Systems & Control Letters 145 (2020) 104788

Contents lists available at ScienceDirect

Systems & Control Letters

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

Data-based stabilization of unknown bilinear systems with guaranteed basin of attraction[☆]

Andrea Bisoffi [a,*], Claudio De Persis [a], Pietro Tesi [b]

[a] ENTEG and the J.C. Willems Center for Systems and Control, University of Groningen, 9747 AG Groningen, The Netherlands
[b] DINFO, University of Florence, 50139 Florence, Italy

ARTICLE INFO

Article history:
Received 24 April 2020
Received in revised form 22 July 2020
Accepted 13 September 2020
Available online 1 October 2020

Keywords:
Direct data-driven control
Bilinear systems
Lyapunov functions
Basin of attraction
Linear matrix inequalities

ABSTRACT

Motivated by the goal of having a building block in the design of direct data-driven controllers for nonlinear systems, we show how, for an unknown discrete-time bilinear system, the data collected in an offline open-loop experiment enable us to design a feedback controller and provide a guaranteed underapproximation of its basin of attraction. Both can be obtained by solving a linear matrix inequality for a fixed scalar parameter, and possibly iterating on different values of that parameter. The results of this data-based approach are compared with the ideal case when the model is known perfectly.

1. Introduction

Direct data-driven control aims at learning control laws through input–output data collected from on-line or off-line experiments on the system, avoiding the explicit identification of a model. Most of the research works in this area focused on linear systems, including the design of model-reference controllers [1,2] and, more recently, robust and optimal control design [3–6]. An overview of early accounts on this topic is in [7]. In contrast, direct data-driven control for *nonlinear* systems has been much less explored, but is gaining more and more attention also thanks to many impressive experimental results achieved by machine learning algorithms in, e.g., self-driving cars [8]. Contributions to data-driven control for nonlinear systems can be found in the context of intelligent-PID design [9], finite-gain stabilization for Lipschitz continuous nonlinear systems [10], feedback linearization [11], safety control [12], and predictive control [13]. Our paper contributes to this research area with data-driven design of stabilizing controllers for bilinear systems.

Direct data-driven control has the potential to overcome the difficulties related to learning an accurate model of the system to control. However, stability guarantees are more difficult to obtain. To address the intrinsic difficulty of dealing with the control design of unknown nonlinear systems, a natural approach is to reduce their complexity by considering the system evolution along a given Lyapunov function. This classical control theoretic analysis is enhanced by nonparametric regression methods from machine learning to cope with the large uncertainty in the model [14] and is performed using a sufficiently dense set of samples taken from the system. Analytical guarantees of stability and safety are then obtained relying on additional tools from robust control and optimization [12]. The approach of [9,11] to reduce the complexity of controlling unknown nonlinear systems consists of considering systems with a well-defined relative degree, in such a way that the uncertainty only appears in the form of two Lie derivatives of the output function along the system vector fields. Once the dynamics has been discretized, the key observation from sampled-data control theory is that these uncertain functions are constant between sampling times for a sufficiently high sampling rate.

A different approach to data-driven control of nonlinear systems has been recently taken in a series of works that use the nonparametric representation of dynamical systems via Hankel matrices of finite-size input–output data proposed in [15]. On one hand, this representation has given rise to data-enabled predictive controllers where the effect of the nonlinearity is taken into account by a regularized optimization problem [13,16]. On the other hand, it inspired a data-dependent parametrization of the closed-loop system that reduces the control design to semidefinite programs where the nonlinearity is dealt with as a process disturbance [17]. Further results along this research thread have been proposed in [18]. While these results make possible to deal with nonlinear systems, they provide local stability results. Very

☆ The research for this article was partially conducted under the umbrella of the Centre for Data Science and Systems Complexity (DSSC) at the University of Groningen.

* Corresponding author.
E-mail addresses: a.bisoffi@rug.nl (A. Bisoffi), c.de.persis@rug.nl (C. De Persis), pietro.tesi@unifi.it (P. Tesi).

https://doi.org/10.1016/j.sysconle.2020.104788

recently, within the research thread of [15, Thm. 1], there have been efforts to go beyond the local nature of the results for special classes of nonlinear systems, studying data-driven control of second-order discrete Volterra systems [19] and polynomial systems [20].

The goal of this paper is to characterize another notable class of nonlinear systems for which nonlocal data-driven control results can be established, namely bilinear systems. The reason for focusing on bilinear systems is threefold. In spite of their simple nonlinear structure, applying Carleman linearization to a generic continuous-time input-affine nonlinear system yields a continuous-time bilinear system with a larger state plus a remainder (see [21,22]), so bilinear systems can be used as universal approximators of input-affine nonlinear systems [23, p. 110]. This last consideration specifically motivates the proposed data-driven control scheme for bilinear systems, which is envisioned to be a building block in future work on direct data-driven control of input-affine nonlinear systems (see also the discussion in Remark 3). A second motivation is to provide a method alternative to sum-of-squares programming for polynomial control systems [20] to directly design data-driven controllers of bilinear systems. Finally, bilinear systems are interesting *per se* as meaningful models for a number of relevant applications in engineering, biology and ecology [24,25].

Many model-based approaches have been proposed for control of bilinear systems such as [26–29], and we refer the reader to [29, §1] for a thorough overview. Such model-based approaches assume the knowledge of the parameters of the bilinear system. When these are not known from first-principles considerations, one can resort to system identification techniques tailored for bilinear systems, and then apply one of the model-based approaches above. Some of these *indirect* data-driven methods for system identification are [30–32], see also [33, Part II] for an overview. Although combining the aforementioned system identification techniques with model-based design constitutes a natural and valid way to control a bilinear system, we aim here at exploring the less-investigated direct control design of a bilinear system based on data (avoiding altogether a system identification step generally nontrivial in a nonlinear setting). We show that under mild assumptions (see Assumption 1), it is indeed possible to design stabilizing control policies directly from data. We also show via simulations that our approach compares well with a model-based design that has *perfect* knowledge of the parameters of the system, regardless of whether this knowledge derives from first-principles considerations or from a preliminary system identification step.

In the case of data generated by an underlying linear system, the fundamental result [15, Thm. 1] has been shown in [17] to allow direct data-driven design of feedback controllers (with robustness to noise) for linear systems through linear matrix inequalities (LMI) [34] and the *local* stabilization of nonlinear systems through semidefinite programs. In the case of data generated by an underlying bilinear system, the arguments in [17] need substantial modifications to counteract the nonlinear term appearing in the bilinear system and to explicitly provide an estimate of the region of attraction. Thus, we need to resort to tools from robust control (such as [29,35], see Fact 1) besides more standard ones from linear matrix inequalities. Some conservatism is introduced in these steps compared to a model-based approach, as illustrated in Section 4.

Similar to the model-based approaches [27,29] and, partially, to [26,28], we also adopt a linear state feedback and a quadratic Lyapunov function in the design of the closed-loop system. Alternatives are based on rational polynomial controllers and sum-of-squares programming [36]. The choice of linear controllers is restrictive compared to nonlinear state feedback (and the actual basin of attraction has not an ellipsoidal shape), but are dictated by the desire of obtaining a computationally tractable result in the form of linear matrix inequalities (after fixing a scalar parameter). However, the main difference with those model-based approaches is that we design here the linear state feedback and the quadratic Lyapunov function *without* relying on the knowledge of the bilinear system matrices, which we aim to substitute instead through data collected from the bilinear system.

Tuning a feedback controller based only on a limited number of open-loop data, which gives a guaranteed subset of the basin of attraction for a bilinear system, is the main contribution of this paper.

Structure. The considered problem is formulated in Section 2. In Section 3 we provide our data-based controller for the unknown bilinear system with a guaranteed underapproximation of its basin of attraction, as a main result. Section 4 compares this solution with a model-based one on a numerical example.

Notation. For a matrix A, $\|A\|$ denotes the induced 2-norm. For a symmetric matrix $\left[\begin{smallmatrix} A & B \\ B^\top & C \end{smallmatrix}\right]$, we may use the shorthand writing $\left[\begin{smallmatrix} A & B \\ \star & C \end{smallmatrix}\right]$. I denotes an identity matrix of appropriate dimensions.

