Data-Based Control Design for Output-Error Linear Discrete-Time Systems With Probabilistic Stability Guarantees

William D'Amico, Andrea Bisoffi, and Marcello Farina

***Abstract*—In this letter we propose a novel method for direct data-based design of an output feedback controller for output-error processes in the single-input-single-output case. We consider a finite number of input/output data points collected from the system. Based on them, we define a set of state-space perturbed models consistent with data, where a bound on the disturbance is obtained by scenario optimization, and the structural properties of the models in this set are theoretically analyzed. This leads to a linear matrix inequality for the design of the feedback control law with probabilistic asymptotic stability guarantees. A simulated non-minimum phase system illustrates the approach.**

***Index Terms*—Data-driven control, uncertain systems, robust control.**

I. Introduction

When input/state/output data are available from a plant to be controlled dynamically, algorithms that design a feedback controller *directly* from such data are appealing because they circumvent the possibly time-consuming tasks of creating a model for the underlying plant from first principles and, then, of identifying its parameters, especially if some of these parameters affect the closed-loop performance only mildly. Such algorithms are even more appealing if, from the *finite* amount of input/state/output data, they provide guarantees on asymptotic stability of the resulting closed loop obtained by the feedback interconnection of the unknown plant and the designed controller [1]. This is the philosophy behind many recent works that have considered the so-called fundamental lemma by Willems et al. [2] as a starting point, and have ranged over different settings in terms of underlying (unknown) plant dynamics (linear [3], polynomial [4], [5], nonlinear [6]), control objective (stabilization [3], dissipativity [7], invariance [8]) and control design method (predictive [9], optimal [10] or robust control [5], [11], [12]).

In the majority of these works, the data points used for controller design take the form of noisy input/state data since this facilitates the characterization of the plant dynamics that are consistent with the collected data and, in turn, allows obtaining necessary and sufficient conditions for robust stabilization of all such dynamics. How to carry out the design directly through noisy input/output data while guaranteeing robust stabilization is less understood and has comparatively been less investigated; at the same time, it is more relevant and realistic in real-world applications. Among the works that, within the aforementioned thread, have treated input/output data, we mention [3], [12], [13], [14], [15]. Reference [3, Sec. VI] shows how an output feedback controller can be designed from noise-free input/output data of a single-input single-output (SISO) system. In [15] a data pre-processing procedure is proposed to mitigate the effect of the measurement noise in a given batch of input/output data. Reference [12, Sec. IV.B] allows for the design of a dynamic output feedback controller in the presence of noisy input/output data while discussing some challenges in the multi-input multi-output case. Both [12] and [13] consider input/output data where the corrupting disturbance acts in the difference equation in the same fashion as in ARX (AutoRegressive with eXogenous variable) processes. Alternatively, noise can corrupt the input and output signals in an additive fashion in the so-called error-in-variables (EIV) setting. This setting includes the case of output-error (OE) processes where, quite naturally, measurements are considered to be composed of signal and noise; it is known to be potentially challenging [16], and was considered for set-membership identification, e.g., [17], [18]. Direct design of a dynamic compensator for superstabilization is pursued in [14] from input/output data in an EIV setting; such a design is formulated as a polynomial optimization problem, which is effectively solved via sum-of-squares hierarchies.

In this letter we focus on direct design of a SISO output feedback controller with output measurements corrupted by

Manuscript received 7 March 2023; revised 3 May 2023; accepted 26 May 2023. Date of publication 8 June 2023; date of current version 22 June 2023. This work was supported in part by the Project Digital Twin of the Research Programme TTW Perspective which is (partly) financed by the Dutch Research Council (NWO) under Grant P18-03. Recommended by Senior Editor A. P. Aguiar. *(Corresponding author: William D'Amico.)*

William D'Amico and Marcello Farina are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: william.damico@polimi.it).

Andrea Bisoffi was with ENTEG, University of Groningen, 9747 AG Groningen, The Netherlands. He is now with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy.

Digital Object Identifier 10.1109/LCSYS.2023.3284391

measurement noise, i.e., for OE processes. Unlike superstability in [14], we consider classical asymptotic stability. Our approach consists of defining a suitable set of state-space perturbed models consistent with the data, where an energy bound for the whole disturbance sequence is obtained via scenario optimization, leading to a linear matrix inequality for the design of the feedback control law. Also, we prove that all the matrices in the set share a particular structure in agreement with the setting of OE processes. A key feature of our scheme is that asymptotic stability is guaranteed, in a probabilistic sense, despite the uncertainty induced by additive noise corrupting the output measurements.

This letter is organized as follows: in Section II the control problem is stated, whereas in Section III a disturbance uncertainty set is constructed through a novel algorithm that relies on scenario optimization. In Section IV the main tool for control design is provided. Section V shows the application of the proposed method to a simulation example.

Notation: k denotes the discrete-time index belonging to the integers $\mathbb{Z}$. The transpose of a matrix M is M^T. $0_{n,m}$ (or 0) denotes a zero matrix with n rows and m columns (or with a suitable number of rows and columns). I_n is the identity matrix of dimension n. $|a|$ denotes the absolute value of a real number a. $\mathcal{P}_\alpha\{A\}$ denotes the probability of an event A with respect to random variable α, and $\mathcal{P}_\alpha\{A^C\}$ the probability of the complement of A.

II. Problem Statement

We consider a discrete-time linear-time-invariant (LTI) SISO system $\mathcal{S}$ of order n described by the input-output representation

$$\mathcal{S}\colon \begin{cases} z(k+1) = \theta^{oT}\phi(k) \\ y(k) = z(k) + w(k) \end{cases} \tag{1}$$

where the regressor vector $\phi(k)\colon \mathbb{Z} \to \mathbb{R}^{2n}$ is defined as

$$\phi(k) := [z(k), \dots, z(k-n+1), u(k), \dots, u(k-n+1)]^T, \tag{2}$$

u is the manipulated input, z is the nominal output, y is the measured output, w is the measurement noise; $\theta^o = [\theta_1^o, \dots, \theta_{2n}^o]^T \in \mathbb{R}^{2n}$ contains the *unknown* system parameters.

