_a & \mathbf{F}_{\ell^\star}+\mathbf{B}_{\ell^\star}\mathbf{K} \end{bmatrix} \begin{bmatrix} x_a \\ \chi \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{L}_{\ell^\star} \end{bmatrix} \overset{\text{m}}{y},\ \overset{\text{m}}{u} = \mathbf{K}\chi .$$

We select $x(0) = (-0.20, -0.60, -1.71)$, $x_a(0) = 0$, $\chi(0) = 0$, the input noise as $d_1^u, d_2^u \sim U[-0.01, 0.01]$, and the output noise as $d_1^y, d_2^y \sim U[-0.01, 0.01]$. As expected, by bounded-input-bounded-output stability properties (Antsaklis & Michel, 2005, §6.10, Thm. 10.17), the outputs remain in a neighborhood of the origin as shown in Fig. 6.

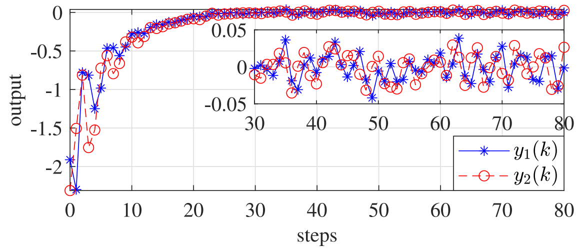

Fig. 6. Controller execution with measurement noise.

Table 1
Comparison with respect to n.

n	4	5	6	7
Feasibility (33)	✓	✓	✓	✗
Comp. time (s)	0.45	0.51	0.63	0.74
Feasibility (Miller et al., 2023b, (28))	✓	✗	✗	✗
Comp. time (s)	126	531	1571	3616

5.3. Comparison

To the best of our knowledge, Miller et al. (2023b) is the only other existing work that designs a controller directly from input–output data with measurement noise. Miller et al. (2023b) treats single-input single-output systems, and the controller is designed for superstability, which is a stronger property than asymptotic stability. To evaluate feasibility and scalability, we consider a set of systems with $n = 4, \ldots, 7$. Set $A_i := \begin{bmatrix} \sin(0.1i\pi) & \cos(0.1i\pi) \\ -\cos(0.1i\pi) & \sin(0.1i\pi) \end{bmatrix}$ for $i = 1, 2, 3$, $A^\star = \operatorname{blkdiag}(A_1, A_2)$ for $n = 4$, $A^\star = \operatorname{blkdiag}(A_1, A_2, 1)$ for $n = 5$, $A^\star = \operatorname{blkdiag}(A_1, A_2, A_3)$ for $n = 6$, $A^\star = \operatorname{blkdiag}(A_1, A_2, A_3, 1)$ for $n = 7$, $B^\star = \mathbb{1}_n$, and $C^\star = \mathbb{1}_n^\top$, where $\operatorname{blkdiag}(\cdot)$ is the block diagonal matrix obtained from its matrix arguments and $\mathbb{1}_n$ is the column vector of n ones. The eigenvalues of the system are all located on the unit circle. For each system, we apply $\overset{\text{m}}{u} \sim U[-1, 1]$, $d^u \sim U[-0.0001, 0.0001]$ and $d^y \sim U[-0.0001, 0.0001]$ to collect data. The data length is set to $T = 4n$. We solve (33) and Miller et al. (2023b, Eq. (28)) using the same data set, verify whether each approach returns a stabilizing or superstabilizing controller, and record the computational time (for a desktop with Intel Core™ i7-9700 CPU 3.00 GHz). The comparison results are in Table 1. While the approach in Miller et al. (2023b) can handle higher levels of noise in lower-dimensional systems, the proposed method demonstrates superior scalability, making it particularly effective in higher-dimensional settings.

6. Conclusion

In this paper, we have synthesized a dynamic controller directly from noisy-input noisy-output data. To deal with input–output data, an auxiliary state–space representation is utilized. We distinguish between two cases: $p\ell^\star = n$ and $p\ell^\star \neq n$. In the former case, we capitalize on the structure of the auxiliary representation, and design a controller that stabilizes all systems consistent with data via a common Lyapunov function. In the latter case, we introduce an augmented system comprising the original system. The controller for the augmented system is designed using the same procedure as in the first case and, subsequently, the controller for the original system is derived from the controller for the augmented system. The results are validated on numerical examples. This work specifically focuses on the stabilization problem because it serves as a prototypical control problem. Future research includes the extension of the approach to address other control problems such as $\mathcal{H}_2$, $\mathcal{H}_\infty$ and tracking control using input–output data. We also anticipate potential applications and extensions of the proposed methodology to nonlinear systems.

Appendix A. Ancillary results for linear systems (1), (3)

In this section we present some straightforward results for the linear system in (1) and the auxiliary linear system in (3). Relevant relations between input, state, output of solutions to (1) are characterized next.

