Low-complexity learning of Linear Quadratic Regulators from noisy data

De Persis, Claudio; Tesi, Pietro

Published in:
Automatica

DOI:
10.1016/j.automatica.2021.109548

Automatica 128 (2021) 109548

Contents lists available at ScienceDirect

Automatica

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

Low-complexity learning of Linear Quadratic Regulators from noisy data[☆]

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

[a] ENTEG, Faculty of Science and Engineering, University of Groningen, 9747 AG Groningen, The Netherlands
[b] Department of Information Engineering (DINFO), University of Florence, 50139 Florence, Italy

ARTICLE INFO

Article history:
Received 3 May 2020
Received in revised form 21 January 2021
Accepted 28 January 2021
Available online 22 March 2021

ABSTRACT

This paper considers the Linear Quadratic Regulator problem for linear systems with unknown dynamics, a central problem in data-driven control and reinforcement learning. We propose a method that uses data to directly return a controller without estimating a model of the system. Sufficient conditions are given under which this method returns a stabilizing controller with guaranteed relative error when the data used to design the controller are affected by noise. This method has low complexity as it only requires a finite number of samples of the system response to a sufficiently exciting input, and can be efficiently implemented as a semi-definite programme.

1. Introduction

Control theory is witnessing an increasing renewed interest towards *data-driven* (*data-based*) control. This terminology refers to all those cases where the dynamics of the system are unknown and the control law must be designed using data alone. This can be done either by identifying a model of the system from data and then use the model for control design, or by directly designing the control law bypassing the system identification (ID) step. Methods in the first category are usually called *indirect* (sequential system ID and control design), while methods in the second category are usually called *direct* or *model-free*.

The interest for data-driven control has several motivations. As systems become more complex, first-principle models may be difficult to obtain or may be too complex for control design. Fully automated (end-to-end) procedures may also facilitate the online tuning or re-design of controllers, which is needed in all those applications where the system to be controlled or the environment are subject to changes that are difficult to predict. Dozens of publications on data-driven control have appeared in the last few years. We mention works on predictive control (Alpago, Lygeros, & Dörfler, 2020; Coulson, Lygeros, & Dörfler, 2019; Salvador, Muñoz de la Peña, Alamo, & Bemporad, 2018), optimal control (Baggio, Katewa, & Pasqualetti, 2019; De Persis & Tesi, 2019; Gonçalves da Silva, Bazanella, Lorenzini, & Campestrini, 2019; Recht, 2019), robust and nonlinear control (Berberich, Koch, Scherer, & Allgöwer, 2019; Dai & Sznaier, 2018; De Persis & Tesi, 2020a; Novara, Formentin, Savaresi, & Milanese, 2016; Wabersich & Zeilinger, 2018). We refer the interested reader to Hou and Wang (2013) for a survey on earlier contributions.

The Linear Quadratic Regulator problem

This paper considers the *infinite horizon* Linear Quadratic Regulator (LQR) problem for linear time-invariant systems. Besides its relevance, this problem is a prime example of the challenges encountered in data-driven control. Specifically, we consider the problem of determining the solution to the LQR problem from a finite set of (noisy) data collected from the system.

Early data-driven methods for LQR originate from adaptive control, examples being *self-tuning regulators* (Åström & Wittenmark, 1989) and *policy iteration* schemes (Bradtke, Ydstie, & Barto, 1994). While the specific methods are different, the common idea is to study the convergence of an adaptive law to the optimal one as time goes to infinity. Starting from Fiechter (1997), much effort has been made to derive *non-asymptotic* properties of data-driven methods, that is for a *finite* number of steps. The interest towards non-asymptotic properties is both theoretical and practical: they help to derive performance guarantees for iterative (online) methods (Fazel, Ge, Kakade, & Mesbahi, 2018), and are the basis of non-iterative (offline) methods (De Persis & Tesi, 2019; Recht, 2019) which use only a finite number of data points.

It turns out that non-asymptotic properties are very difficult to derive if one departs from the assumption that the data are noise-free. Most of the works dealing with noisy data are of indirect type. The idea is to estimate the parameters of the system along with sample guarantees on the estimation error and then design or update the control law according to the estimate. Early

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Tongwen Chen under the direction of Editor Ian R. Petersen.