We assume that one batch of input/output data, obtained from a suitable experiment on the system $\mathcal{S}$, is available. The objective of this letter is to propose a data-based control design technique for regulation with asymptotic stability guarantees for the feedback control system in case only input/output data are available and the output is affected by noise as in (1). The next assumption is required on the system.

Assumption 1:

a) The system in (1) is open-loop asymptotically stable;

b) The system order n is known;

c) w is bounded, i.e., $\exists \bar{w} \in \mathbb{R}$ such that $|w(k)| \leq \bar{w}$ for all $k \geq -n+1$.

Assumption 1a) is required to ensure the stationarity properties needed in Section III-C; unstable plants can be handled by a cascade control architecture, provided that a preliminary (possibly low-performing) stabilizing controller is available. Assumption 1b) is necessary to avoid problems with structural identifiability [19, Sec. 5.3]. Data-based estimation of the system order can be obtained by the procedure in [20], subspace identification methods [21, Th. 2], or in view of our considerations in Remark 1. Assumption 1c), which is commonly verified in practice, is required to obtain a bounded disturbance uncertainty set, as we discuss in the next section.

III. Uncertainty Set

The objective of this section is the construction of a disturbance uncertainty set suitable for the application of the data-dependent stability condition in [5, Th. 1].

A. State-Space Model and Disturbance Uncertainty Set

In this section we obtain a state-space model for (1) and define the disturbance uncertainty set. Firstly, we assume that a finite number N_d of output/regressor data pairs $(y(k+1), \hat{\phi}(k))$ is available, for $k = 0, \dots, N_d - 1$, where

$$\hat{\phi}(k) := [y(k), \dots, y(k-n+1), u(k), \dots, u(k-n+1)]^T.$$

In view of this, system (1) can be rewritten as

$$y(k+1) = \theta^{oT}\hat{\phi}(k) + \tilde{d}(k), \tag{3}$$

where $\tilde{d}(k) := -\theta^{oT}\hat{w}(k) + w(k+1)$ and

$$\hat{w}(k) := [w(k), \dots, w(k-n+1), 0_{1,n}]^T.$$

System (3) can be recast in state space form as

$$\begin{cases} x(k+1) = F_\star x(k) + G_\star u(k) + d(k) \\ y(k) = Hx(k) \end{cases} \tag{4}$$

where

$$d(k) := [\tilde{d}(k), 0_{1,2n-2}]^T \in \mathbb{R}^{2n-1} \tag{5}$$

$$x(k) := [y(k), \dots, y(k-n+1), u(k-1), \dots, u(k-n+1)]^T \in \mathbb{R}^{2n-1}, \tag{6}$$

$$F_\star := \begin{bmatrix} \theta_1^o \dots \theta_n^o & \theta_{n+2}^o \dots \theta_{2n}^o \\ I_{n-1}\ 0_{n-1,1} & 0_{n-1,n-1} \\ 0_{1,n} & 0_{1,n-1} \\ 0_{n-2,n} & I_{n-2}\ 0_{n-2,1} \end{bmatrix},$$

$G_\star := \begin{bmatrix}\theta_{n+1}^o & 0_{1,n-1} & 1 & 0_{1,n-2}\end{bmatrix}^T$, and $H := \begin{bmatrix}1 & 0_{1,2n-2}\end{bmatrix}$. Using the available *input/output* data we define the matrices

$$U_0 := \begin{bmatrix} u(0) & u(1) & \cdots & u(N_d-1) \end{bmatrix}, \tag{7a}$$

$$X_0 := \begin{bmatrix} x(0) & x(1) & \cdots & x(N_d-1) \end{bmatrix}, \tag{7b}$$

$$X_1 := \begin{bmatrix} x(1) & x(2) & \cdots & x(N_d) \end{bmatrix}. \tag{7c}$$

We require the next mild and easily verifiable assumption.

Assumption 2: Matrix $\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}$ has full row rank.

Remark 1: Assumption 2 is closely related to the identifiability properties of the system. To understand this, consider noise-free data and recall that the system order is n; through a minimal realization of the difference equation in (1) with matrices $(A_{\min}, B_{\min}, C_{\min}, D_{\min})$ (i.e., where $(A_{\min}, B_{\min})$ is controllable and $(A_{\min}, C_{\min})$ is observable), $\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}$ can be shown to have full row rank by the so-called fundamental lemma [2, Th. 1] if $u(-n+1), \dots, u(N_d-1)$ is persistently exciting of order $2n$.

We can use the considerations above to estimate the system order. Assume that one considers $\hat{n} > n$ as system order and constructs the matrix $\begin{bmatrix} \hat{X}_0 \\ U_0 \end{bmatrix}$ based on $\hat{n}$. Indeed, in view of the fundamental lemma, with noise-free data and $u(-n+1), \ldots, u(N_d - 1)$ persistently exciting of order $2\hat{n}$, the matrix $\begin{bmatrix} \hat{X}_0 \\ U_0 \end{bmatrix}$ cannot be full row rank since $\hat{n}-n$ rows are linearly dependent on the others. More generally, for *noisy* input/output data and fairly good signal-to-noise ratios, $\hat{n}-n$ singular values of $\begin{bmatrix} \hat{X}_0 \\ U_0 \end{bmatrix}$ are negligible with respect to the others. This is analogous to using the numerical rank of a suitable Hankel matrix in subspace identification [19, pp. 181 and 182] to estimate the system order. ◁

We define the unknown disturbance sequence matrix as

$$D_0 := \begin{bmatrix} d(0) & d(1) & \cdots & d(N_d - 1) \end{bmatrix}, \tag{8}$$

so that data satisfy $X_1 = F_\star X_0 + G_\star U_0 + D_0$. In view of Assumption 1c), we have that the disturbance sequence D_0 has bounded energy, i.e., $D_0 \in \mathcal{D}$, where

$$\mathcal{D} := \left\{ D \in \mathbb{R}^{(2n-1)\times N_d} : DD^T \preceq \Delta\Delta^T \right\} \tag{9}$$

for some matrix Δ with $\Delta\Delta^T \in \mathbb{R}^{(2n-1)\times(2n-1)}$. $\mathcal{D}$ is the disturbance uncertainty set. The knowledge of $\Delta\Delta^T$ is required for the application of [5, Th. 1]. Deriving a deterministic bound $\Delta\Delta^T$ only from $\bar{w}$ may be conservative and, hence, we provide an alternative scenario-based estimate of $\Delta\Delta^T$ in the next section.