2. System description and problem formulation

Consider the discrete-time bilinear system

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

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}$ is the input, and the system matrices have dimensions $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^n$, $D \in \mathbb{R}^{n\times n}$. Our choice to consider a *scalar* input in (1) is motivated in Remark 2 after we have outlined our approach. The matrices A, B, D are completely unknown apart from a bound on the matrix norm of D as follows.

Assumption 1. For some known $\delta > 0$, the matrix D satisfies $\|D\| \le \delta$ (equivalently, $D^\top D \preceq \delta^2 I$).

Assumption 1 amounts to having prior information on the *strength* of the nonlinear coupling. An upper bound on $\|D\|$ can be obtained, e.g., from the knowledge of a Lipschitz constant for the system on some compact set [10]. Clearly, as exemplified numerically in Section 4, such prior information influences the solution in the sense that looser bounds on $\|D\|$ lead to less performing control laws.

Our objective is to design a controller $u = Kx$ for the bilinear system in (1) based only on data collected from an off-line experiment (namely, without identifying the matrices A, B, D) and give a guaranteed underapproximation of the basin of attraction of the origin for the closed-loop system. The off-line experiment of duration T (with $T > 0$) collects the input and state sequences $u(0), u(1), \dots, u(T-1)$ and $x(0), x(1), \dots, x(T)$. These are organized as

$$U_{0,T} := \begin{bmatrix} u(0) & u(1) & \dots & u(T-1) \end{bmatrix} \tag{2a}$$

$$X_{0,T} := \begin{bmatrix} x(0) & x(1) & \dots & x(T-1) \end{bmatrix} \tag{2b}$$

$$X_{1,T} := \begin{bmatrix} x(1) & x(2) & \dots & x(T) \end{bmatrix}, \tag{2c}$$

and allow computing the auxiliary quantity

$$V_{0,T} := \begin{bmatrix} x(0)u(0) & x(1)u(1) & \dots & x(T-1)u(T-1) \end{bmatrix}. \tag{2d}$$

Following [17], we reparametrize the gain K by a matrix G_K and give in the next lemma an equivalent representation of (1) in closed loop with $u = Kx$, which depends on data, except for the matrix D.

Lemma 1. *Let $G_K \in \mathbb{R}^{T\times n}$ satisfy*

$$I = X_{0,T} G_K. \tag{3}$$

Then, system (1) with state feedback $u = Kx$ and $K = U_{0,T} G_K$ has the equivalent representation

$$x^+ = (X_{1,T} - DV_{0,T} + DxU_{0,T})G_K x =: g_D(x)x. \tag{4}$$

Proof. (1) with state feedback $u = Kx$ becomes $x^+ = (A + BK + DxK)x$. This closed-loop matrix is, by (3),

$$\begin{aligned} &A + BK + DxK = A \cdot I + BK + DxK \\ &= AX_{0,T}G_K + BU_{0,T}G_K + DxU_{0,T}G_K \\ &= (AX_{0,T} + BU_{0,T} + DxU_{0,T})G_K \\ &= (X_{1,T} - DV_{0,T} + DxU_{0,T})G_K, \end{aligned}$$

since the data in (2) satisfy $X_{1,T} = AX_{0,T} + BU_{0,T} + DV_{0,T}$, and this proves the statement. □

The reparametrization G_K is a decision variable that we tune to achieve our control objective. Based on G_K and on data, we define for compactness

$$\mathcal{A}_c := X_{1,T}G_K,\ \mathcal{F} := I,\ \mathcal{H} := -V_{0,T}G_K,\ \mathcal{K} := U_{0,T}G_K, \tag{5}$$

so that the closed-loop representation in (4) becomes

$$x^+ = (\mathcal{A}_c + \mathcal{F}D\mathcal{H} + Dx\mathcal{K})x = g_D(x)x, \tag{6}$$

where D is highlighted and its presence will be removed in Section 3 thanks to Assumption 1. We aim at giving a guaranteed underapproximation of the basin of attraction of the closed-loop system in (6). We do so by considering a quadratic Lyapunov function

$$V(x) = x^\top Q x \tag{7}$$

with $Q = Q^\top \succ 0$ and imposing the strict decrease of $V\big(g_D(x)x\big) - V(x)$ for the dynamics in (6). The last quantity is easily computed as in the next lemma.

Lemma 2. *We have that $V\big(g_D(x)x\big) - V(x) = x^\top \mathcal{N}_D(x)x$ with $\mathcal{N}_D(x)$ defined as*

$$\begin{aligned} &\mathcal{N}_D(x) := (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top Q(\mathcal{A}_c + \mathcal{F}D\mathcal{H}) - Q \\ &+ (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QDx\mathcal{K} + \mathcal{K}^\top x^\top D^\top Q(\mathcal{A}_c + \mathcal{F}D\mathcal{H}) \\ &+ \mathcal{K}^\top x^\top D^\top QDx\mathcal{K}. \end{aligned} \tag{8}$$

Proof. The expression for $\mathcal{N}_D(x)$ is immediate by substituting (6) in $V\big(g_D(x)x\big) - V(x)$. □

Note that for $D = 0$, (1) becomes linear and (8) reduces to $\mathcal{N}_D(x) = \mathcal{A}_c^\top Q\mathcal{A}_c - Q$, corresponding to the classical Lyapunov condition for discrete-time linear systems. We impose $V\big(g_D(x)x\big) - V(x) < 0$ for all $x \neq 0$ in the ellipsoid

$$\mathcal{E}_Q := \{x \in \mathbb{R}^n : x^\top Q x \le 1\}, \tag{9}$$

by designing the decision variables G_K, which determines $K = U_{0,T}G_K$, and Q, which will be optimized to maximize the volume of the ellipsoid $\mathcal{E}_Q$. The design will be based only on data, and return the ellipsoid $\mathcal{E}_Q$ as a guaranteed underapproximation of the basin of attraction. With the outlined method using decision variables G_K and Q, the problem we address is stated as follows:

Problem 1. Based only on the data in (2) collected from an off-line experiment and the bound δ in Assumption 1, obtain a controller $u = Kx$ for (1) such that for the closed-loop system, the origin has a guaranteed basin of attraction.

Some remarks are in order.

Remark 1 (*Quality of Data*)**.** The existence of a matrix G_K satisfying (3) is related to the "quality" of the experimental data. In fact, condition (3) expresses the property that the data are sufficiently rich so that the system dynamics can be parametrized directly in terms of the matrices in (2). A key property established in [15] is that, for linear systems, $X_{0,T}$ is full-row rank (thus, a solution G_K to (3) exists) when the experiment is carried out using a sufficiently exciting input signal. An extension of this property to nonlinear systems is discussed in [37] where it is shown that under prior knowledge of an upper bound on the nonlinearity (in fact, on D in the present case of bilinear systems) one can always design experiments so that (3) is feasible.

Remark 2 (*Multi-Input Bilinear Systems*)**.** The present analysis can be extended to bilinear systems with input $u \in \mathbb{R}^m$ and $m \ge 2$. For $m = 2$, (1) can be written for $u = \left[\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right]$ as

$$x^+ = Ax + B_1u_1 + B_2u_2 + D_1xu_1 + D_2xu_2. \tag{10}$$

We can define $U_{0,T}^{(1)}$ and $U_{0,T}^{(2)}$ as in (2a), but considering respectively the components u_1 and u_2. Similarly, we can define $V_{0,T}^{(1)}$ and $V_{0,T}^{(2)}$ as in (2d). Based on the very same steps as in Lemma 1, we can obtain for $U_{0,T} = \begin{bmatrix} U_{0,T}^{(1)} \\ U_{0,T}^{(2)} \end{bmatrix}$ the next equivalent representation of (10)

$$x^+ = (X_{1,T} - D_1V_{0,T}^{(1)} - D_2V_{0,T}^{(2)} + D_1xU_{0,T}^{(1)} + D_2xU_{0,T}^{(2)})G_K x.$$

This expression shows by comparison with (4) that the case for $m = 2$ can be treated using the same procedure we develop in the presence of a single unknown D, and this consideration easily generalizes to m larger than 2. For this reason we focus on the essential case with input $u \in \mathbb{R}$.