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

https://doi.org/10.1016/j.automatica.2021.109548
0005-1098/

results on the finite sample properties of system ID methods can be traced back to Campi and Weyer (2002), but it is only recently that these ideas (with different estimation methods) have been considered for the LQR problem (Cohen, Koren, & Mansour, 2019; Dean, Mania, Matni, Recht, & Tu, 2020; Mania, Tu, & Recht, 2019). Direct methods have been instead much less explored (Abbasi-Yadkori, Lazíc, & Szepesvári, 2018).

Contribution and outline of the paper

Our contribution is a new approach to design LQ controllers from noisy data. The method builds on the framework of De Persis and Tesi (2020a) and has the following features:

(i) *Low complexity*. The proposed method requires a finite (pre-computable) number of data points obtained from a single or multiple system's trajectories, and it can be implemented as a convex program.
(ii) *Stability and performance guarantees*. As long as the noise satisfies suitable inequalities our method returns a stabilizing controller with guarantees on the *relative error* (the gap between the solution and the unknown optimal controller).
(iii) *No assumptions on statistical properties of noise*. We do not make assumptions regarding the noise statistics such as the noise being a martingale or white.

As in Dean et al. (2020), Mania et al. (2019) and Recht (2019), we focus on non-iterative methods. The main difference with respect to the existing works is that our method is direct and makes no assumptions on the noise statistics. The advantage of not relying on noise statistics is twofold. Although the solution to LQR can be viewed as the one minimizing the variance of the system output in response to white Gaussian noise, experimental data need not comply with such setting, and show correlation and dependence. Our method is free from this issue, but it can also be specialized to such noise. Not relying on noise statistics also enables us to extend the analysis to the stabilization of nonlinear systems around an equilibrium point since, around an equilibrium, a nonlinear system can be expressed via its first-order approximation plus a remainder acting as a noise source. We will elaborate on these points in the paper.

Our method is direct (model-free). Indirect methods show excellent performance (Dean et al., 2020; Mania et al., 2019) and they can capitalize on the many analysis and design tools available for model-based control. On the other hand, although much less explored, also direct methods have their own strengths. First, while generalizations are intuitively possible, indirect methods for data-driven LQR are currently restricted to special types of noise. Our method overcomes this limitation. Further, direct methods seem to be applicable in a more straightforward way to settings where the ID step is usually involved such as with nonlinear and time-varying systems. Examples in this direction, not covering LQR, are Dai and Sznaier (2018), Guo, De Persis, and Tesi (2020) and Wabersich and Zeilinger (2018).

Our method rests on a key result by Willems and co-authors (Willems, Rapisarda, Markovsky, & De Moor, 2005) which is recalled in Section 2. This result states that a (noise-free) system trajectory generated by a *persistently exciting* input is an equivalent system representation. We exploit this result to develop our model-free method. In Section 3, we cast the LQR as an $\mathcal{H}_2$ problem (Scherer & Weiland, 2019) and derive a data-based solution based on convex programming for the ideal case of noise-free data (Theorem 1). The main results are given throughout Sections 4 and 5 . The first one (Theorem 2) provides stability properties and error bounds of the baseline solution in case of noisy data. One variant to the baseline solution is discussed in Theorem 3. This variant ensures more noise tolerance at the cost of possibly reduced performance bounds. This matches what has been observed also in indirect methods (Mania et al., 2019). The results are discussed in Section 5.2. Section 6 extends the analysis to nonlinear systems and de-noising strategies. In Section 7 we provide some numerical simulations and Section 8 gives concluding remarks.

2. Notation and auxiliary facts

We denote by $\mathbb{N}_0$ and $\mathbb{N}_1$ non-negative and positive integers, respectively. Given a signal $z : \mathbb{N}_0 \to \mathbb{R}^\sigma$ and two integers $k, r \in \mathbb{N}_0$ with $r \geq k$ we let $z_{[k,r]} := \{z(k), \dots, z(r)\}$. Given a signal z and a positive integer T, we also define $Z_i := [z(i)\ z(i+1)\ \dots\ z(T+i-1)]$. As we will always consider experiments of length T, we omit the dependence of Z_i on T. Finally, the prime denotes transpose.