B. Computation of the Disturbance Upper Bound $\Delta\Delta^T$ via Scenario Optimization

In this section we propose a method based on the scenario approach [22] to obtain an upper bound $\Delta\Delta^T$ for the disturbance with probabilistic guarantees.

We define $\Lambda := \Theta \times \mathbb{W} \subseteq \mathbb{R}^{N_d+3n}$ where $\Theta \subseteq \mathbb{R}^{2n}$ contains parameters θ of LTI SISO asymptotically stable systems of order n in the same representation as (1) and $\mathbb{W} \subseteq \mathbb{R}^{N_d+n}$ contains noise sequences $\mathbf{w} = \begin{bmatrix} w(-n+1) & \ldots & w(N_d) \end{bmatrix}^T$. So, any element of Λ is denoted by $\lambda = \begin{bmatrix} \theta^T & \mathbf{w}^T \end{bmatrix}^T$. In view of Assumption 1a), it is guaranteed that $\lambda^o = \begin{bmatrix} \theta^{oT} & \mathbf{w}^{oT} \end{bmatrix}^T \in \Lambda$, where $\mathbf{w}^o = \begin{bmatrix} w^o(-n+1) & \ldots & w^o(N_d) \end{bmatrix}^T \in \mathbb{W}$ is the unknown noise sequence of the real experiment. With some abuse of notation, we use λ^o in the following to denote its unknown value, but also the related random variable, as clear from the context. Indeed, although θ^o is an unknown constant, it is common procedure (e.g., see [23, Sec. 5]) to assign a probability distribution to it and consider it as a random variable when its estimation is uncertain due to the presence of noise. We require the next technical assumption.

Assumption 3: Λ is endowed with a known probability distribution $\mathbb{P}_\lambda$.

Some guidelines for the estimation of $\mathbb{P}_\lambda$ from data in the absence of any a-priori information are in Section III-C.

To estimate the disturbance upper bound $\Delta\Delta^T$, we propose Algorithm 1 where

$$\mathbf{D} := \begin{bmatrix} 1 & 0_{1,2n-2} \\ 0_{2n-2,1} & 0_{2n-2,2n-2} \end{bmatrix}. \tag{10}$$

Algorithm 1 Estimation of $\Delta\Delta^T$

i) Choose a violation probability $\epsilon \in (0, 1)$, a confidence parameter $\beta \in (0, 1)$, and a number p of scenarios to be discarded.

ii) By bisection, find the minimum integer $N \geq 1$ solving

$$\sum_{j=0}^{p} \binom{N}{j} \epsilon^j (1-\epsilon)^{N-j} \leq \beta. \tag{11}$$

iii) Generate a sample $(\lambda^1, \lambda^2, \ldots, \lambda^N)$ of N independent random elements from $(\Lambda, \mathbb{P}_\lambda)$, where $\lambda^i = [\theta^{i^T} \ \mathbf{w}^{i^T}]^T$, $\mathbf{w}^i = \begin{bmatrix} w^i(-n+1) & \ldots & w^i(N_d) \end{bmatrix}^T$, for $i = 1, \ldots, N$.

iv) For each scenario λ^i, for $k = 0, \ldots, N_d - 1$, define $\hat{w}^i(k) := [w^i(k), \ldots, w^i(k-n+1), 0_{1,n}]^T$, $\tilde{d}^i(k) := -\theta^{i^T}\hat{w}^i(k) + w^i(k+1)$, and the row vector

$$\tilde{\mathbf{d}}^i := \begin{bmatrix} \tilde{d}^i(0) & \tilde{d}^i(1) & \cdots & \tilde{d}^i(N_d - 1) \end{bmatrix}. \tag{12}$$

v) For each scenario λ^i, compute $\alpha^{\lambda^i} := \tilde{\mathbf{d}}^i \tilde{\mathbf{d}}^{iT}$.

vi) Discard p scenarios corresponding to the ones with the greatest α^{λ^i}.

vii) Among the remaining $N-p$ scenarios, take the maximum α^{λ^i}, denoted by α_p^*, and set

$$\Delta\Delta^T = \alpha_p^* \mathbf{D}. \tag{13}$$

As comments to Algorithm 1, we emphasize that the N scenarios involve only λ and *not* output/regressor data pairs, which may be costly to collect. Furthermore, [24, Th. 1.2] provides an explicit (albeit conservative) inequality to upperbound N such that (11) holds, namely,

$$N \geq \epsilon^{-1}\left(p + \sqrt{p} + (\sqrt{p}+1)\ln(1/\beta)\right). \tag{14}$$

From (11) and (14), we note that N does not depend on the dimension of Λ, i.e., on N_d and n. Removing scenarios allows us to avoid too conservative estimates of the disturbance upper bound $\Delta\Delta^T$ while providing probabilistic guarantees that $D_0 \in \mathcal{D}$, as stated next.

Lemma 1: Let Assumption 3 hold. For all $N \geq 1$ fulfilling (11), if $\Delta\Delta^T$ in the set $\mathcal{D}$ in (9) is computed according to Algorithm 1 as in (13), it holds that $\mathcal{P}_{D_0}\{D_0 \in \mathcal{D}\} \geq 1 - \epsilon$ with probability at least $1 - \beta$.