Remark 3 (*Continuous Time*)**.** The universal approximation property of bilinear systems mentioned in Section 1 holds with respect to continuous-time nonlinear systems. We focus here on discrete-time bilinear systems since the data in (2) are samples obtained from experiments. However, analogous results can be obtained for continuous-time bilinear systems if $X_{1,T}$ in (2c) is replaced by samples of the state derivative. These results would then lend themselves to the analysis of a bilinear approximation of continuous-time nonlinear systems (provided disturbances are accounted for, e.g., using the result in Section 3.1).

3. Data-based solution with guaranteed basin of attraction

In Section 2, we showed that data allow expressing (1) in closed loop with $u = Kx$ as (6) (by introducing the reparametrization G_K of K). Data, however, did not allow us to completely remove the matrices of model (1). In particular, $g_D(x)$ in (6) still contains two instances of the matrix D (namely, $Dx\mathcal{K}$ and $\mathcal{F}D\mathcal{H}$), which can both be interpreted as a perturbation of the matrix $\mathcal{A}_c$. In this section we first address the former, which is more standard and occurs analogously for model-based design of a bilinear system (see, e.g., [29]), and then the latter, which is motivated by our desire to solve Problem 1 based only on data and calls for the matrix norm bound in Assumption 1.

Before presenting the developments of this section, we recall an auxiliary result from [35], which has been reported in a convenient form as [29, Lemma 1] and is related to the S-procedure [34, §2.6.3]. In particular, [29, Lemma 1] implies the next fact.

Fact 1 (*[29, Lemma 1]*)**.** *Let $\mathsf{G} = \mathsf{G}^\top \in \mathbb{R}^{\mathsf{n}\times\mathsf{n}}$, $\mathsf{M} \in \mathbb{R}^{\mathsf{n}\times\mathsf{p}}$, $\mathsf{N} \in \mathbb{R}^{\mathsf{n}\times\mathsf{q}}$.*

$$\mathsf{G} + \mathsf{M}\mathsf{D}\mathsf{N}^\top + \mathsf{N}\mathsf{D}^\top\mathsf{M}^\top \prec 0 \quad \text{for all } \mathsf{D} \in \mathbb{R}^{\mathsf{p}\times\mathsf{q}} \text{ with } \|\mathsf{D}\| \le 1 \tag{11}$$

if there exists a scalar e *such that*

$$\begin{bmatrix} \mathsf{G} + \mathsf{e}\mathsf{M}\mathsf{M}^\top & \mathsf{N} \\ \mathsf{N}^\top & -\mathsf{e}I \end{bmatrix} \prec 0. \tag{12}$$

With Fact 1 we are in a position to develop this section. The next lemma addresses the term DxK in $g_D(x)$ in (6). Specifically, it shows that as long as we restrict the analysis to a sublevel set $\mathcal{E}_Q$ (defined in (9)) of the Lyapunov function V in (7) (where Q itself is a decision variable determining the size of this sublevel set), strict decrease of V along solutions is guaranteed ($\mathcal{N}_D(x) \prec 0$) since $\mathcal{N}_D(x)$ determines $V\big(g_D(x)x\big) - V(x)$ as in Lemma 2.

Lemma 3. *If there exist* $\tau \in \mathbb{R}$ *and* $Q = Q^\top \in \mathbb{R}^{n\times n}$ *such that*

$$\begin{bmatrix} -Q & 0 & \mathcal{K}^\top & (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top \\ \star & -\tau Q & 0 & D^\top \\ \star & \star & -\frac{1}{\tau}I & 0 \\ \star & \star & \star & -Q^{-1} \end{bmatrix} \prec 0, \tag{13}$$

then $\mathcal{N}_D(x) \prec 0$ *for all* $x \in \mathcal{E}_Q$.

Proof. The proof follows closely [29], but is reported for self-containedness. Define for compactness

$$\mathcal{R} := -Q + (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top Q(\mathcal{A}_c + \mathcal{F}D\mathcal{H}) \tag{14}$$

and note for the following that (13) implies $Q \succ 0$ and $\tau > 0$. By Schur's complement (with respect to lowest block $-Q^{-1}$), (13) is equivalent, by (14), to

$$\begin{bmatrix} \mathcal{R} & (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QD & \mathcal{K}^\top \\ \star & -\tau Q + D^\top QD & 0 \\ \star & \star & -\frac{1}{\tau}I \end{bmatrix} \prec 0.$$

By Schur's complement, this inequality is equivalent to

$$\begin{bmatrix} \mathcal{R} & (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QD & \mathcal{K}^\top & 0 \\ \star & -\tau Q & 0 & D^\top Q \\ \star & \star & -\frac{1}{\tau}I & 0 \\ \star & \star & \star & -Q \end{bmatrix} \prec 0.$$

Rearranging rows and columns of this inequality gives

$$\begin{bmatrix} \mathcal{R} & 0 & (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QD & \mathcal{K}^\top \\ \star & -Q & QD & 0 \\ \star & \star & -\tau Q & 0 \\ \star & \star & \star & -\frac{1}{\tau}I \end{bmatrix} \prec 0,$$

which is equivalent to (13). We want to put this inequality in a form where we can apply Fact 1. Then, we pre- and post-multiply the previous inequality by the block diagonal matrix with entries I, I, $(Q^{1/2})^{-1}$, I (where $Q^{1/2}$ is the unique symmetric, positive definite square root matrix for $Q = Q^\top \succ 0$ [38, Thm. 7.2.6], so that $Q = Q^{1/2}Q^{1/2}$) and apply Schur's complement (with respect to the lowest block $-\frac{1}{\tau}I$) to obtain with some computations

$$\left[\begin{array}{c|c} \begin{bmatrix} \mathcal{R} & 0 \\ 0 & -Q \end{bmatrix} + \tau \begin{bmatrix} \mathcal{K}^\top \\ 0 \end{bmatrix} \begin{bmatrix} \mathcal{K} & 0 \end{bmatrix} & \begin{bmatrix} (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QD(Q^{1/2})^{-1} \\ QD(Q^{1/2})^{-1} \end{bmatrix} \\ \hline \star & -\tau I \end{array}\right] \prec 0.$$

Note that $x^\top Qx = (x^\top Q^{1/2})(Q^{1/2}x)$, hence for all x such that $x^\top Qx \le 1$, $\|x^\top Q^{1/2}\| \le 1$. With this observation and by Fact 1 we conclude, after some simplifications, that

$$\begin{bmatrix} \mathcal{R} & 0 \\ 0 & -Q \end{bmatrix} + \begin{bmatrix} \mathcal{K}^\top \\ 0 \end{bmatrix} x^\top \begin{bmatrix} D^\top Q(\mathcal{A}_c + \mathcal{F}D\mathcal{H}) & D^\top Q \end{bmatrix} + \begin{bmatrix} (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QD \\ QD \end{bmatrix} x \begin{bmatrix} \mathcal{K} & 0 \end{bmatrix} \prec 0 \tag{15}$$

for all x such that $x^\top Qx \le 1$. We show now that this is equivalent to the conclusion of the lemma. Define for compactness

$$\mathcal{P} := \mathcal{R} + (\mathcal{A}_c + \mathcal{F}D\mathcal{H})^\top QDx\mathcal{K} + \mathcal{K}^\top x^\top D^\top Q(\mathcal{A}_c + \mathcal{F}D\mathcal{H}),$$

so that (15) is equivalent, after some computations, to

$$\begin{bmatrix} \mathcal{P} & \mathcal{K}^\top x^\top D^\top Q \\ QDx\mathcal{K} & -Q \end{bmatrix} \prec 0.$$

By Schur's complement, we obtain that

$$\mathcal{P} + \mathcal{K}^\top x^\top D^\top QDx\mathcal{K} \prec 0 \quad \text{for all } x \text{ such that } x^\top Qx \le 1,$$

which is equivalent, by (8), to $\mathcal{N}_D(x) \prec 0$ for all $x \in \mathcal{E}_Q$. □

The next lemma addresses the term $\mathcal{F}D\mathcal{H}$ in $g_D(x)$ in (6). Specifically, it shows that as long as the matrix D is bounded in norm by δ as in Assumption 1, we can obtain a matrix inequality depending only on δ and guarantee that Lemma 3 and its conclusions hold for all such D, which is key to obtain a fully data-based solution to our problem.