Claim 1. *For each $\hat{x} \in \mathbb{R}^n$ and sequence $\{u(k)\}_{k=0}^{\infty}$, let $x(\cdot)$ be the solution to (1) with initial condition $x(0) = \hat{x}$ and input $\{u(k)\}_{k=0}^{\infty}$, and $y(\cdot) = Cx(\cdot)$ be the corresponding output response. If (A, C) is observable, then, for each $k \geq \ell$,*

$$\begin{bmatrix} y(k-\ell) \\ \vdots \\ y(k-1) \end{bmatrix} = \mathcal{O}_\ell x(k-\ell) + \mathcal{T}_\ell \begin{bmatrix} u(k-\ell) \\ \vdots \\ u(k-1) \end{bmatrix} \tag{A.1a}$$

$$x(k) = A^\ell \mathcal{O}_\ell^{\mathrm{L}} \begin{bmatrix} y(k-\ell) \\ \vdots \\ y(k-1) \end{bmatrix} + \left(\mathcal{R}_\ell - A^\ell \mathcal{O}_\ell^{\mathrm{L}} \mathcal{T}_\ell\right) \begin{bmatrix} u(k-\ell) \\ \vdots \\ u(k-1) \end{bmatrix} \tag{A.1b}$$

$$y(k) = CA^\ell \mathcal{O}_\ell^{\mathrm{L}} \begin{bmatrix} y(k-\ell) \\ \vdots \\ y(k-1) \end{bmatrix} + \left(C\mathcal{R}_\ell - CA^\ell \mathcal{O}_\ell^{\mathrm{L}} \mathcal{T}_\ell\right) \begin{bmatrix} u(k-\ell) \\ \vdots \\ u(k-1) \end{bmatrix}. \tag{A.1c}$$

Proof. See Li et al. (2024, Proof of Claim 1). □

In Lemma 1, we need to determine the reachable subspace of $(\mathbf{A}_\ell, \mathbf{B}_\ell)$ in (3). Hence, we provide an analytic expression for the forced response of (3) next.

Claim 2. *Let $\xi(\cdot)$ be the solution to (3) with initial condition equal to 0 and input sequence $\{v(k)\}_{k=0}^{\infty}$, and introduce the fictitious $v(-\ell-1) = v(-\ell) = \cdots = v(-1) = 0$. The solution $\xi(\cdot)$ satisfies, for each $k \geq 0$,*

$$\xi(k) = \left[\begin{array}{c} \xi_1(k) \\ \xi_2(k) \\ \vdots \\ \xi_{\ell-1}(k) \\ \xi_\ell(k) \\ \hline \xi_{\ell+1}(k) \\ \xi_{\ell+2}(k) \\ \vdots \\ \xi_{\ell+\ell-1}(k) \\ \xi_{\ell+\ell}(k) \end{array}\right] = \left[\begin{array}{c} \sum_{j=0}^{k-1-\ell} CA^{k-1-\ell-j}Bv(j) \\ \sum_{j=0}^{k-\ell} CA^{k-\ell-j}Bv(j) \\ \vdots \\ \sum_{j=0}^{k-3} CA^{k-3-j}Bv(j) \\ \sum_{j=0}^{k-2} CA^{k-2-j}Bv(j) \\ \hline v(k-\ell) \\ v(k-\ell+1) \\ \vdots \\ v(k-2) \\ v(k-1) \end{array}\right] \tag{A.2}$$

or, equivalently,

$$\xi(k) = \left[\begin{array}{c} \mathcal{O}_\ell \sum_{j=0}^{k-1-\ell} A^{k-1-\ell-j}Bv(j) + \mathcal{T}_\ell \begin{bmatrix} v(k-\ell) \\ \vdots \\ v(k-1) \end{bmatrix} \\ \hline v(k-\ell) \\ \vdots \\ v(k-1) \end{array}\right]. \tag{A.3}$$

Proof. See Li et al. (2024, Proof of Claim 2). □

An immediate consequence of Claim 2 is the next result.

Claim 3. *For (A, C) observable, let $x(\cdot)$ be the solution to (1) with initial condition $x(0) = 0$ and input $\{u(k)\}_{k=0}^{\infty}$, and $y(\cdot) = Cx(\cdot)$ be the corresponding output response. Then, the solution $\xi(\cdot)$ to (3) with initial condition $\xi(0) = 0$ and input $\{v(k)\}_{k=0}^{\infty} = \{u(k)\}_{k=0}^{\infty}$ satisfies*

$$\xi(k) = \left[\begin{array}{c} y(k-\ell) \\ \vdots \\ y(k-1) \\ \hline u(k-\ell) \\ \vdots \\ u(k-1) \end{array}\right] \quad \forall k \geq \ell. \tag{A.4}$$

Proof. See Li et al. (2024, Proof of Claim 3). □

Appendix B. Proof of Lemma 4

By (A.1c) in Claim 1, the quantities of the experiment satisfy, for $k = \ell^\star, \ldots, T$,

$$y^{\mathrm{m}}(k) - d^y(k) = y(k) = Z_{1\ell^\star} \begin{bmatrix} y(k-\ell^\star) \\ \vdots \\ y(k-1) \end{bmatrix} + Z_{2\ell^\star} \begin{bmatrix} u(k-\ell^\star) \\ \vdots \\ u(k-1) \end{bmatrix}$$
$$= Z_{1\ell^\star} \begin{bmatrix} y^{\mathrm{m}}(k-\ell^\star) - d^y(k-\ell^\star) \\ \vdots \\ y^{\mathrm{m}}(k-1) - d^y(k-1) \end{bmatrix} + Z_{2\ell^\star} \begin{bmatrix} u^{\mathrm{m}}(k-\ell) - d^u(k-\ell^\star) \\ \vdots \\ u^{\mathrm{m}}(k-1) - d^u(k-1) \end{bmatrix}$$