Proof: See the Appendix. ■

Note that $\mathcal{P}_{D_0}\{D_0 \in \mathcal{D}\} \geq 1 - \epsilon$ is itself an event and, in Lemma 1 and in the following Theorem 1, probability $1 - \beta$ is indeed intended with respect to the random multisample extracted from Λ^N, see [25]. As stated in [22], [25], β can be selected small enough to be negligible, e.g., $\beta = 10^{-10}$, without significantly increasing N.

C. Estimation of $\mathbb{P}_\lambda$ From Data

In this section we provide some guidelines on how to estimate $\mathbb{P}_\lambda$. $\mathbb{P}_\theta$ indicates the probability distribution of θ. Furthermore, we consider the noise realizations $w(k)$ as independent and identically distributed random variables with probability distribution $\mathbb{P}_w$, for all $k = -n+1, \ldots, N_d$. It is also reasonable to assume each element of θ to be uncorrelated with $w(k)$, for all $k = -n+1, \ldots, N_d$. Hence, we address separately the problem of estimating $\mathbb{P}_\theta$ and $\mathbb{P}_w$.

Estimation of $\mathbb{P}_\theta$: While referring to [23, Sec. 5] and references therein for more details concerning the estimation of $\mathbb{P}_\theta$, we recall here the main ideas. Firstly, an estimate $\hat{\theta}$ of θ^o

is obtained by using a prediction error identification algorithm and the corresponding covariance matrix $\Sigma_{\hat{\theta}}$ is estimated from data. In case a prior distribution for θ is available, given $\hat{\theta}$, a Bayesian setting can be adopted to obtain a refined estimate of $\mathbb{P}_\theta$, see [23, Proposition 2].

Estimation of $\mathbb{P}_w$: To estimate $\mathbb{P}_w$ a preliminary data collection experiment in a steady-state condition with zero (or constant) output can be performed in view of Assumption 1a). From the analysis of the collected data, a suitable probability distribution can be selected, and the corresponding parameters can be estimated from data using classical methods, see [26] for more details.

D. Set of Models Consistent With Data

In this section we define the set $\mathcal{C}$ of system matrices consistent with data, and we analyze the structure of the matrices therein contained, for $\Delta\Delta^T$ as in (13). Using the data in (7) and the set $\mathcal{D}$ in (9), we define $\mathcal{C}$ as

$$\mathcal{C} := \{[F\ G] : X_1 = FX_0 + GU_0 + D, D \in \mathcal{D}\}, \quad (15)$$

i.e., the set of all matrices $[F\ G]$ that, given U_0, could generate data X_0 and X_1 according to (4) and keep the disturbance sequence in $\mathcal{D}$. Note that $D_0 \in \mathcal{D}$ is equivalent to $[F_\star\ G_\star] \in \mathcal{C}$.

Let us define $\widetilde{D}(\tilde{\mathbf{d}}) := \begin{bmatrix} \tilde{\mathbf{d}} \\ 0_{2n-2,N_d} \end{bmatrix}$, for $\tilde{\mathbf{d}} \in \mathbb{R}^{1\times N_d}$, and the set $\mathcal{C}_s$ of structured matrices, for any $\alpha \geq 0$, as

$$\mathcal{C}_s := \left\{\left[F_s(\theta)\ G_s(\theta)\right] : X_1 = F_s(\theta)X_0 + G_s(\theta)U_0 + \widetilde{D}(\tilde{\mathbf{d}}), \tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \leq \alpha\right\}, \quad (16)$$

where $F_s(\theta) := \begin{bmatrix} \overline{F}_s(\theta) \\ \underline{F}_s \end{bmatrix}$, $G_s(\theta) := \begin{bmatrix} \overline{G}_s(\theta) \\ \underline{G}_s \end{bmatrix}$, $\theta \in \mathbb{R}^{2n}$,

$$\overline{F}_s(\theta) := \begin{bmatrix} \theta_1 & \cdots & \theta_n & \theta_{n+2} & \cdots & \theta_{2n} \end{bmatrix}, \ \overline{G}_s(\theta) := \theta_{n+1},$$

$$\underline{F}_s := \begin{bmatrix} I_{n-1}\ 0_{n-1,1} & 0_{n-1,n-1} \\ 0_{1,n} & 0_{1,n-1} \\ 0_{n-2,n} & I_{n-2}\ 0_{n-2,1} \end{bmatrix}, \ \underline{G}_s := \begin{bmatrix} 0_{n-1,1} \\ 1 \\ 0_{n-2,1} \end{bmatrix}.$$

Also, from (7c), let $X_1 := \begin{bmatrix} \overline{X}_1 \\ \underline{X}_1 \end{bmatrix}$, where $\overline{X}_1 \in \mathbb{R}^{1\times N_d}$ and $\underline{X}_1 \in \mathbb{R}^{2n-2\times N_d}$.

We have the next result on the set $\mathcal{C}_s$.

Proposition 1: Under Assumption 2, if $\Delta\Delta^T = \alpha\mathbf{D}$ for any $\alpha \geq 0$ and $\mathbf{D}$ defined in (10), then $\mathcal{C} = \mathcal{C}_s$.

Proof: See the Appendix. ∎

Proposition 1 shows that, when $\Delta\Delta^T = \alpha\mathbf{D}$ for any $\alpha \geq 0$, the matrices F and G in the set $\mathcal{C}$ in (15) cannot be "full" but need to have the same structure as $F_s(\theta)$ and $G_s(\theta)$ in $\mathcal{C}_s$, which coincides with the structure of $F_\star$ and $G_\star$ in (4).