Consider a linear time-invariant system

$$x(k+1) = Ax(k) + Bu(k) \quad k \in \mathbb{N} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state and $u \in \mathbb{R}^m$ is the control input, and suppose that we have access to T-long data sequences $u_{[0,T-1]}$ and $x_{[0,T-1]}$ of (1). In this paper, the following condition plays a key role:

$$\operatorname{rank} W_0 = n + m \tag{2}$$

where

$$W_0 := \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} \tag{3}$$

Condition (2) ensures that any T-long input-state trajectory of the system can be expressed as a linear combination of the columns of W_0, meaning that W_0 encodes full information on the dynamics of the system. A key property established in Willems et al. (2005) is that one can ensure (2) when the input is sufficiently *exciting*.

Definition 1 (*Willems et al., 2005*)**.** A signal $z_{[0,T-1]} \in \mathbb{R}^\sigma$ is said to be *persistently exciting* of order $s \in \mathbb{N}_1$ if

$$\mathscr{Z}_0 := \begin{bmatrix} z(0) & z(1) & \cdots & z(T-s) \\ z(1) & z(2) & \cdots & z(T-s+1) \\ \vdots & \vdots & \ddots & \vdots \\ z(s-1) & z(s) & \cdots & z(T-1) \end{bmatrix}$$

has full rank σs. ■

Lemma 1 (*Willems et al., 2005, Corollary 2*)**.** *Suppose that system (1) is controllable. If $u_{[0,T-1]}$ is persistently exciting of order $n+1$, then condition (2) holds.* ■

3. Problem definition and data-driven formulation

In this section, we introduce the problem of interest and our baseline direct (model-free) data-driven method, which rests on condition (2).

3.1. The linear quadratic regulator problem

Consider a linear time-invariant system

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + d(k) \\ z(k) = \begin{bmatrix} W_x^{1/2} & 0 \\ 0 & W_u^{1/2} \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \end{cases} \tag{4}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, and where d is a disturbance term; z is a performance signal of interest; (A, B) is controllable $W_x, W_u \succ 0$ are weighting matrices. Controllability is actually not needed for the LQR problem, but we assume it in view of Lemma 1.

We consider the problem of designing a state-feedback controller K that renders $A+BK$ Schur and minimizes the $\mathcal{H}_2$-norm of the transfer function $\mathscr{T}: d \to z$ of the closed-loop system

$$\begin{bmatrix} x(k+1) \\ z(k) \end{bmatrix} = \left[\begin{array}{c|c} A+BK & I \\ \hline \begin{bmatrix} W_x^{1/2} \\ W_u^{1/2}K \end{bmatrix} & 0 \end{array}\right] \begin{bmatrix} x(k) \\ d(k) \end{bmatrix} \tag{5}$$

where (Chen & Francis, 1995, Section 4.4)

$$\|\mathscr{T}\|_2 := \left[\frac{1}{2\pi}\int_0^{2\pi} \mathrm{tr}\left(\mathscr{T}\left(e^{j\theta}\right)' \mathscr{T}\left(e^{j\theta}\right)\right) d\theta\right]^{\frac{1}{2}} \tag{6}$$

In the sequel, we will write $\mathscr{T} = \mathscr{T}(K)$ so as to emphasize the dependence of $\mathscr{T}$ on K. It is known (Chen & Francis, 1995, Section 4.4) that when $A+BK$ is Schur

$$\|\mathscr{T}(K)\|_2^2 = \mathrm{tr}\,(W_xP) + \mathrm{tr}\left(W_uKPK'\right) \tag{7}$$

where P is the controllability Gramian of the closed-loop system (5), which coincides with the unique solution to the Lyapunov equation $(A+BK)P(A+BK)'-P+I=0$. The $\mathcal{H}_2$-norm corresponds in the time domain to the 2-norm of the output z when impulses are applied to the input channels, and it can be interpreted as the mean-square deviation of z when d is a white process with unit covariance, which is the classic stochastic LQR formulation. Here, we view the LQR problem as a $\mathcal{H}_2$-norm minimization problem as our method is based on the minimization of (7).