Lemma 4. *Let Assumption 1 hold. If there exist* $\tau \in \mathbb{R}$, $\epsilon_2 \in \mathbb{R}$ *and* $Q = Q^\top \in \mathbb{R}^{n\times n}$ *such that*

$$\begin{bmatrix} -Q & 0 & \mathcal{K}^\top & \mathcal{A}_c^\top & \delta\mathcal{H}^\top \\ \star & -\tau Q & 0 & 0 & \delta I \\ \star & \star & -\frac{1}{\tau}I & 0 & 0 \\ \star & \star & \star & -Q^{-1} + \epsilon_2 I & 0 \\ \star & \star & \star & \star & -\epsilon_2 I \end{bmatrix} \prec 0, \tag{16}$$

then (13) *holds.*

Proof. Note that from $\mathcal{F} = I$ in (5), (13) is equivalent to

$$\begin{bmatrix} -Q & 0 & \mathcal{K}^\top & \mathcal{A}_c^\top \\ 0 & -\tau Q & 0 & 0 \\ \mathcal{K} & 0 & -\frac{1}{\tau}I & 0 \\ \mathcal{A}_c & 0 & 0 & -Q^{-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathcal{F} \end{bmatrix} \frac{D}{\delta} \begin{bmatrix} \delta\mathcal{H} & \delta I & 0 & 0 \end{bmatrix}$$
$$+ \begin{bmatrix} \delta\mathcal{H}^\top \\ \delta I \\ 0 \\ 0 \end{bmatrix} \frac{D^\top}{\delta} \begin{bmatrix} 0 & 0 & 0 & \mathcal{F}^\top \end{bmatrix} \prec 0$$

and this equation has the same structure as $\mathsf{G} + \mathsf{M}D\mathsf{N}^\top + \mathsf{N}D^\top \mathsf{M}^\top \prec 0$ in Fact 1, since $\|D\| \le \delta$ ($\delta > 0$) by Assumption 1. Indeed, by making the suitable correspondences between the quantities of this lemma and those of Fact 1, the existence of ϵ_2 such that (16) holds (corresponding to (12) of Fact 1) guarantees that (13) (corresponding to (11) of Fact 1) holds for D as in Assumption 1. □

Lemma 4 enables us to generalize the conclusions of Lemma 3 for all D with $\|D\| \le \delta$, so that we do not need to rely on the knowledge of D (as it would be the case in a model-based scheme), but just on its (possibly loose) norm bound δ. The matrix inequality (16) of Lemma 4 (where only δ appears), however, contains products of decision variables and inverses of decision variables. We address this in the next proposition, which obtains a matrix inequality that is as close as possible to an LMI (hence efficient to solve) and expresses explicitly the matrix inequality in terms of the available data. This proposition is the main result of this paper.

Proposition 1 (*Stabilization with Guaranteed Basin of Attraction*). *Under Assumption 1, suppose there exist* $\epsilon_1 \in \mathbb{R}$, $\epsilon_2 \in \mathbb{R}$, $Y \in \mathbb{R}^{n\times T}$ *and* $P = P^\top \in \mathbb{R}^{n\times n}$ *such that*

$$\begin{bmatrix} -P & 0 & YU_{0,T}^\top & YX_{1,T}^\top & -\delta YV_{0,T}^\top \\ \star & -\epsilon_1 P & 0 & 0 & \delta\epsilon_1 P \\ \star & \star & -\epsilon_1 I & 0 & 0 \\ \star & \star & \star & -P + \epsilon_2 I & 0 \\ \star & \star & \star & \star & -\epsilon_2 I \end{bmatrix} \prec 0 \tag{17a}$$

$$P = X_{0,T}Y^\top, \tag{17b}$$

and set $Q = P^{-1}$, $G_K = Y^\top P^{-1}$. *Then,* *(i) for the dynamics in (6) corresponding to D, the Lyapunov function* $V(x)=x^\top Qx = x^\top(X_{0,T}Y^\top)^{-1}x$ *satisfies*

$$V(g_D(x)x)-V(x)<0 \quad \text{for all } x\in\mathcal{E}_Q\setminus\{0\};$$

(ii) the origin is asymptotically stable for (1) with controller $u = Kx = U_{0,T}G_Kx = U_{0,T}Y^\top(X_{0,T}Y^\top)^{-1}x$ *and its basin of attraction contains the set* $\mathcal{E}_Q$.

Proof. We begin showing that inequalities (16) and (17a) are equivalent, noting for the following that (17a) implies $P\succ 0$. With the definitions in (5), (16) is equivalent to

$$\begin{bmatrix} -Q & 0 & G_K^\top U_{0,T}^\top & G_K^\top X_{1,T}^\top & -\delta G_K^\top V_{0,T}^\top \\ \star & -\tau Q & 0 & 0 & \delta I \\ \star & \star & -\frac{1}{\tau}I & 0 & 0 \\ \star & \star & \star & -Q^{-1}+\epsilon_2 I & 0 \\ \star & \star & \star & \star & -\epsilon_2 I \end{bmatrix} \prec 0.$$

By pre- and post-multiplying this inequality by the block diagonal matrix with entries Q^{-1}, Q^{-1}, I, I, I and by setting $Q=P^{-1}$, $G_K = Y^\top P^{-1}$ as in the statement of the proposition, the last inequality is equivalent to

$$\begin{bmatrix} -P & 0 & YU_{0,T}^\top & YX_{1,T}^\top & -\delta YV_{0,T}^\top \\ \star & -\tau P & 0 & 0 & \delta P \\ \star & \star & -\frac{1}{\tau}I & 0 & 0 \\ \star & \star & \star & -P+\epsilon_2 I & 0 \\ \star & \star & \star & \star & -\epsilon_2 I \end{bmatrix} \prec 0.$$

To avoid the simultaneous presence of τ and $1/\tau$, this inequality is equivalent to the next one by pre- and post-multiplying by the block diagonal matrix with entries I, $\frac{1}{\tau}I$, I, I, I and setting $\epsilon_1 = 1/\tau$:

$$\begin{bmatrix} -P & 0 & YU_{0,T}^\top & YX_{1,T}^\top & -\delta YV_{0,T}^\top \\ \star & -\epsilon_1 P & 0 & 0 & \delta\epsilon_1 P \\ \star & \star & -\epsilon_1 I & 0 & 0 \\ \star & \star & \star & -P+\epsilon_2 I & 0 \\ \star & \star & \star & \star & -\epsilon_2 I \end{bmatrix} \prec 0,$$

which is exactly (17a). After these manipulations, the conclusions of the proposition follow readily. Indeed, the fact that (17a) holds, implies that (16) holds, and then, by Lemmas 3 and 4, that D as in Assumption 1 satisfies $\mathcal{N}_D(x)\prec 0$ for all $x\in\mathcal{E}_Q$. By Lemma 2, (i) follows. (17b), which is equivalent to $I = X_{0,T}G_K$, and Lemma 1 ensure that (4) or, equivalently, (6) are an equivalent representation of (1) with controller $u=Kx=U_{0,T}G_Kx$. Standard Lyapunov theorems give then (ii). □

Proposition 1 effectively solves Problem 1. Indeed, if a solution to (17) is found (which is based on data from an off-line experiment), then we have a controller K and a guaranteed basin of attraction in terms of the set $\mathcal{E}_Q$.

The matrix inequality (17a) in Proposition 1 is convenient because, after fixing the scalar ϵ_1, it is an LMI in the decision variables ϵ_2, Y, P. A line search with respect to ϵ_1 on top of solving this LMI is typically preferable than solving directly the bilinear matrix inequality in (17a). Note that also model-based approaches for controlling bilinear systems encounter such a situation, and fix one of the parameters directly [29] or in an iterative way [27].