and this yields (15). (15) yields, for $k = \ell^\star, \ldots, T$,

$$\left[\begin{array}{c} y^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ y^{\mathrm{m}}(k-1) \\ y^{\mathrm{m}}(k) \\ \hline u^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ u^{\mathrm{m}}(k-1) \\ u^{\mathrm{m}}(k) \end{array}\right] = \left[\begin{array}{c} y^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ y^{\mathrm{m}}(k-1) \\ \left\{\begin{array}{l} Z_{1\ell^\star} \begin{bmatrix} y^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ y^{\mathrm{m}}(k-1) \end{bmatrix} + Z_{2\ell^\star} \begin{bmatrix} u^{\mathrm{m}}(k-\ell^\star) \\ \vdots \\ u^{\mathrm{m}}(k-1) \end{bmatrix} \\ - Z_{1\ell^\star} \begin{bmatrix} d^y(k-\ell^\star) \\ \vdots \\ d^y(k-1) \end{bmatrix} - Z_{2\ell^\star} \begin{bmatrix} d^u(k-\ell^\star) \\ \vdots \\ d^u(k-1) \end{bmatrix} + d^y(k) \end{array}\right\} \\ \hline u^{\mathrm{m}}(k-\ell^\star+1) \\ \vdots \\ u^{\mathrm{m}}(k-1) \\ u^{\mathrm{m}}(k) \end{array}\right],$$

and this yields (16) for $\mathbf{F}_{\ell^\star}$, $\mathbf{L}_{\ell^\star}$ and $\mathbf{B}_{\ell^\star}$ as in (4a). Finally, by the definitions in (13), (14), (11), stack (16) column after column from $k = \ell^\star$ to $k = T$ to get (17).

Appendix C. Proof of Lemma 5

If $\mathscr{A} \succ 0$, then $\mathscr{Z}$ and $\mathscr{Q}$ in (30) are well-defined, and algebraic computations yield (31) from (29a). (17) in Lemma 4 and $\Delta_{10} \in \mathcal{D}$ imply $Z_{\ell^\star} \in \mathcal{C}$ and, by $Z_{\ell^\star} \in \mathcal{C}$, (31) yields that $\mathscr{Q} \succeq (Z_{\ell^\star} - \mathscr{Z})\mathscr{A}(Z_{\ell^\star} - \mathscr{Z})^\top \succeq 0$ by $\mathscr{A} \succ 0$. $\mathscr{A} \succ 0$ and $\mathscr{Q} \succeq 0$ allow applying (Bisoffi et al., 2022, Proposition 1) and obtaining (32) from (31); they also allow applying (Bisoffi et al., 2022, Lemma 2) and showing with analogous arguments that the nonempty set $\mathcal{C}$ is bounded.

Appendix D. Proof of Lemma 6

By $P \succ 0$ and Schur complement, (27b) is equivalent to

$$\begin{bmatrix} -P & (\mathbf{F}_{\ell^\star} + \mathbf{L}_{\ell^\star} Z + \mathbf{B}_{\ell^\star} \mathbf{K})P \\ \star & -P \end{bmatrix} \prec 0 \quad \forall Z \in \mathcal{C}. \tag{D.1}$$

There exist $\mathbf{K}$ and $P = P^\top \succ 0$ satisfying (D.1) if and only if there exist $\mathbf{Y}$ and $P = P^\top \succ 0$ satisfying

$$\begin{bmatrix} -P & \mathbf{F}_{\ell^\star} P + \mathbf{L}_{\ell^\star} ZP + \mathbf{B}_{\ell^\star} \mathbf{Y} \\ \star & -P \end{bmatrix} \prec 0 \quad \forall Z \in \mathcal{C}$$

with $\mathbf{Y} = \mathbf{K}P$. By $\mathcal{C}$ in (32), which is obtained under Assumption 2 in Lemma 5, this holds if and only if

$$0 \succ \begin{bmatrix} -P & \mathbf{F}_{\ell^\star} P + \mathbf{L}_{\ell^\star} \mathscr{Z} P + \mathbf{B}_{\ell^\star} \mathbf{Y} \\ \star & -P \end{bmatrix} + \begin{bmatrix} \mathbf{L}_{\ell^\star} \mathscr{Q}^{1/2} \\ 0 \end{bmatrix} \Upsilon \begin{bmatrix} 0 \\ P\mathscr{A}^{-1/2} \end{bmatrix}^\top$$
$$+ \begin{bmatrix} 0 \\ P\mathscr{A}^{-1/2} \end{bmatrix} \Upsilon^\top \begin{bmatrix} \mathbf{L}_{\ell^\star} \mathscr{Q}^{1/2} \\ 0 \end{bmatrix}^\top \quad \forall \Upsilon \text{ with } \|\Upsilon\| \leq 1.$$

By Petersen's lemma (Petersen, 1987) as reported in Bisoffi et al. (2022, Fact 1), this holds if and only if there exists $\lambda > 0$ such that

$$0 \succ \begin{bmatrix} -P & \mathbf{F}_{\ell^\star} P + \mathbf{L}_{\ell^\star} \mathscr{Z} P + \mathbf{B}_{\ell^\star} \mathbf{Y} \\ \star & -P \end{bmatrix}$$
$$+ \frac{1}{\lambda} \begin{bmatrix} \mathbf{L}_{\ell^\star} \mathscr{Q}^{1/2} \\ 0 \end{bmatrix} \begin{bmatrix} \mathbf{L}_{\ell^\star} \mathscr{Q}^{1/2} \\ 0 \end{bmatrix}^\top + \lambda \begin{bmatrix} 0 \\ P\mathscr{A}^{-1/2} \end{bmatrix} \begin{bmatrix} 0 \\ P\mathscr{A}^{-1/2} \end{bmatrix}^\top$$

$$=\begin{bmatrix} -P+\frac{1}{\lambda}\mathbf{L}_{\ell^\star}\mathscr{Q}\mathbf{L}_{\ell^\star}^\top & \mathbf{F}_{\ell^\star}P+\mathbf{L}_{\ell^\star}\mathscr{Z}P+\mathbf{B}_{\ell^\star}\mathbf{Y} \\ \star & -P+\lambda P\mathscr{A}^{-1}P \end{bmatrix}.$$