IV. Data-Based Control Design

In this section the stability condition in [5, Th. 1] is extended to the case when only input/output data are available. To this end, we consider the control law

$$u(k) = Kx(k) \quad (17)$$

where $K^T \in \mathbb{R}^{2n-1}$ and $x(k)$ is defined in (6) as a stack of (past) values of y and u. From data U_0, X_0, X_1 in (7), we define

$$\mathbf{A} := \begin{bmatrix} X_0 \\ U_0 \end{bmatrix}\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}^T, \ \mathbf{B} := -\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}X_1^T, \ \mathbf{C} := X_1X_1^T - \Delta\Delta^T.$$

The next theorem provides a sufficient condition for closed-loop stability and is our main tool for control design.

Theorem 1: Under Assumptions 1, 2 and 3, consider (4) with the feedback control law in (17). If $\Delta\Delta^T$ in the set $\mathcal{D}$ in (9) is computed according to Algorithm 1 as in (13) and there exist $P = P^T \succ 0$ and L such that

$$\begin{bmatrix} -P-\mathbf{C} & 0_{2n-1,2n-1} & \mathbf{B}^T \\ 0_{2n-1,2n-1} & -P & [P\ L^T] \\ \mathbf{B} & \begin{bmatrix} P \\ L \end{bmatrix} & -\mathbf{A} \end{bmatrix} \prec 0, \quad (18)$$

then $K = LP^{-1}$ ensures that, with probability at least $1-\beta$, the unknown closed-loop system is asymptotically stable with probability at least $1-\epsilon$.

Proof: See the Appendix. ∎

V. Simulation Example

In this section we show the effectiveness of the proposed algorithm for a non-minimum phase asymptotically stable system of order $n = 3$. The system is drawn from [27, Sec. IV], with the addition of a nonminimum phase zero. The corresponding discrete-time system, discretized using the zero-order-hold method with a sample time $T_s = 0.125$ s, has unknown parameter vector

$$\theta^o = [1.883\ \ -1.276\ \ 0.2346\ \ 0.7617\ \ -0.346\ \ -0.4948]^T.$$

An additive uniform random noise w acting in the range $[-3, 3]$ affects the output z of the system as in (1).

A dataset composed of $N_d = 1000$ output/regressor data pairs $(y(k+1), \hat{\phi}(k))$ is collected from the plant in open loop. The input signal is a pseudorandom binary sequence in the range $[-10, 10]$, and the signal-to-noise ratio is 25.448 dB.

Algorithm 1 is applied for the estimation of the disturbance upper bound $\Delta\Delta^T$. $\mathbb{P}_\theta$ is estimated according to Section III-C and $\mathbb{P}_w$ is considered uniform in the range $[-3, 3]$. We choose a violation parameter $\epsilon = 0.05$, a confidence parameter $\beta = 10^{-10}$, and the number of scenarios to be discarded $p = 20$. In view of this, the number of required scenarios obtained by applying the bisection algorithm in (11) is $N = 1265$. As a result, we obtain $\alpha_p^* = 21108$. By considering a validation set of 1265 new scenarios λ^i, the percentage of α^{λ^i} greater than α_p^* is 1.66%, less than $\epsilon_\% = 5\%$. The mean and standard deviation of the validation α^{λ^i} are equal to 18705 and 1071, respectively. The actual matrices $[F_\star\ G_\star] \in \mathcal{C}$.

The controller gain K in (17) is obtained by solving (18) using YALMIP [28] and MOSEK. In Figures 1 and 2 we show the regulation results obtained by applying the control law (17) to (1), in terms of trajectories of measured output y and control input u, respectively. In the simulations, an impulse disturbance δ is added to the input-output equation in (1) to perturb the system, i.e., $z(k+1) = \theta^{oT}\phi(k) + \delta(k)$. By inspection of Figure 1 it is possible to see that the closed-loop system

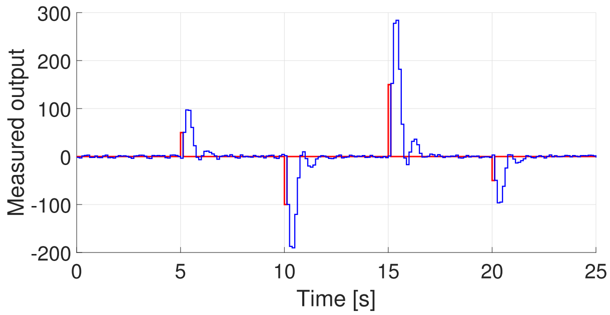

Fig. 1. Discrete-time closed-loop system. Blue line: measured output y; red line: impulse disturbance δ.

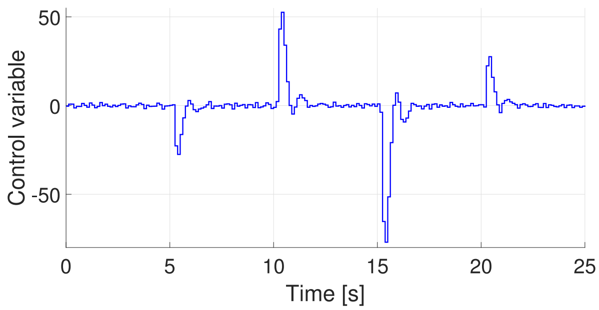

Fig. 2. Discrete-time closed-loop system. Blue line: control variable u.

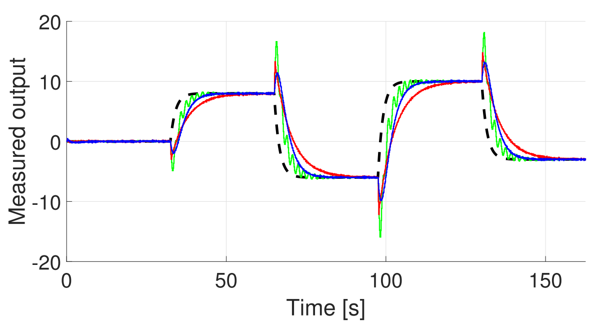

Fig. 3. Measured output trajectories. Black dashed line: reference closed-loop trajectory; blue line: proposed method; red line: [29, Algorithm 2]; green line: [30].

is asymptotically stable. The spectral radius of the unknown closed-loop-system state matrix $F_\star + G_\star K$ is equal to 0.7583.