A conclusion of Proposition 1 is that the basin of attraction of the origin contains the set $\mathcal{E}_Q = \mathcal{E}_{P^{-1}}$. It is quite natural to maximize the volume of this ellipsoid, which is proportional to the square root of $\det(P)$, as is done in the model-based setting of [29]. (Other size criteria can be optimized, see the discussion in [39, §2.2.5.1].) This leads to the next immediate corollary.

Corollary 1 (*Ellipsoid Maximization*)**.** *Let Assumption 1 hold. If there exist a solution to the next optimization problem in the decision variables* $\epsilon_1\in\mathbb{R}$, $\epsilon_2\in\mathbb{R}$, $Y\in\mathbb{R}^{n\times T}$ *and* $P=P^\top\in\mathbb{R}^{n\times n}$

minimize $\quad -\log\det(P)$

subject to $\quad$ (17a), (17b),

then the conclusion of Proposition 1 holds.

Finally, since we are considering a quadratic Lyapunov function and as is done in the model-based solutions [27,29], the very same arguments leading to Proposition 1 yield exponential (instead of asymptotic) stability by strengthening a little the matrix inequality in (17a). This is stated in the next corollary, whose proof is thus omitted.

Corollary 2 (*Exponential Convergence*)**.** *For* $\mu\in(0,1)$, *suppose that the assumptions of Proposition 1 can be satisfied after replacing the element* $(1,1)$ *of the matrix in (17a) (i.e.,* $-P$*) with* $-\mu P$*. Then,*

(i) for the dynamics in (6) corresponding to D, the Lyapunov function $V(x)=x^\top Qx = x^\top(X_{0,T}Y^\top)^{-1}x$ *satisfies*

$$V(g_D(x)x) < \mu V(x) \quad \text{for all } x\in\mathcal{E}_Q\setminus\{0\};$$

(ii) the origin is exponentially stable for (1) with controller $u = Kx = U_{0,T}G_Kx = U_{0,T}Y^\top(X_{0,T}Y^\top)^{-1}x$ *and its basin of attraction contains the set* $\mathcal{E}_Q$.

3.1. Noisy data

In this section we show how the design in Proposition 1 can be made robust with noisy data. To this end, we consider that for all $t=0,\dots,T$, the state $x(t)$ is perturbed by the noise $n(t)$ resulting in a measured state $\tilde{x}(t)$, i.e.,

$$\tilde{x}(t) = x(t)+n(t).$$

In other words, we still consider (1) as underlying system and $u(0), u(1), \dots, u(T-1)$ as input sequence, but we can only rely on the noisy state sequence $\tilde{x}(0), \tilde{x}(1), \dots, \tilde{x}(T)$ for the design of the controller. Instead of (2), we then employ the noisy data

$$\begin{aligned} \tilde{X}_{0,T} &:= \begin{bmatrix}\tilde{x}(0) & \tilde{x}(1) & \dots & \tilde{x}(T-1)\end{bmatrix} \\ \tilde{X}_{1,T} &:= \begin{bmatrix}\tilde{x}(1) & \tilde{x}(2) & \dots & \tilde{x}(T)\end{bmatrix}, \\ \tilde{V}_{0,T} &:= \begin{bmatrix}\tilde{x}(0)u(0) & \tilde{x}(1)u(1) & \dots & \tilde{x}(T-1)u(T-1)\end{bmatrix} \end{aligned} \tag{18}$$

and define also the *unknown* quantities

$$\begin{aligned} N_{0,T} &:= \begin{bmatrix}n(0) & n(1) & \dots & n(T-1)\end{bmatrix} \\ N_{1,T} &:= \begin{bmatrix}n(1) & n(2) & \dots & n(T)\end{bmatrix} \\ W_{0,T} &:= \begin{bmatrix}n(0)u(0) & n(1)u(1) & \dots & n(T-1)u(T-1)\end{bmatrix}. \end{aligned} \tag{19}$$

By assuming now $I = \tilde{X}_{0,T}G_K$, we can reproduce the parametrization of Lemma 1 for the noisy data as

$$\begin{aligned} x^+ = (&\tilde{X}_{1,T} - N_{1,T} + AN_{0,T} + DW_{0,T} \\ &- D\tilde{V}_{0,T} + DxU_{0,T})G_Kx =: \tilde{g}_D(x)x, \end{aligned} \tag{20}$$

which boils down to (4) for $n(0)=\dots=n(T)=0$. With

$$\begin{aligned} &\hat{\mathcal{A}}_c := \tilde{X}_{1,T}G_K, \ \hat{\mathcal{F}} := I, \ \hat{\mathcal{H}} := -\tilde{V}_{0,T}G_K, \\ &\tilde{D} := -N_{1,T} + AN_{0,T} + DW_{0,T}, \ \tilde{\mathcal{F}} := I, \ \tilde{\mathcal{H}} := G_K, \end{aligned} \tag{21}$$

(20) can be written as

$$x^+ = (\hat{\mathcal{A}}_c + \tilde{\mathcal{F}}\tilde{D}\tilde{\mathcal{H}} + \hat{\mathcal{F}}D\hat{\mathcal{H}} + Dx\mathcal{K})x. \tag{22}$$

By comparison with (6), this expression reveals that noisy data result in an additional perturbation of the known $\hat{\mathcal{A}}_c$ through the unknown $\tilde{D}$ in (21). Similarly to D, we thus consider the next assumption for $\tilde{D}$.

Assumption 2. For some known $\tilde{\delta} > 0$, the matrix $\tilde{D}$ in (21) satisfies $\|\tilde{D}\| \le \tilde{\delta}$.

We give next a result with noisy data, which we then discuss together with Assumption 2.

Proposition 2 (*Stabilization with Guaranteed Basin of Attraction from Noisy Data*). *Under Assumptions 1 and 2, suppose there exist $\epsilon_1 \in \mathbb{R}$, $\epsilon_2 \in \mathbb{R}$, $\epsilon_3 \in \mathbb{R}$, $Y \in \mathbb{R}^{n\times T}$ and $P = P^\top \in \mathbb{R}^{n\times n}$ such that*

$$\begin{bmatrix} -P & 0 & YU_{0,T}^\top & Y\tilde{X}_{1,T}^\top & -\delta Y\tilde{V}_{0,T}^\top & \delta Y \\ \star & -\epsilon_1 P & 0 & 0 & \delta\epsilon_1 P & 0 \\ \star & \star & -\epsilon_1 I & 0 & 0 & 0 \\ \star & \star & \star & -P+(\epsilon_2+\epsilon_3)I & 0 & 0 \\ \star & \star & \star & \star & -\epsilon_2 I & 0 \\ \star & \star & \star & \star & \star & -\epsilon_3 I \end{bmatrix} \prec 0 \tag{23a}$$

$$P = \tilde{X}_{0,T} Y^\top, \tag{23b}$$

and set $Q = P^{-1}$, $G_K = Y^\top P^{-1}$. Then,

(i) for the dynamics in (20) corresponding to D and noisy data, the Lyapunov function $V(x) = x^\top Q x = x^\top(\tilde{X}_{0,T}Y^\top)^{-1}x$ satisfies

$$V(\tilde{g}_D(x)x) - V(x) < 0 \quad \text{for all } x \in \mathcal{E}_Q \setminus \{0\};$$

(ii) the origin is asymptotically stable for (1) with controller $u = Kx = U_{0,T}G_K x = U_{0,T}Y^\top(\tilde{X}_{0,T}Y^\top)^{-1}x$ and its basin of attraction contains the set $\mathcal{E}_Q$.