The existence of $\mathbf{Y}$, $P=P^\top\succ 0$, $\lambda>0$ such that this holds is equivalent to the existence of $\mathbf{Y}$, $P=P^\top\succ 0$ such that

$$0\succ\begin{bmatrix} -P+\mathbf{L}_{\ell^\star}\mathscr{Q}\mathbf{L}_{\ell^\star}^\top & \mathbf{F}_{\ell^\star}P+\mathbf{L}_{\ell^\star}\mathscr{Z}P+\mathbf{B}_{\ell^\star}\mathbf{Y} \\ \star & -P+P\mathscr{A}^{-1}P \end{bmatrix} \tag{D.2}$$

as can be shown by multiplying or dividing by $\lambda>0$. By $\mathscr{Q}$ and $\mathscr{Z}$ in (30), this is equivalent to

$$0\succ\begin{bmatrix} -P-\mathbf{L}_{\ell^\star}\mathscr{C}\mathbf{L}_{\ell^\star}^\top & \mathbf{F}_{\ell^\star}P+\mathbf{B}_{\ell^\star}\mathbf{Y} \\ \star & -P \end{bmatrix}+\begin{bmatrix}\mathbf{L}_{\ell^\star}\mathscr{B} \\ -P\end{bmatrix}\mathscr{A}^{-1}\begin{bmatrix}\mathbf{L}_{\ell^\star}\mathscr{B} \\ -P\end{bmatrix}^\top$$

and, by $\mathscr{A}\succ 0$ from Assumption 2 and Schur complement, this is equivalent to (33b).

Appendix E. Proof of Lemma 7

For (A, C) observable, we show that for each $\hat{x}$, $\hat{\chi}$ and sequence $\{u(k)\}_{k=0}^{\infty}$, (i) the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$ to (35) with initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0)\end{smallmatrix}\right]=\left[\begin{smallmatrix} \hat{x} \\ \hat{\chi}\end{smallmatrix}\right]$ and with input $\{u(k)\}_{k=0}^{\infty}$ and the corresponding output response $y(\cdot)=Cx(\cdot)$ satisfy, for all $k\geq\ell$,

$$\chi(k)=\big(y(k-\ell),\dots,y(k-1),u(k-\ell),\dots,u(k-1)\big);$$

(ii) there exists $\hat{\xi}$ such that such solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$, the corresponding output response $y(\cdot)=Cx(\cdot)$, and the solution $\xi(\cdot)$ to (3) with initial condition $\xi(\ell)=\hat{\xi}$ and input $\{v(k)\}_{k=\ell}^{\infty}=\{u(k)\}_{k=\ell}^{\infty}$ satisfy, for all $k\geq\ell$

$$\xi(k)=\big(y(k-\ell),\dots,y(k-1),u(k-\ell),\dots,u(k-1)\big).$$

As for (ii), (35) has a "triangular" structure so that the component $x(\cdot)$ of the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$ to (35) is also a solution to (1). Hence, Lemma 3 applies and (ii) holds. As for (i), we use mathematical induction. As for the base case we need to show that $\chi(\ell)=\big(y(0),\dots,y(\ell-1),u(0),\dots,u(\ell-1)\big)$. Indeed, by (35b),

$$\chi(1)=\left[\begin{array}{c}\hat{\chi}_2\\ \vdots\\ \hat{\chi}_\ell\\ y(0)\\ \hline \hat{\chi}_{\ell+2}\\ \vdots\\ \hat{\chi}_{\ell+\ell}\\ u(0)\end{array}\right],\dots,\ \chi(\ell-1)=\left[\begin{array}{c}\hat{\chi}_\ell\\ y(0)\\ \vdots\\ y(\ell-2)\\ \hline \hat{\chi}_{\ell+\ell}\\ u(0)\\ \vdots\\ u(\ell-2)\end{array}\right],\ \chi(\ell)=\left[\begin{array}{c}y(0)\\ \vdots\\ y(\ell-1)\\ \hline u(0)\\ \vdots\\ u(\ell-1)\end{array}\right].$$

As for the induction step, we suppose that, for $k\geq\ell$,

$$\chi(k)=\left[\begin{array}{c}y(k-\ell)\\ \vdots\\ y(k-1)\\ \hline u(k-\ell)\\ \vdots\\ u(k-1)\end{array}\right] \tag{E.1}$$

and we need to show that $\chi(k+1)=\big(y(k-\ell+1),\dots,y(k),u(k-\ell+1),\dots,u(k)\big)$. We have

$$\chi(k+1)\overset{(35b)}{=}\mathbf{F}_\ell\chi(k)+\mathbf{L}_\ell y(k)+\mathbf{B}_\ell u(k)$$
$$=\left[\begin{array}{c}\chi_2(k)\\ \vdots\\ \chi_\ell(k)\\ y(k)\\ \hline \chi_{\ell+2}(k)\\ \vdots\\ \chi_{\ell+\ell}(k)\\ u(k)\end{array}\right]\overset{(E.1)}{=}\left[\begin{array}{c}y(k-\ell+1)\\ \vdots\\ y(k-1)\\ y(k)\\ \hline u(k-\ell+1)\\ \vdots\\ u(k-1)\\ u(k)\end{array}\right]$$

as we needed to show.