Finally, the feasibility program in (18) can be easily upgraded to an optimization program to account for performance, e.g., by relying in part on the method in [29, Sec. IV]. Specifically, condition (18) can be used to provide robust stability guarantees in a probabilistic sense when tracking constant references, while minimizing the mismatch of the designed closed loop from a reference closed-loop model, i.e., in place of [29, Th. 1, Eq. (24)]. In Figure 3 the reference tracking trajectories of the proposed method are depicted and compared with the methods in [29, Algorithm 2] and [30], where the same system, control objective, disturbance level, and data are considered. Comparable tracking results are achieved.

VI. Conclusion

In this letter a novel approach based on scenario optimization has been proposed for the data-based control design of linear discrete-time SISO systems with probabilistic asymptotic stability guarantees in case of available input-output data. A simulation example has validated the proposed method.

Future work will investigate the extension to multiple-input multiple-output systems, the application of the method to experimental case studies, and the possibility to consider mismatches in the estimation of $\mathbb{P}_\lambda$.

Appendix

Proof of Lemma 1: Under Assumption 3, we consider a sample $(\lambda^1, \lambda^2, \ldots, \lambda^N)$ of N independent random elements from $(\Lambda, \mathbb{P}_\lambda)$, where N fulfills (11). Without loss of generality, we consider M as a sufficiently large real number such that $\tilde{\mathbf{d}}^i\tilde{\mathbf{d}}^{iT} < M$ for all $i = 1, \ldots, N$. Then, we define $\mathbb{A} := [0, M]$. Note that, in case of no scenario removal, α_0^* is the solution to the scenario program

$$\begin{aligned} &\text{minimize}_{\alpha\in\mathbb{A}} \quad \alpha \\ &\text{subject to} \quad \alpha \in \bigcap_{i=1,\ldots,N} \mathbb{A}_{\lambda^i} \end{aligned} \tag{19}$$

where $\mathbb{A}_{\lambda^i} := \{\alpha \in \mathbb{A} : \tilde{\mathbf{d}}^i\tilde{\mathbf{d}}^{iT} \preceq \alpha\}$. Note that the cost function is linear in α. Moreover, $\mathbb{A}$ and $\mathbb{A}_\lambda$, $\lambda \in \Lambda$, are convex and closed sets in α, and the solution to (19) obtained after discarding p scenarios, denoted by α_p^*, exists and is unique.

In view of these facts, from scenario optimization theory with constraint removal [22], if N fulfills (11), we can state that, with probability at least $1-\beta$, it holds that $\mathcal{P}_\lambda\{\alpha_p^* \notin \mathbb{A}_\lambda\} \leq \epsilon$. By recalling that $\lambda^o \in \Lambda$ and $\tilde{\mathbf{d}}^o := \left[\tilde{d}(0) \cdots \tilde{d}(N_d - 1)\right]$, cf. (12), the previous statement holds also when setting $\lambda = \lambda^o$, meaning that $\mathcal{P}_{\lambda^o}\{\alpha_p^* \notin \mathbb{A}_{\lambda^o}\} \leq \epsilon$. Hence, we have that

$$\epsilon \geq \mathcal{P}_{\lambda^o}\left\{\alpha_p^* \notin \mathbb{A}_{\lambda^o}\right\} = \mathcal{P}_{\tilde{\mathbf{d}}^o}\left\{\tilde{\mathbf{d}}^o\tilde{\mathbf{d}}^{oT} > \alpha_p^*\right\} \overset{(10)}{=}$$
$$\mathcal{P}_{\tilde{\mathbf{d}}^o}\left\{\begin{bmatrix}\tilde{\mathbf{d}}^o\tilde{\mathbf{d}}^{oT} & 0\\ 0 & s0\end{bmatrix} \succ \alpha_p^*\mathbf{D}\right\} \overset{(5),(8)}{=} \mathcal{P}_{D_0}\left\{(D_0 D_0^T \preceq \alpha_p^*\mathbf{D})^C\right\}.$$

Hence, $\mathcal{P}_{D_0}\{D_0D_0^T \preceq \alpha_p^*\mathbf{D}\} \geq 1-\epsilon$ and, with the choice $\Delta\Delta^T = \alpha_p^*\mathbf{D}$, $\mathcal{P}_{D_0}\{D_0 \in \mathcal{D}\} \geq 1-\epsilon$.

Proof of Proposition 1: ($\mathcal{C}_s \subseteq \mathcal{C}$) Immediate since $\widetilde{D}(\tilde{\mathbf{d}}) \in \mathcal{D}$ under the hypothesis.

($\mathcal{C} \subseteq \mathcal{C}_s$) Firstly, let $D =: \begin{bmatrix}\tilde{\mathbf{d}}\\ \underline{D}\end{bmatrix}$ with $\tilde{\mathbf{d}} \in \mathbb{R}^{1\times N_d}$ and $\underline{D} \in \mathbb{R}^{2n-2\times N_d}$. Secondly, note that $DD^T = \begin{bmatrix}\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T & \tilde{\mathbf{d}}\underline{D}^T\\ \underline{D}\tilde{\mathbf{d}}^T & \underline{D}\underline{D}^T\end{bmatrix}$ and that $\underline{D}\underline{D}^T \succeq 0$. Moreover, for any $D \in \mathcal{D}$, if $\Delta\Delta^T = \alpha\mathbf{D}$, for any $\alpha \geq 0$, with $\mathbf{D}$ defined as in (10), it must also hold that $\underline{D}\underline{D}^T \preceq 0$. It follows that $\underline{D}\underline{D}^T = 0_{2n-2,2n-2}$ and that $\underline{D} = 0_{2n-2,N_d}$. Since the actual disturbance sequence D_0 satisfies $D_0 \in \mathcal{D}$, we have $D_0 = \begin{bmatrix}\tilde{\mathbf{d}}^o\\ 0_{2n-2,N_d}\end{bmatrix}$ for some $\tilde{\mathbf{d}}^o \in \mathbb{R}^{1\times N_d}$.