Proof. Due to space constraints and the close similarity to the developments leading to Proposition 1, we summarize only the key steps. Since the ideal data generated by (1) still satisfy $X_{1,T} = AX_{0,T} + BU_{0,T} + DV_{0,T}$, substituting in it (18) and (19) yields (20) for $I = \tilde{X}_{0,T}G_K$, as in Lemma 1. By clear correspondences between the matrices in (6) and (22), lemmas analogous to Lemmas 2 and 3 are obtained. By applying Fact 1, the term $\tilde{\mathcal{F}}\tilde{D}\tilde{\mathcal{H}}$ is addressed for the unknown $\tilde{D}$. Finally, the same steps as in the proof of Proposition 1 yield (23a), and (23b) is equivalent to $I = \tilde{X}_{0,T}G_K$. □

By comparing (17a) and (23a), one can see by continuity arguments that if (17a) is feasible then also (23a) is feasible provided that the noise has sufficiently small magnitude (corresponding to a sufficiently small $\tilde{\delta}$). This shows that our method is intrinsically robust to sufficiently small noise. On the other hand, Proposition 2 does not provide an *explicit quantification* of admissible signal-to-noise levels, as is done for instance in [17, §V-A] for linear systems. We believe that this analysis is possible also in this context and we leave it as future work.

We note that obtaining nonconservative values for $\tilde{\delta}$ clearly depends on the possibility of having nonconservative estimates on the noise level and on $\|A\|$. Upper bounds on $\|A\|$ can be obtained from the knowledge of a Lipschitz constant of the function $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \mapsto Ax + Bu + Dxu$ on compact sets, as we commented for Assumption 1.

4. Numerical example

We consider for (1) the matrices

$$A = \begin{bmatrix} 0.8 & 0.5 \\ 0.4 & 1.2 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, D = \begin{bmatrix} 0.45 & 0.45 \\ 0.3 & -0.3 \end{bmatrix}, \tag{24}$$

which are taken from [26, §5]. Our design does *not* rely on their knowledge, but simply on the data generated according to them and a bound δ of $\|D\|$. In particular, we consider $\delta = 0.7637$, which overapproximates by 20% the actual $\|D\| = 0.6364$ ($\delta/\|D\| = 1.2$), and we illustrate in Section 4.1 the effect of different δ. Moreover, we will use the matrices in (24) to compare

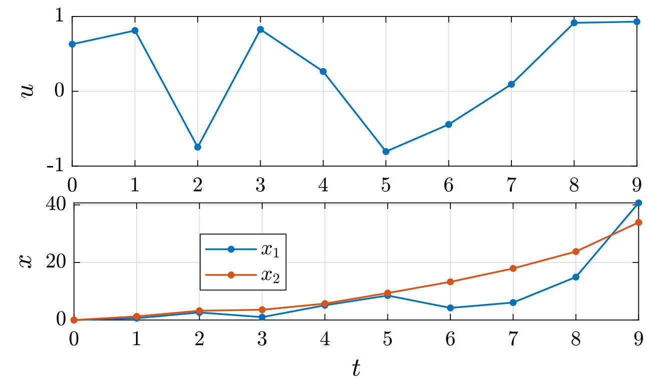

Fig. 1. Input and state sequences giving the quantities in (2).

our data-based design with a model-based design in Section 4.2. We note that the comparison is made with a model-based design that has perfect knowledge of the parameters of the system. Getting to perfectly know the parameters would correspond to the ideal case even for a preliminary system identification step. We show in this section that our designed controller performs comparably to such a model-based design, in spite of being tuned only on an offline experiment.

We consider $T = 10$. In Fig. 1, we show the input and state sequences giving (2) and generated according to the matrices in (24). We note that A being unstable is challenging because a suitable control action has to be designed to modify by feedback the system evolution in a neighborhood of the origin (without the "help" of a stable linear part), and the diverging data pose a practical limit on the length of the open loop experiment, besides possibly impacting the numerical accuracy of the procedure.

4.1. Data-based solution

In the following implementation, we present Corollary 1 because the size criterion of the determinant allows quantitative comparisons (as opposed to Proposition 1), and (as opposed to Corollary 2) the benefits of guaranteed exponential convergence are outweighed by the reduction of the size of the guaranteed basin of attraction in the present example, despite the theoretical interest of Corollary 2. By using Corollary 1, the data-based solution implemented in this section is as follows, where we opt for fixing the scalar variable ϵ_1, solve an LMI, and perform a line search on ϵ_1.

1. We fix $\epsilon_1 > 0$.
2. We solve the next optimization problem in the decision variables $\epsilon_2 \in \mathbb{R}$, $Y \in \mathbb{R}^{n\times T}$ and $P = P^\top \in \mathbb{R}^{n\times n}$

   minimize $-\log\det(P)$

   subject to (17a), (17b)

   which corresponds to an LMI. By denoting the solution $P =: P_{\mathrm{DB}}$, we then obtain $G_K = Y^\top P_{\mathrm{DB}}^{-1}$ and the controller gain as $K_{\mathrm{DB}} := U_{0,T}G_K$.
3. We iterate on the selection of ϵ_1 in case of, e.g., infeasibility.

We implement this scheme (and the others in this section) through the toolbox YALMIP [40] and the solver MOSEK. For a value of $\epsilon_1 = 0.8$, we obtain

$$P_{\mathrm{DB}} = \begin{bmatrix} 3.2827 & -0.9642 \\ -0.9642 & 2.4388 \end{bmatrix}, \ K_{\mathrm{DB}} = \begin{bmatrix} -0.3175 & -0.5649 \end{bmatrix}.$$

The evolution of x when $u = K_{\mathrm{DB}}x$ is used in (1) is given in Fig. 2 in the top plot as a phase portrait (*solid* colored lines) and in the middle plot as a time evolution.

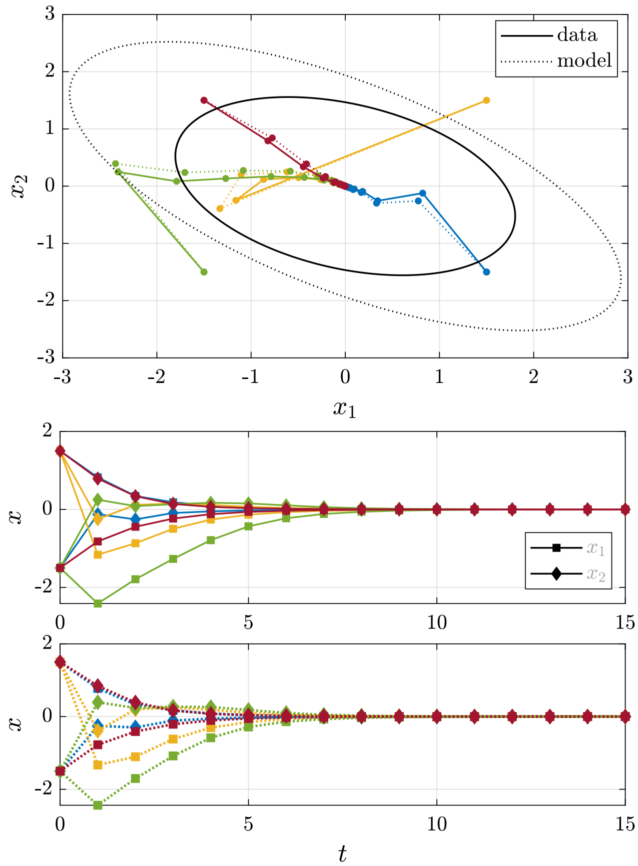

Fig. 2. Evolution of the data-based and model-based solutions of Sections 4.1 and 4.2, corresponding to the selected value of ϵ_1. The same color corresponds to solutions with the same initial condition. Solid and dotted lines correspond respectively to the data-based and model-based solutions. (Top) Phase portrait. The area within the ellipsoids is guaranteed to be in the basin of attraction of the origin, by the existence of the Lyapunov functions corresponding to the matrices P_{DB} and P_{MB}. (Middle) Time evolutions of the state x for the data-based solution, where traces with squares and diamonds identify respectively the components x_1 and x_2. (Bottom) Time evolutions of the state x for the model-based solution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

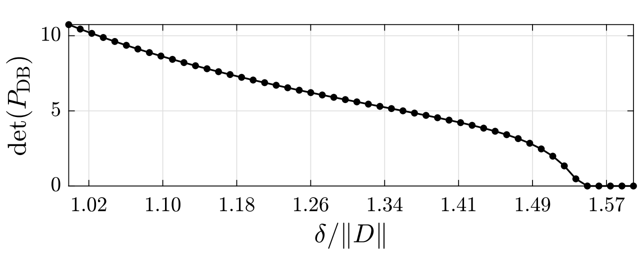

Fig. 3. Effect of δ on the size of the guaranteed basin of attraction.