Appendix F. Proof of Lemma 8

Since $\mathbf{K}$ makes the matrix $\mathbf{A}_\ell+\mathbf{B}_\ell\mathbf{K}$ Schur, $\xi=0$ is globally asymptotically stable for

$$\xi^+=\mathbf{A}_\ell\xi+\mathbf{B}_\ell v \tag{F.1a}$$

$$v=\mathbf{K}\xi. \tag{F.1b}$$

For each $\hat{x}$, $\hat{\chi}$, the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$ to the closed-loop system (36) with initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0)\end{smallmatrix}\right]=\left[\begin{smallmatrix} \hat{x} \\ \hat{\chi}\end{smallmatrix}\right]$ has $\{u(k)\}_{k=0}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=0}^{\infty}$. From Lemma 7, for each $\hat{x}$, $\hat{\chi}$ and the sequence $\{u(k)\}_{k=0}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=0}^{\infty}$, there exists $\hat{\xi}$ such that: the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$ to (36a)–(36b) with initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0)\end{smallmatrix}\right]=\left[\begin{smallmatrix} \hat{x} \\ \hat{\chi}\end{smallmatrix}\right]$ and with input $\{u(k)\}_{k=0}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=0}^{\infty}$, the corresponding output response $y(\cdot)=Cx(\cdot)$ and the solution $\xi(\cdot)$ to (F.1a) with initial condition $\xi(\ell)=\hat{\xi}$ and input $\{v(k)\}_{k=\ell}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=\ell}^{\infty}$ satisfy

$$\xi(k)=\left[\begin{array}{c}y(k-\ell)\\ \vdots\\ y(k-1)\\ \hline u(k-\ell)\\ \vdots\\ u(k-1)\end{array}\right]=\chi(k)\quad\forall k\geq\ell. \tag{F.2}$$

In other words, we have considered an arbitrary initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0)\end{smallmatrix}\right]=\left[\begin{smallmatrix} \hat{x} \\ \hat{\chi}\end{smallmatrix}\right]$ and the input $\{u(k)\}_{k=0}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=0}^{\infty}$ for the open-loop system (36a)–(36b) and obtained a solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$; Lemma 7 allows us to associate, with this solution, a solution $\xi(\cdot)$ to the open-loop system (F.1a) with some initial condition $\xi(\ell)=\hat{\xi}$ and input $\{v(k)\}_{k=\ell}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=\ell}^{\infty}$ such that $\xi(k)=\chi(k)$ for all $k\geq\ell$. As a consequence, $\{v(k)\}_{k=\ell}^{\infty}=\{\mathbf{K}\chi(k)\}_{k=\ell}^{\infty}=\{\mathbf{K}\xi(k)\}_{k=\ell}^{\infty}$. Hence, for some initial condition $\xi(\ell)=\hat{\xi}$, $\xi(\cdot)$ is also a solution to the closed-loop system (F.1). Thanks to the assumption of $\mathbf{A}_\ell+\mathbf{B}_\ell\mathbf{K}$ being Schur, $\xi(\cdot)$ converges asymptotically to 0. By (F.2), if $k\mapsto\xi(k)$ converges asymptotically to 0, so do $k\mapsto\left[\begin{smallmatrix} y(k-\ell) \\ \vdots \\ y(k-1)\end{smallmatrix}\right]$, $k\mapsto\left[\begin{smallmatrix} u(k-\ell) \\ \vdots \\ u(k-1)\end{smallmatrix}\right]$ and $k\mapsto\chi(k)$. By (A, C) observable and (A.1b) in Claim 1, for each $k\geq\ell$,

$$x(k)=A^\ell\mathcal{O}_\ell^{\mathrm{L}}\begin{bmatrix}y(k-\ell)\\ \vdots\\ y(k-1)\end{bmatrix}+(\mathcal{R}_\ell-A^\ell\mathcal{O}_\ell^{\mathrm{L}}\mathcal{T}_\ell)\begin{bmatrix}u(k-\ell)\\ \vdots\\ u(k-1)\end{bmatrix}$$

and so also $k\mapsto x(k)$ converges asymptotically to 0. In summary, we have shown that for each $\hat{x}$, $\hat{\chi}$ the solution $\left[\begin{smallmatrix} x(\cdot) \\ \chi(\cdot)\end{smallmatrix}\right]$ to (36) with initial condition $\left[\begin{smallmatrix} x(0) \\ \chi(0)\end{smallmatrix}\right]=\left[\begin{smallmatrix} \hat{x} \\ \hat{\chi}\end{smallmatrix}\right]$ converges asymptotically to 0. For linear time-invariant systems, attractivity implies global attractivity and Lewis (2017, Prop. 4) stability. Then, we can conclude that $(x,\chi)=0$ is globally asymptotically stable for (36).