Hence, under the hypothesis, we have that $\mathcal{C} = \widetilde{\mathcal{C}}$ for

$$\begin{aligned}\widetilde{\mathcal{C}} &:= \left\{\left[F\ G\right] : X_1 = FX_0 + GU_0 + \widetilde{D}(\tilde{\mathbf{d}}), \widetilde{D}(\tilde{\mathbf{d}}) \in \mathcal{D}\right\}\\ &= \left\{\left[F\ G\right] : X_1 = FX_0 + GU_0 + \widetilde{D}(\tilde{\mathbf{d}}), \tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \leq \alpha\right\}.\end{aligned}$$

Let us consider an arbitrary element $\begin{bmatrix} F & G \end{bmatrix} \in \widetilde{\mathcal{C}}$ and let us show that $\begin{bmatrix} F & G \end{bmatrix} \in \mathcal{C}_s$. $\begin{bmatrix} F & G \end{bmatrix}$ can be partitioned as $\begin{bmatrix} F & G \end{bmatrix} =: \begin{bmatrix} \overline{F} & \overline{G} \\ \underline{F} & \underline{G} \end{bmatrix}$ for $\overline{F} =: \begin{bmatrix} \overline{F}_1 & \overline{F}_2 \end{bmatrix} \in \mathbb{R}^{1\times 2n-1}$, $\overline{F}_1 \in \mathbb{R}^{1\times n}$, $\overline{F}_2 \in \mathbb{R}^{1\times n-1}$, $\underline{F} \in \mathbb{R}^{2n-2\times 2n-1}$, $\overline{G} \in \mathbb{R}^{1\times 1}$, $\underline{G} \in \mathbb{R}^{2n-2\times 1}$. By these definitions, $\begin{bmatrix} F & G \end{bmatrix} \in \widetilde{\mathcal{C}}$ if and only if

$$\overline{X}_1 = \overline{F}X_0 + \overline{G}U_0 + \tilde{\mathbf{d}}, \ \ \tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \leq \alpha, \tag{20a}$$

$$\underline{X}_1 = \underline{F}X_0 + \underline{G}U_0. \tag{20b}$$

With the definitions in (16) and $\tilde{\theta} := \begin{bmatrix} \overline{F}_1 & \overline{G} & \overline{F}_2 \end{bmatrix}$, (20a) rewrites

$$\overline{X}_1 = \overline{F}_s(\tilde{\theta})X_0 + \overline{G}_s(\tilde{\theta})U_0 + \tilde{\mathbf{d}}, \ \ \tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \leq \alpha. \tag{21a}$$

Data satisfy $X_1 = F_\star X_0 + G_\star U_0 + D_0$ since they are generated by (4), and D_0 satisfies $D_0 = \begin{bmatrix} \tilde{\mathbf{d}}^o \\ 0_{2n-2,N_d} \end{bmatrix}$ for some $\tilde{\mathbf{d}}^o \in \mathbb{R}^{1\times N_d}$ as noted above. Hence,

$$\underline{X}_1 = \underline{F}_s X_0 + \underline{G}_s U_0 \tag{21b}$$

by the definitions of $F_\star$ and $G_\star$. Suppose by contradiction that, for the considered $\begin{bmatrix} F & G \end{bmatrix} \in \widetilde{\mathcal{C}}$, $\begin{bmatrix} \underline{F} & \underline{G} \end{bmatrix}$ as in (20b) satisfy $\begin{bmatrix} \underline{F} & \underline{G} \end{bmatrix} \neq \begin{bmatrix} \underline{F}_s & \underline{G}_s \end{bmatrix}$. Let us subtract (21b) from (20b) so that

$$\left(\begin{bmatrix} \underline{F} & \underline{G} \end{bmatrix} - \begin{bmatrix} \underline{F}_s & \underline{G}_s \end{bmatrix}\right)\begin{bmatrix} X_0 \\ U_0 \end{bmatrix} = 0_{2n-2,N_d}.$$

If $\begin{bmatrix} \underline{F} & \underline{G} \end{bmatrix} \neq \begin{bmatrix} \underline{F}_s & \underline{G}_s \end{bmatrix}$, there would exist some non-zero vector in the left kernel of $\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}$, but this contradicts Assumption 2. Thus, $\begin{bmatrix} \underline{F} & \underline{G} \end{bmatrix} = \begin{bmatrix} \underline{F}_s & \underline{G}_s \end{bmatrix}$ is the only solution to (20b). In summary, we have shown that for an arbitrary $\begin{bmatrix} F & G \end{bmatrix} \in \widetilde{\mathcal{C}}$, there exist $\tilde{\theta}$ and $\tilde{\mathbf{d}}$ such that $\begin{bmatrix} F & G \end{bmatrix} = \begin{bmatrix} \overline{F}_s(\tilde{\theta}) & \overline{G}_s(\tilde{\theta}) \\ \underline{F}_s & \underline{G}_s \end{bmatrix}$ and $\tilde{\mathbf{d}}\tilde{\mathbf{d}}^T \leq \alpha$, by (21a) and (21b). In other words, we have shown that $\begin{bmatrix} F & G \end{bmatrix} \in \mathcal{C}_s$.