We illustrate the effect of different bounds δ on $\|D\|$ using the same parameters and data as before, and report the corresponding $\det(P_{\mathrm{DB}})$ in Fig. 3. The guaranteed basin of attraction shrinks when $\delta/\|D\|$ increases, which is the price to pay for not knowing D and having only an upper bound on its norm. However, the figure shows that this deterioration is tolerable for $\delta/\|D\|$ as loose as 1.4.

Remark 4 (*Practical Considerations on Number of Data and Computation Times*)**.** Our method manages to provide stability guarantees with small datasets. This feature is very appealing in contexts like the one just considered where the system is open-loop unstable and collecting large datasets can be problematic. The method, however, can handle datasets of larger size, which can instead be convenient when the system to control is open-loop stable or mildly unstable. We exemplify this point by considering the same parameters as before, except for substituting A with $A/1.485$, which is only mildly unstable. We generate the data in (2) according to the new A, and consider different values of T. For each of them, we measure the wall-clock time for solving the previous optimization problem using the MATLAB® function `timeit` (MATLAB® R2018a on a machine with processor Intel® Core™ i7 with 4 cores and 1.80 GHz). The resulting wall-clock times in Fig. 4 from $T = 10$ up to $T = 1000$ data points, are at most 1.2 s.

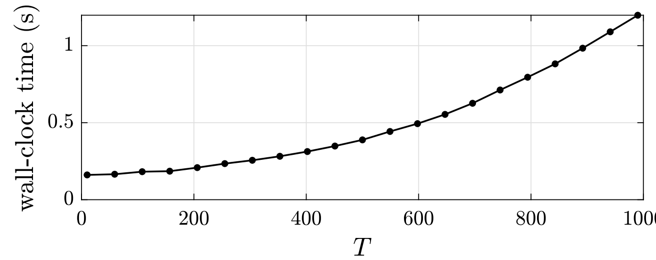

Fig. 4. Execution times as a function of the number T of data.

4.2. Model-based solution

For (1) with matrices in (24), we use the model-based solution in [29] for comparison. This model-based solution is also not an LMI, unless the scalar parameter ϵ_1 is fixed (as in the data-based solution) and a line search is performed.

1. We fix $\epsilon_1 > 0$.
2. We solve the next optimization problem in the decision variables $y \in \mathbb{R}^n$ and $P = P^\top \in \mathbb{R}^{n\times n}$

$$\begin{aligned}&\text{minimize } -\log\det(P)\\&\text{subject to } P \succ 0\\&\begin{bmatrix} -P & 0 & y & PA^\top + yB^\top\\ 0 & -\epsilon_1 P & 0 & PD^\top\\ y^\top & 0 & -\epsilon_1 I & 0\\ AP + By^\top & DP & 0 & -P\end{bmatrix} \prec 0,\end{aligned}$$

which corresponds to an LMI. By denoting the solution $P =: P_{\mathrm{MB}}$, we then obtain the controller gain as $K_{\mathrm{MB}} := y^\top P_{\mathrm{MB}}^{-1}$.
3. We iterate on the selection of ϵ_1 in case of, e.g., infeasibility.

For $\epsilon_1 = 0.8$ as in Section 4.1, we obtain

$$P_{\mathrm{MB}} = \begin{bmatrix} 8.5623 & -4.7253\\ -4.7253 & 6.3616\end{bmatrix},\ K_{\mathrm{MB}} = \begin{bmatrix} -0.3572 & -0.5738\end{bmatrix}.$$

The evolution of x when $u = K_{\mathrm{MB}}x$ is used in (1) is given in Fig. 2 in the top plot as a phase portrait (*dotted* colored lines) and in the bottom plot as a time evolution.

4.3. Comparison of data-based and model-based solutions

We compare the performance of the data-based solution against the model-based solution by performing a thorough line search on the parameter ϵ_1, which we fixed before in order to be able to solve an LMI. The result is in Fig. 5. Only values of ϵ_1 where an optimal solution was returned by YALMIP, are displayed

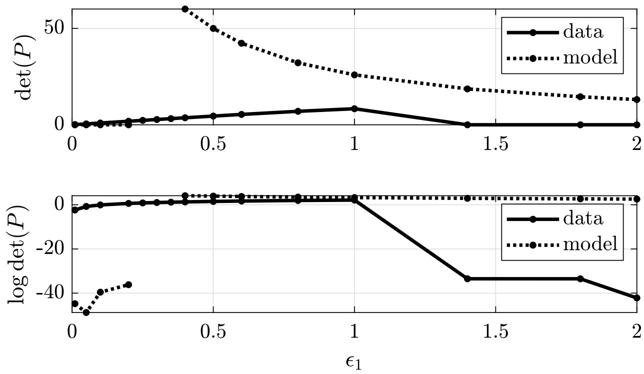

Fig. 5. Characterization of the determinants of matrices $P_{\rm DB}$ and $P_{\rm MB}$ (top) and their logarithms (bottom) as a function of the parameter ϵ_1.

(in particular, this did not happen for the model-based solution with values of ϵ_1 between 0.2 and 0.4).

The top plot represents the determinants of the matrices $P_{\rm DB}$ and $P_{\rm MB}$, which was considered since its square root is proportional to the volume of the ellipsoids guaranteed to be in the basin of attraction of the closed-loop system. In the bottom plot, the logarithms of these determinants are also provided since they are the actual objective functions in the optimization problems of Sections 4.1–4.2.

As expected, the model-based solution provides ellipsoids with larger sizes (e.g., $\det(P_{\rm MB}) = 60.03$ for $\epsilon_1 = 0.4$). For the given example, it appears from Fig. 5 that the data-based solution performs better for small ϵ_1, whereas it performs worse than the model-based solution for large ϵ_1. We note that log det is actually more representative of the actual difference between the two solutions. Indeed, for values of ϵ_1 around 1, the two solutions are not so distant, as is confirmed by the illustration of Fig. 2 where the corresponding ellipsoids are also depicted in the top plot (solid and dotted black curves).

In summary, our designed controller presents in these simulations a similar performance to the model-based design, where the former relies on an offline experiment and the latter on the perfect knowledge of system parameters.

4.4. Data-based solution with noisy data

Finally, we briefly illustrate the robust design of Proposition 2 in the presence noisy data. As in Section 4.1, we take an input signal (uniformly) distributed in $[-1, 1]$ but we now assume that the data are corrupted by a measurement noise with all components (uniformly) distributed in $[-\bar{n}, \bar{n}]$. For $\tilde{D}$ in (21), we have

$$\|\tilde{D}\| \le \|N_{1,T}\| + \|A\|\|N_{0,T}\| + \|D\|\|W_{0,T}\|. \quad (25)$$

Accordingly, we overapproximate $\|\tilde{D}\|$ using

$$\tilde{\delta} := \bar{n}\left(\sqrt{nT} + \alpha\sqrt{nT} + \delta\sqrt{n}\|U_{0,T}\|\right) \quad (26)$$

where the scalar n is the system order ($n = 2$ in this example), T is the number of samples, α is an upper bound on $\|A\|$ and δ is the upper bound on $\|D\|$. Here, we consider $\delta = 0.7637$ and $T = 10$ as in Section 4.1, and $\alpha = 2.9874$, which overapproximates by 100% the actual $\|A\| = 1.4937$. We solve (23) still using $\epsilon_1 = 0.8$ for different values of $\bar{n}$. Fig. 6 reports the behavior of $\det(P_{\rm DB})$ as a function of $\bar{n}$.

As discussed in Section 3.1, the numerical results show that the performance is close to the ideal one for small values of noise despite the overapproximation on $\|\tilde{D}\|$.