Appendix G. Proof of Lemma 9

Let $\{u(k)\}_{k=0}^{T-1}$ be the data-collection input sequence applied to (8), which generates the state and output sequences $\{x(k)\}_{k=0}^{T-\ell^\star}$ and $\{y(k)\}_{k=0}^{T-1}$. Let

$$W_0:=\left[\begin{array}{cccc} x(0) & x(1) & \dots & x(T-\ell^\star)\\ \hline u(0) & u(1) & \dots & u(T-\ell^\star)\\ \vdots & \vdots & & \vdots\\ u(\ell^\star-1) & u(\ell^\star) & \dots & u(T-1)\end{array}\right].$$

By the definition of solution to the linear system (8) and the definitions of $\mathcal{O}_{\ell^\star}$ and $\mathcal{T}_{\ell^\star}$, data satisfy

$$\begin{bmatrix}y(0)\\ \vdots\\ y(\ell^\star-1)\end{bmatrix}=\mathcal{O}_{\ell^\star}x(0)+\mathcal{T}_{\ell^\star}\begin{bmatrix}u(0)\\ \vdots\\ u(\ell^\star-1)\end{bmatrix},\dots,$$

$$\begin{bmatrix}y(T-\ell^\star)\\ \vdots\\ y(T-1)\end{bmatrix}=\mathcal{O}_{\ell^\star}x(T-\ell^\star)+\mathcal{T}_{\ell^\star}\begin{bmatrix}u(T-\ell^\star)\\ \vdots\\ u(T-1)\end{bmatrix},$$

$$\begin{bmatrix}u(0)\\ \vdots\\ u(\ell^\star-1)\end{bmatrix}=\begin{bmatrix}u(0)\\ \vdots\\ u(\ell^\star-1)\end{bmatrix},\dots,\begin{bmatrix}u(T-\ell^\star)\\ \vdots\\ u(T-1)\end{bmatrix}=\begin{bmatrix}u(T-\ell^\star)\\ \vdots\\ u(T-1)\end{bmatrix}.$$

By (38b), these relations can be assembled as

$$S_0=\begin{bmatrix}\mathcal{O}_{\ell^\star} & \mathcal{T}_{\ell^\star}\\ 0 & I_{m\ell^\star}\end{bmatrix}W_0\in\mathbb{R}^{(p\ell^\star+m\ell^\star)\times(T-\ell^\star+1)}. \tag{G.1}$$

Based on (G.1), the proof is carried out in three steps.

Step 1: $\Psi_0\Psi_0^\top \succ \Theta_{22} \implies \operatorname{rank} S_0 = (p+m)\ell^\star$. Suppose by contradiction that $\Psi_0\Psi_0^\top \succ \Theta_{22}$ but S_0 has not full row rank. Then, there exists a nonzero vector $v \in \mathbb{R}^{p\ell^\star + m\ell^\star}$ such that $v^\top S_0 = 0$. We have

$$v^\top \Psi_0\Psi_0^\top v \overset{(39)}{=} v^\top(S_0+N_0)(S_0+N_0)^\top v$$

$$= v^\top N_0 N_0^\top v > v^\top \Theta_{22} v \tag{G.2}$$

by $\Psi_0\Psi_0^\top \succ \Theta_{22}$. On the other hand, the definitions of Δ_{10} in (11d) and N_0 in (38a), and $\Delta_{10} \in \mathcal{D}$ yield

$$\Delta_{10}\Delta_{10}^\top = \begin{bmatrix} d^y(\ell^\star)\cdots d^y(T) \\ \hline N_0 \end{bmatrix} \begin{bmatrix} d^y(\ell^\star)\cdots d^y(T) \\ \hline N_0 \end{bmatrix}^\top \preceq \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{12}^\top & \Theta_{22} \end{bmatrix},$$

which implies $N_0N_0^\top \preceq \Theta_{22}$. This is contradicted by (G.2), thus S_0 has full row rank.

Step 2: $\operatorname{rank} S_0 = (p+m)\ell^\star \implies \operatorname{rank} \mathcal{O}_{\ell^\star} = p\ell^\star$. From (G.1), we have that $\begin{bmatrix} \mathcal{O}_{\ell^\star} & \mathcal{T}_{\ell^\star} \\ 0 & I_{m\ell^\star} \end{bmatrix}$ must have full row rank, i.e., its rank is $(p+m)\ell^\star$. We have

$$(p+m)\ell^\star = \operatorname{rank} \begin{bmatrix} \mathcal{O}_{\ell^\star} & \mathcal{T}_{\ell^\star} \\ 0 & I_{m\ell^\star} \end{bmatrix}$$

$$= \operatorname{rank}\left(\begin{bmatrix} I_{p\ell^\star} & \mathcal{T}_{\ell^\star} \\ 0 & I_{m\ell^\star} \end{bmatrix} \begin{bmatrix} \mathcal{O}_{\ell^\star} & 0_{p\ell^\star \times m\ell^\star} \\ 0_{m\ell^\star \times n} & I_{m\ell^\star} \end{bmatrix} \right)$$

$$= \operatorname{rank} \mathcal{O}_{\ell^\star} + m\ell^\star.$$

Hence, $p\ell^\star = \operatorname{rank} \mathcal{O}_{\ell^\star}$.

Step 3: $\operatorname{rank} \mathcal{O}_{\ell^\star} = p\ell^\star \implies p\ell^\star = n$. By observability of $(A^\star, C^\star)$, $\operatorname{rank} \mathcal{O}_{\ell^\star} = n$ so $p\ell^\star = n$.