Proof of Theorem 1: Under Assumption 1 it is possible to define the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, which are necessary to express (18). By Assumption 2 and [5, Th. 1], if (18) holds for some $P = P^T \succ 0$ and L, we have that the closed-loop system is asymptotically stable (i.e., $F + GK$ is Schur) for all $\begin{bmatrix} F & G \end{bmatrix} \in \mathcal{C}$. Hence, the probability that the unknown closed-loop system is asymptotically stable is greater than or equal to $\mathcal{P}_{\theta^o}\{\begin{bmatrix} F_\star & G_\star \end{bmatrix} \in \mathcal{C}\} = \mathcal{P}_{D_0}\{D_0 \in \mathcal{D}\}$. Under Assumption 3, if $\Delta\Delta^T$ is computed according to Algorithm 1, the claim follows straightforwardly from Lemma 1.

Acknowledgment

The authors thank Prof. Claudio De Persis, Prof. Pietro Tesi, and Prof. Simone Garatti for useful suggestions and fruitful discussions.

References

[1] L. Fraile, M. Marchi, and P. Tabuada, "Data-driven stabilization of SISO feedback linearizable systems," 2020, *arXiv:2003.14240*.

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

[3] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, Mar. 2020.

[4] M. Guo, C. De Persis, and P. Tesi, "Data-driven stabilization of nonlinear polynomial systems with noisy data," *IEEE Trans. Autom. Control*, vol. 67, no. 8, pp. 4210–4217, Aug. 2022.

[5] A. Bisoffi, C. De Persis, and P. Tesi, "Data-driven control via Petersen's lemma," *Automatica*, vol. 145, Nov. 2022, Art. no. 110537.

[6] C. De Persis, M. Rotulo, and P. Tesi, "Learning controllers from data via approximate nonlinearity cancellation," *IEEE Trans. Autom. Control*, early access, Jan. 6, 2023, doi: 10.1109/TAC.2023.3234889.

[7] A. Koch, J. Berberich, and F. Allgöwer, "Provably robust verification of dissipativity properties from data," *IEEE Trans. Autom. Control*, vol. 67, no. 8, pp. 4248–4255, Aug. 2022.

[8] A. Bisoffi, C. De Persis, and P. Tesi, "Controller design for robust invariance from noisy data," *IEEE Trans. Autom. Control*, vol. 68, no. 1, pp. 636–643, Jan. 2023.

[9] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *Proc. Eur. Control Conf.*, 2019, pp. 307–312.

[10] G. Baggio, V. Katewa, and F. Pasqualetti, "Data-driven minimum-energy controls for linear systems," *IEEE Control Syst. Lett.*, vol. 3, no. 3, pp. 589–594, Jul. 2019.

[11] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, "From noisy data to feedback controllers: Non-conservative design via a matrix S-lemma," *IEEE Trans. Autom. Control*, vol. 67, no. 1, pp. 162–175, Jan. 2022.

[12] J. Berberich, C. W. Scherer, and F. Allgöwer, "Combining prior knowledge and data for robust controller design," *IEEE Trans. Autom. Control*, early access, Sep. 26, 2022, doi: 10.1109/TAC.2022.3209342.

[13] H. J. van Waarde, J. Eising, M. K. Camlibel, and H. L. Trentelman, "A behavioral approach to data-driven control with noisy input-output data," *IEEE Trans. Autom. Control*, early access, May 10, 2023, doi: 10.1109/TAC.2023.3275014.

[14] J. Miller, T. Dai, and M. Sznaier, "Superstabilizing control of discrete-time ARX models under error in variables," 2022, *arXiv:2210.14893*.

[15] A. Sassella, V. Breschi, and S. Formentin, "Noise handling in data-driven predictive control: A strategy based on dynamic mode decomposition," in *Proc. Learn. Dyn. Control Conf.*, 2022, pp. 74–85.

[16] T. Söderström, "Errors-in-variables methods in system identification," *IFAC Proc. Vol.*, vol. 39, no. 1, pp. 1–19, 2006.

[17] V. Cerone, D. Piga, and D. Regruto, "Set-membership error-in-variables identification through convex relaxation techniques," *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 517–522, Feb. 2012.

[18] V. Cerone, V. Razza, and D. Regruto, "Set-membership errors-in-variables identification of MIMO linear systems," *Automatica*, vol. 90, pp. 25–37, Apr. 2018.

[19] S. Bittanti, *Model Identification and Data Analysis*. Hoboken, NJ, USA: Wiley, 2019.

[20] M. Lauricella and L. Fagiano, "Set membership identification of linear systems with guaranteed simulation accuracy," *IEEE Trans. Autom. Control*, vol. 65, no. 12, pp. 5189–5204, Dec. 2020.

[21] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Boston, MA, USA: Kluwer, 1996.

[22] M. C. Campi and S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," *J. Optim. Theory Appl.*, vol. 148, no. 2, pp. 257–280, 2011.

[23] M. C. Campi, A. Lecchini, and S. M. Savaresi, "Virtual reference feedback tuning (VRFT): A new direct approach to the design of feedback controllers," in *Proc. 39th IEEE Conf. Decis. Control*, vol. 1, 2000, pp. 623–629.

[24] M. C. Campi and S. Garatti, *Introduction to the Scenario Approach*. Philadelphia, PA, USA: SIAM, 2018.

[25] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 742–753, May 2006.

[26] S. M. Ross, *Introduction to Probability and Statistics for Engineers and Scientists*. London, U.K.: Academic, 2020.

[27] W. D'Amico and M. Farina, "Data-based control design for linear discrete-time systems with robust stability guarantees," in *Proc. 61st IEEE Conf. Decis. Control*, 2022, pp. 1429–1434.

[28] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, 2004, pp. 284–289.

[29] W. D'Amico and M. Farina, "Virtual reference feedback tuning for linear discrete-time systems with robust stability guarantees based on set membership," 2022, *arXiv:2203.00088*.

[30] G. Battistelli, D. Mari, D. Selvi, and P. Tesi, "Direct control design via controller unfalsification," *Int. J. Robust Nonlinear Control*, vol. 28, no. 12, pp. 3694–3712, 2018.

Open Access funding provided by 'Politecnico di Milano' within the CRUI CARE Agreement