As shown in Chen and Francis (1995, Section 6.4) the state-feedback controller that minimizes the $\mathcal{H}_2$-norm of $\mathscr{T}(K)$ is unique and can be computed as

$$K_\star = -(W_u + B'XB)^{-1}B'XA \tag{8}$$

where X is the unique positive definite solution to the classic discrete-time algebraic Riccati (DARE) equation

$$A'XA - X - (A'XB)(W_u + B'XB)^{-1} \times (B'XA) + W_x = 0$$

We are interested in computing $K_\star$ when a model of the system is not available, and we have only access to a T-long stream of (noisy) data $u_{[0,T-1]}$ and $x_{[0,T-1]}$ collected during some experiment on system (4) where by *noisy* we mean that the data collected from (4) might have been generated with *nonzero* disturbance d. We aim at establishing properties of the data-driven solution with respect to the one that we can compute under exact model knowledge.

3.2. A data-driven SDP formulation

The problem of finding $K_\star$ can be equivalently formulated as a *semi-definite programme* (SDP):[1]

$$\min_{(\gamma,K,P,L)} \gamma$$

subject to

$$\begin{cases} (A+BK)P(A+BK)' - P + I \preceq 0 \\ P \succeq I \\ L - KPK' \succeq 0 \\ \mathrm{tr}\,(W_xP) + \mathrm{tr}\,(W_uL) \le \gamma \end{cases} \tag{9}$$

This formulation is the natural discrete-time counterpart of the formulation proposed in Feron, Balakrishnan, Boyd, and El Ghaoui (1992) for continuous-time systems. We will not discuss the properties associated to (9). Rather, we will discuss the properties associated to an *equivalent data-based* version of (9).

Consider system (4) along with data sequences $d_{[0,T-1]}$, $u_{[0,T-1]}$ and $x_{[0,T]}$ resulting from an experiment of length T. Define corresponding matrices D_0, U_0, X_0 and X_1, which satisfy the relation

$$X_1 = AX_0 + BU_0 + D_0 \tag{10}$$

It turns out that $K_\star$ can be parametrized directly in terms of the data matrices D_0, U_0, X_0 and X_1. Specifically, as shown in next Theorem 1, under condition (2) the controller $K_\star$ can be expressed as $K_\star = U_0Q_\star P_\star^{-1}$ where $(\gamma_\star, Q_\star, P_\star, L_\star)$ is any optimal solution to the SDP:

$$\min_{(\gamma,Q,P,L)} \gamma$$

subject to

$$\begin{cases} (X_1 - D_0)QP^{-1}Q'(X_1 - D_0)' - P + I \preceq 0 \\ P \succeq I \\ L - U_0QP^{-1}Q'U_0' \succeq 0 \\ X_0Q = P \\ \mathrm{tr}\,(W_xP) + \mathrm{tr}\,(W_uL) \le \gamma \end{cases} \tag{11}$$

which only depends on data. The idea behind (11) is that, under condition (2), any feedback interconnection $A+BK$ can be written in a form that does not involve the matrices A and B. In fact, under condition (2), for any K there exists a matrix G that solves the system of equations

$$\begin{bmatrix} K \\ I \end{bmatrix} = W_0G \tag{12}$$

which implies

$$A+BK = [B \quad A]W_0G = (X_1 - D_0)G \tag{13}$$

Under (2), the SDP (11) thus coincides with the one in (9) with $Q = GP$. In particular, $K = U_0QP^{-1}$ and $X_0Q = P$ give an equivalent characterization of the two constraints in (12). As noted in van Waarde, Eising, Trentelman, and Camlibel (2020), condition (2) is not restrictive for the LQR problem. In fact, it is necessary for reconstructing A and B from data (thus for indirect methods) and is generically needed for data-driven methods in general.

The SDP (11) first appeared in De Persis and Tesi (2020a) under the assumption that the collected data are noise-free, that is with $D_0 = 0$, in which case (11) can be solved. Here we revisit this result providing some additional properties related to this formulation.

Theorem 1. *Suppose that condition (2) is satisfied. Then problem (11) is feasible. Moreover, any optimal solution $(\gamma_\star, Q_\star, P_\star, L_\star)$ is such that $K_\star = U_0Q_\star P_\star^{-1}$ and*

$$\|\mathscr{T}(K_\star)\|_2^2 = \mathrm{tr}(W_xP_\star) + \mathrm{tr}(W_uL_\star) \tag{14}$$

The proof of Theorem 1 relies on two auxiliary results which are proven in the Appendix.