Appendix H. Proof of Lemma 10

From the hypothesis of the lemma and $\Theta_2 = \Theta_2^\top \succeq 0$, $\lambda_{\min}(S_0S_0^\top) > 0$. Thus, the rows of S_0 are fewer or as many as the columns of S_0. The same holds for Ψ_0, see (39), and thus $\lambda_{\min}(\Psi_0\Psi_0^\top) = \sigma_{\min}(\Psi_0)^2$ (Horn & Johnson, 2013, Theorem 2.6.3). Analogously, $\lambda_{\min}(S_0S_0^\top) = \sigma_{\min}(S_0)^2$. Since $S_0 = \Psi_0 + (-N_0)$ by (39), Horn and Johnson (2013, (7.3.13)) yields

$$\sigma_{\min}(\Psi_0) \geq \sigma_{\min}(S_0) - \sigma_{\max}(N_0)$$

$$= \sqrt{\lambda_{\min}(S_0S_0^\top)} - \sigma_{\max}(N_0)$$

$$> 2\sqrt{\lambda_{\max}(\Theta_{22})} - \sigma_{\max}(N_0) \tag{H.1}$$

where the last inequality follows from the hypothesis of the lemma. Since $\Delta_{10} \overset{(11d)}{=} \left[\frac{d^y(\ell^\star)\cdots d^y(T)}{N_0}\right] \in \mathcal{D}$, we have $N_0N_0^\top \preceq \Theta_{22}$; by $\Theta_{22} = \Theta_{22}^\top$ and $N_0N_0^\top \preceq \Theta_{22}$, Horn and Johnson (2013, Cor. 7.7.4(c)) yields $\lambda_{\max}(N_0N_0^\top) \leq \lambda_{\max}(\Theta_{22})$ or, equivalently, $\sigma_{\max}(N_0) \leq \sqrt{\lambda_{\max}(\Theta_{22})}$. By using this condition in (H.1), we obtain $\sigma_{\min}(\Psi_0) > \sqrt{\lambda_{\max}(\Theta_{22})}$ and $\lambda_{\min}(\Psi_0\Psi_0^\top) > \lambda_{\max}(\Theta_{22})$. Hence,

$$\Psi_0\Psi_0^\top \succeq \lambda_{\min}(\Psi_0\Psi_0^\top)I \succ \lambda_{\max}(\Theta_{22})I \succeq \Theta_{22}$$

and thus $\Psi_0\Psi_0^\top \succ \Theta_{22}$, i.e., Assumption 2 holds.

Appendix I. Proof of Lemma 11

We note that in the proof of Lemma 9

$$\Psi_0\Psi_0^\top \succ \Theta_{22} \implies \operatorname{rank} S_0 = (p+m)\ell^\star \implies \operatorname{rank} \mathcal{O}_{\ell^\star} = p\ell^\star$$

hold independently of the observability of $(A^\star, C^\star)$ and the fact that $\ell^\star$ is the observability index of $(A^\star, C^\star)$. Hence, from analogous arguments to those in the proof of Lemma 9, $\Psi_0^{\mathrm{aug}}\Psi_0^{\mathrm{aug}\top} \succ \Theta_{22}^{\mathrm{aug}}$ implies that $\operatorname{rank} \mathcal{O}_{\ell^\star}^{\mathrm{aug}} = p\ell^\star$. This equality implies that $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ in (41) is nonsingular, since $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ is square by construction. As $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ has full rank, the pair $(A_{\mathrm{aug}}, C_{\mathrm{aug}})$ is observable. Again since $\mathcal{O}_{\ell^\star}^{\mathrm{aug}}$ is square and nonsingular, the observability index of $(A_{\mathrm{aug}}, C_{\mathrm{aug}})$ is, by its definition, equal to $\ell^\star$.

References

Al Makdah, A. A., & Pasqualetti, F. (2024). Model-based and data-based output feedback for external positivity. In *IEEE conference on decision and control* (pp. 4119–4124).

Alsalti, M., Lopez, V. G., & Müller, M. A. (2025). Notes on data-driven output-feedback control of linear MIMO systems. *IEEE Transactions on Automatic Control*, *70*(9), 6143–6150.

Antsaklis, P. J., & Michel, A. N. (2005). *Linear systems*. Springer Science & Business Media.

Barmish, B. R. (1985). Necessary and sufficient conditions for quadratic stabilizability of an uncertain system. *Journal of Optimization Theory and Applications*, *46*, 399–408.

Berberich, J., Scherer, C. W., & Allgöwer, F. (2023). Combining prior knowledge and data for robust controller design. *IEEE Transactions on Automatic Control*, *68*(8), 4618–4633.

Bisoffi, A., De Persis, C., & Tesi, P. (2022). Data-driven control via Petersen's lemma. *Automatica*, *145*, Article 110537.

Bisoffi, A., Li, L., De Persis, C., & Monshizadeh, N. (2024). Controller synthesis for input-state data with measurement errors. *IEEE Control Systems Letters*, *8*, 1571–1576.

Campi, M. C., Lecchini, A., & Savaresi, S. M. (2002). Virtual reference feedback tuning: a direct method for the design of feedback controllers. *Automatica*, *38*(8), 1337–1346.

Coulson, J., Lygeros, J., & Dörfler, F. (2019). Data-enabled predictive control: In the shallows of the DeePC. In *European control conference* (pp. 307–312).

Dai, X., De Persis, C., & Monshizadeh, N. (2023). Data-driven optimal output feedback control of linear systems from input-output data. *IFAC-PapersOnLine*, *56*(2), 1376–1381.