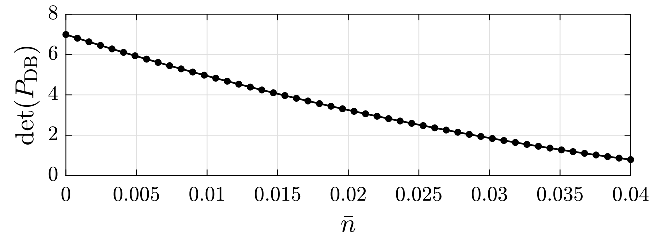

Fig. 6. Effect of the amplitude of noise on the size of the guaranteed basin of attraction.

5. Conclusions

We proposed a direct data-driven design for bilinear systems, which comes with a guaranteed subset of the basin of attraction. This design is best suited for a limited number of open-loop data, and numerical experiments show its applicability for a large number of data. As a proof of concept, we show how to make the design robust in the presence of noisy data.

The main goal of future work is applying this scheme as a building block for data-driven control of input-affine nonlinear systems (by approximating the latter through Carleman linearization). A closely related topic of future work is a study of the tradeoffs with schemes based on sum-of-squares programming for bilinear systems.

CRediT authorship contribution statement

Andrea Bisoffi: Formal analysis, Software, Writing - original draft. **Claudio De Persis:** Formal analysis, Writing - review & editing, Supervision. **Pietro Tesi:** Formal analysis, Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

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

[2] A. Karimi, L. Mišković, D. Bonvin, Iterative correlation-based controller tuning, Int. J. Adapt. Control Signal Process. 18 (8) (2004) 645–664.

[3] R. Gonçalves da Silva, A. Bazanella, C. Lorenzini, L. Campestrini, Data-driven LQR control design, IEEE Control Syst. Lett. 3 (1) (2019) 180–185.

[4] G. Baggio, V. Katewa, F. Pasqualetti, Data-driven minimum-energy controls for linear systems, IEEE Control Syst. Lett. 3 (3) (2019) 589–594.

[5] B. Recht, A tour of reinforcement learning: The view from continuous control, Ann. Rev. Control Robot. Auton. Syst. 2 (2019) 253–279.

[6] C. De Persis, P. Tesi, On persistency of excitation and formulas for data-driven control, in: Proc. IEEE Conf. Decis. Control, 2019, pp. 873–878.

[7] Z.-S. Hou, Z. Wang, From model-based control to data-driven control: Survey, classification and perspective, Inform. Sci. 235 (2013) 3–35.

[8] M. Bojarski, et al., End to end learning for self-driving cars, 2016, arXiv preprint arXiv:1604.07316.

[9] M. Fliess, Model-free control and intelligent PID controllers: Towards a possible trivialization of nonlinear control? IFAC Proc. Vol. 42 (10) (2009) 1531–1550.

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

[11] P. Tabuada, W.-L. Ma, J. Grizzle, A.D. Ames, Data-driven control for feedback linearizable single-input systems, in: Proc. IEEE Conf. Decis. Control, 2017, pp. 6265–6270.

[12] K.P. Wabersich, M.N. Zeilinger, Scalable synthesis of safety certificates from data with application to learning-based control, in: Proc. Eur. Control Conf., 2018, pp. 1691–1697.

[13] J. Coulson, J. Lygeros, F. Dörfler, Data-enabled predictive control: In the shallows of the DeePC, in: Proc. Eur. Contr. Conf., 2019, pp. 307–312.
[14] F. Berkenkamp, M. Turchetta, A.P. Schoellig, A. Krause, Safe model-based reinforcement learning with stability guarantees, in: Proc. Advances in Neural Information Processing Systems, Vol. 30, 2017, pp. 908–918.
[15] J.C. Willems, P. Rapisarda, I. Markovsky, B.L.M. De Moor, A note on persistency of excitation, Systems Control Lett. 54 (4) (2005) 325–329.
[16] L. Huang, J. Coulson, J. Lygeros, F. Dörfler, Data-enabled predictive control for grid-connected power converters, in: Proc. IEEE Conf. Decis. Control, 2019, pp. 8130–8135.
[17] C. De Persis, P. Tesi, Formulas for data-driven control: Stabilization, optimality and robustness, IEEE Trans. Automat. Control 65 (3) (2020) 909–924.
[18] J. Berberich, A. Romer, C.W. Scherer, F. Allgöwer, Robust data-driven state-feedback design, in: Proc. Amer. Control Conf., 2020, pp. 1532–1538.
[19] J.G. Rueda-Escobedo, J. Schiffer, Data-driven internal model control of second-order discrete Volterra systems, 2020, arXiv preprint arXiv:2003.14158.
[20] M. Guo, C. De Persis, P. Tesi, Learning control for polynomial systems using sum of squares, 2020, arXiv preprint arXiv:2004.00850, to appear in: Proc. IEEE Conf. Decis. Control, 2020.
[21] R.W. Brockett, Volterra series and geometric control theory, Automatica 12 (2) (1976) 167–176.
[22] A.J. Krener, Bilinear and nonlinear realizations of input-output maps, SIAM J. Control 13 (4) (1975) 827–834.
[23] W.J. Rugh, Nonlinear System Theory, Johns Hopkins University Press, 1981, Web version prepared by the author in 2002.
[24] R.R. Mohler, Bilinear Control Processes: with Applications to Engineering, Ecology and Medicine, Academic Press, 1973.
[25] C. Bruni, G. DiPillo, G. Koch, Bilinear systems: An appealing class of "nearly linear" systems in theory and applications, IEEE Trans. Automat. Control 19 (4) (1974) 334–348.
[26] G. Bitsoris, N. Athanasopoulos, Constrained stabilization of bilinear discrete-time systems using polyhedral Lyapunov functions, IFAC Proc. Vol. 41 (2) (2008) 2502–2507, 17th IFAC World Congress.
[27] S. Tarbouriech, I. Queinnec, T.R. Calliero, P.L.D. Peres, Control design for bilinear systems with a guaranteed region of stability: An LMI-based approach, in: Med. Conf. Control Automation, 2009, pp. 809–814.
[28] F. Amato, C. Cosentino, A.S. Fiorillo, A. Merola, Stabilization of bilinear systems via linear state-feedback control, IEEE Trans. Circ. Syst. II: Express Briefs 56 (1) (2009) 76–80.
[29] M.V. Khlebnikov, Quadratic stabilization of discrete-time bilinear systems, Autom. Remote Control 79 (7) (2018) 1222–1239.
[30] W. Favoreel, B. De Moor, P. Van Overschee, Subspace identification of bilinear systems subject to white inputs, IEEE Trans. Automat. Control 44 (6) (1999) 1157–1165.
[31] H. Chen, J. Maciejowski, Subspace identification of deterministic bilinear systems, in: Proc. Amer. Control Conf., 2000, pp. 1797–1801.
[32] E.D. Sontag, Y. Wang, A. Megretski, Input classes for identifiability of bilinear systems, IEEE Trans. Automat. Control 54 (2) (2009) 195–207.
[33] V. Verdult, Nonlinear System Identification: a State-Space Approach (Ph.D. thesis), University of Twente, 2002.
[34] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory, Vol. 15, SIAM, 1994.
[35] I.R. Petersen, A stabilization algorithm for a class of uncertain linear systems, Systems Control Lett. 8 (4) (1987) 351–357.
[36] M. Vatani, M. Hovd, S. Olaru, Control design for discrete-time bilinear systems using the scalarized Schur complement, Internat. J. Robust Nonlinear Control 27 (18) (2017) 4492–4506.
[37] C. De Persis, P. Tesi, Designing experiments for data-driven control of nonlinear systems, in: 24th Int. Symp. Mathematical Theory of Networks and Systems, 2020, accepted for publication.
[38] R.A. Horn, C.R. Johnson, Matrix analysis, second ed., Cambridge University Press, 2013.
[39] S. Tarbouriech, G. Garcia, J.M. Gomes da Silva Jr., I. Queinnec, Stability and Stabilization of Linear Systems with Saturating Actuators, Springer, 2011.
[40] J. Löfberg, YALMIP: A toolbox for modeling and optimization in MATLAB, in: Proc. IEEE Int. Symp. Computer Aided Control System Design, 2004, pp. 284–289.