Lemma 2. *Consider any tuple (γ, Q, P, L) feasible for (11). Then, the controller $K = U_0QP^{-1}$ stabilizes (4) and is such that $\|\mathscr{T}(K)\|_2^2 \le \mathrm{tr}\,(W_xP) + \mathrm{tr}\,(W_uL)$.* ■

Lemma 3. *Suppose that condition (2) holds, and consider any controller K which stabilizes (4). Then, there exists a tuple (γ, Q, P, L) feasible for (11) such that $K = U_0QP^{-1}$ and $\|\mathscr{T}(K)\|_2^2 = \mathrm{tr}\,(W_xP) + \mathrm{tr}\,(W_uL)$.* ■

[1] With some abuse of terminology, we refer to (9) and subsequent derivations as an SDP, with the understanding that they can be written as SDP using standard manipulations.

Proof of Theorem 1. Under the controllability assumption there exists a stabilizing and optimal controller $K_\star$. For this stabilizing controller, Lemma 3 ensures the existence of a solution (γ, P, Q, L) for (11), which proves the first part of the claim. Let now $(\gamma_\star, P_\star, Q_\star, L_\star)$ be any optimal solution to (11). By Lemma 2 the controller $K := U_0 Q_\star P_\star^{-1}$ is such that $\|\mathscr{T}(K)\|_2^2 \le \mathrm{tr}(W_x P_\star) + \mathrm{tr}(W_u L_\star)$. On the other hand, since $K_\star$ is stabilizing, Lemma 3 ensures that there exists a solution (γ, P, Q, L) to (11) such that $K_\star = U_0 Q P^{-1}$ and $\|\mathscr{T}(K_\star)\|_2^2 = \mathrm{tr}(W_x P) + \mathrm{tr}(W_u L)$. Since $(\gamma_\star, Q_\star, P_\star, L_\star)$ is by definition optimal for (11) we must therefore have $\mathrm{tr}(W_x P_\star) + \mathrm{tr}(W_u L_\star) \le \mathrm{tr}(W_x P) + \mathrm{tr}(W_u L)$. In turn, this implies $\|\mathscr{T}(K)\|_2 \le \|\mathscr{T}(K_\star)\|_2$. However, since $K_\star$ is the controller minimizing the $\mathcal{H}_2$-norm of the system we must have $\|\mathscr{T}(K)\|_2 = \|\mathscr{T}(K_\star)\|_2$, hence $K = K_\star$ because the optimal controller is unique. ■

In Theorem 1 we say *any* optimal solution because, as it emerges from the proof of Lemma 3, to any stabilizing controller we can associate an infinite number of solutions (γ, Q, P, L) feasible for (11) with $Q = Q_+ + Q_\sim$ where Q_+ is a particular solution and $Q_\sim$ is any matrix in the right kernel of W_0.

We note that (11) is cheap in terms of sample complexity. In fact, in order to find a solution we just need condition (2) fulfilled. By Lemma 1 this holds if the input is persistently exciting of order $n+1$. Thus, in view of Definition 1, even $T = (m+1)n+m$ samples can be sufficient.

4. Certainty-equivalence solution with noisy data

From the previous analysis, when the data are noise-free, $K_\star$ can be computed directly using (11). When $D_0 \neq 0$ the SDP (11) cannot be solved unless we know D_0. A natural alternative which can be computed from data alone consists in disregarding the noise term:

$$\min_{(\gamma,Q,P,L)} \gamma$$
subject to

$$\begin{cases} X_1 Q P^{-1} Q' X_1' - P + I \preceq 0 \\ P \succeq I \\ L - U_0 Q P^{-1} Q' U_0' \succeq 0 \\ X_0 Q = P \\ \mathrm{tr}\,(W_x P) + \mathrm{tr}\,(W_u L) \le \gamma \end{cases} \tag{15}$$

When a solution is found then the controller is computed as $K = U_0 Q P^{-1}$. We will term (15) a *certainty-equivalence* solution since we carry out the design as if the noise were zero. Now, three questions arise for (15):

1. A solution need not exist.
2. Even if a solution is found, the corresponding controller K need not be stabilizing.
3. Even if a solution is found and K is stabilizing, the performance achieved by K might still substantially differ from the performance achieved by $K_\star$.