D'Amico, W., Bisoffi, A., & Farina, M. (2023). Data-based control design for output-error linear discrete-time systems with probabilistic stability guarantees. *IEEE Control Systems Letters*, *7*, 2035–2040.

de Jong, T. O., Breschi, V., Schoukens, M., & Formentin, S. (2023). Data-driven model-reference control with closed-loop stability: The output-feedback case. *IEEE Control Systems Letters*, *7*, 2431–2436.

De Persis, C., & Tesi, P. (2020). Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, *65*(3), 909–924.

Douglas, R. G. (1966). On majorization, factorization, and range inclusion of operators on Hilbert space. *Proceedings of the American Mathematical Society*, *17*(2), 413–415.

Goodwin, G. C., & Sin, K. S. (2009). *Adaptive filtering prediction and control*. Courier Corporation.

Grant, M., & Boyd, S. (2008). Graph implementations for nonsmooth convex programs. In *Recent advances in learning and control* (pp. 95–110). Springer.

Grant, M., & Boyd, S. (2014). *CVX: Matlab software for disciplined convex programming, v. 2.1*. http://cvxr.com/cvx.

Green, M., & Limebeer, D. J. (2012). *Linear robust control*. Courier Corporation.

Horn, R. A., & Johnson, C. R. (2013). *Matrix analysis*. Cambridge University Press.

Jo, N. H., & Shim, H. (2022). Data-driven output-feedback predictive control: Unknown plant's order and measurement noise. arXiv preprint arXiv:2201.03136.

Kailath, T. (1980). *Linear systems*. Prentice-Hall.

Koch, A., Berberich, J., & Allgöwer, F. (2022). Provably robust verification of dissipativity properties from data. *IEEE Transactions on Automatic Control*, *67*(8), 4248–4255.

Lewis, A. D. (2017). Remarks on stability of time-varying linear systems. *IEEE Transactions on Automatic Control*, *62*(11), 6039–6043.

Li, L., Bisoffi, A., De Persis, C., & Monshizadeh, N. (2024). Controller synthesis from noisy-input noisy-output data. arXiv preprint arXiv:2402.02588.

Miller, J., Dai, T., & Sznaier, M. (2023a). Data-driven stabilizing and robust control of discrete-time linear systems with error in variables. arXiv preprint arXiv:2210.13430.

Miller, J., Dai, T., & Sznaier, M. (2023b). Superstabilizing control of discrete-time ARX models under error in variables. *IFAC-PapersOnLine*, *56*(2), 2444–2449.

Pan, G., Ou, R., & Faulwasser, T. (2023). On a stochastic fundamental lemma and its use for data-driven optimal control. *IEEE Transactions on Automatic Control*, *68*(10), 5922–5937.

Petersen, I. R. (1987). A stabilization algorithm for a class of uncertain linear systems. *Systems & Control Letters*, *8*(4), 351–357.

Sadamoto, T. (2023). On equivalence of data informativity for identification and data-driven control of partially observable systems. *IEEE Transactions on Automatic Control*, *68*(7), 4289–4296.

Söderström, T. (2018). *Errors-in-variables methods in system identification*. Springer Cham.

Steentjes, T. R. V., Lazar, M., & Van den Hof, P. M. J. (2022). On data-driven control: Informativity of noisy input-output data with cross-covariance bounds. *IEEE Control Systems Letters*, *6*, 2192–2197.

Tanaskovic, M., Fagiano, L., Novara, C., & Morari, M. (2017). Data-driven control of nonlinear systems: An on-line direct approach. *Automatica*, *75*, 1–10.

van Waarde, H. J., Camlibel, M. K., Eising, J., & Trentelman, H. L. (2023). Quadratic matrix inequalities with applications to data-based control. *SIAM Journal on Control and Optimization*, *61*(4), 2251–2281.

van Waarde, H. J., Eising, J., Camlibel, M. K., & Trentelman, H. L. (2024). A behavioral approach to data-driven control with noisy input–output data. *IEEE Transactions on Automatic Control*, *69*(2), 813–827.

Lidong Li received the B.Sc. degree in 2018 and the M.Sc. degree in 2020, both from the Harbin Institute of Technology, Harbin, China. He is currently pursuing a Ph.D. degree at the Engineering and Technology Institute, Faculty of Science and Engineering, University of Groningen, the Netherlands. His research interests include nonlinear control systems and data-driven control.

Andrea Bisoffi received the M.Sc. degree in Automatic Control Engineering from Politecnico di Milano, Italy, in 2013 and the Ph.D. degree in Mechatronics from University of Trento, Italy, in 2017. He was a post-doctoral researcher at KTH Royal Institute of Technology, Sweden, and a researcher at University of Groningen, The Netherlands. He is currently an associate professor at Politecnico di Milano. His research interests include hybrid and nonlinear control systems, with applications to mechatronic systems, and data-driven control.

Claudio De Persis is a professor with the Engineering and Technology Institute Groningen (ENTEG), University of Groningen, Netherlands, since 2011. He received the Laurea and Ph.D. degree in engineering in 1996 and 2000, both from the University of Rome “La Sapienza”, Italy.

Nima Monshizadeh is an Associate Professor at the Engineering and Technology Institute Groningen, University of Groningen, Netherlands. He received his Ph.D. with honors (cum laude) from the Bernoulli Institute at the University of Groningen. Previously, he was a Research Associate with the control group at the University of Cambridge (2017–2018). He has served as an Associate Editor for Automatica since 2020. His primary research interests lie in the optimization and control of cyber–physical systems.