In the sequel, we will focus on items 2 and 3 above. Item 1 is implicitly addressed in the analysis. Suppose that a solution $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ to (15) is found, and denote by $\overline{K} = U_0 \overline{Q}\,\overline{P}^{-1}$ the corresponding controller. Moreover, let $(\gamma_\star, Q_\star, P_\star, L_\star)$ be any optimal solution to (11) with $K_\star = U_0 Q_\star P_\star^{-1}$. With this notation in place, we aim at establishing the following chain of relations:

$$\begin{aligned} \|\mathscr{T}(\overline{K})\|_2^2 &\le \eta_1(\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L})) \\ &\le \eta_1\eta_2(\mathrm{tr}(W_x P_\star) + \mathrm{tr}(W_u L_\star)) \\ &= \eta_1\eta_2 \|\mathscr{T}(K_\star)\|_2^2 \end{aligned} \tag{16}$$

for some real constants $\eta_1, \eta_2 \ge 1$ along with the property that $\overline{K}$ is stabilizing.

4.1. Stability and performance bounds

We will focus on the two inequalities in (16) as the equality follows from Theorem 1. Consider the first inequality. The idea is to find conditions under which there exists a constant $\eta_1 \ge 1$ such that $\eta_1(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is a feasible solution to (11). Then the inequality follows from Lemma 2. For brevity, we introduce some additional notation. Let

$$\begin{aligned} M &:= Q P^{-1} Q' \\ \Theta &:= X_1 M X_1' - P \\ \Psi &:= D_0 M D_0' - X_1 M D_0' - D_0 M X_1' \end{aligned} \tag{17}$$

With this notation the first constraint in (15) reads $\Theta + I \preceq 0$ while the first constraint in (11) reads $\Theta + \Psi + I \preceq 0$. In the sequel, it is understood that all the solutions of interest inherit the same notation. In particular, we will use $\overline{M}$, $\overline{\Theta}$ and $\overline{\Psi}$ to denote the matrices corresponding to $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ and $M_\star$, $\Theta_\star$ and $\Psi_\star$ to denote the matrices corresponding to $(\gamma_\star, Q_\star, P_\star, L_\star)$.

Lemma 4. *Suppose that (15) is feasible. Let $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ be any optimal solution and let $\overline{K} = U_0\overline{Q}\,\overline{P}^{-1}$. Let $\eta_1 \ge 1$. If the solution satisfies*

$$\overline{\Psi} \preceq \left(1 - \frac{1}{\eta_1}\right) I \tag{18}$$

then the controller $\overline{K}$ is stabilizing and ensures $\|\mathscr{T}(\overline{K})\|_2^2 \le \eta_1 (\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}))$.

Proof. The idea is to show that although $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ need not be feasible for (11), under the condition (18) a feasible solution to (11) is given by $\eta_1(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$. We prove this fact. Since by hypothesis $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is feasible for (15), then it satisfies $\overline{\Theta} + I \preceq 0$. Hence,

$$\begin{aligned} &\eta_1\overline{\Theta} + \eta_1\overline{\Psi} + I = \\ &\eta_1(\overline{\Theta} + \overline{\Psi}) + \eta_1 I + (1-\eta_1) I = \\ &\eta_1(\overline{\Theta} + I) + \eta_1\overline{\Psi} + (1-\eta_1) I \preceq 0 \end{aligned} \tag{19}$$

where the inequality follows from $\eta_1(\overline{\Theta} + I) \preceq 0$ and (18). Hence $\eta_1(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ satisfies the first constraint of (11). Since $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is feasible for (15) then $\eta_1(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ satisfies by construction all the other constraints of (11). Hence, by Lemma 2, $U_0(\eta_1\overline{Q})(\eta_1\overline{P})^{-1} = U_0\overline{Q}\,\overline{P}^{-1} = \overline{K}$ is stabilizing and $\|\mathscr{T}(\overline{K})\|_2^2 \le \mathrm{tr}(W_x\eta_1\overline{P}) + \mathrm{tr}(W_u\eta_1\overline{L})$, which proves the claim. ■

The second inequality in (16) is similar to the first one. The idea is to find conditions under which we can associate to $K_\star$ some tuple $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star)$ feasible for (15).

Lemma 5. *Suppose that condition (2) is satisfied, and let $(\gamma_\star, Q_\star, P_\star, L_\star)$ be any optimal solution to (11). Let $\eta_2 \ge 1$. If the solution satisfies*

$$-\Psi_\star \preceq \left(1 - \frac{1}{\eta_2}\right) I \tag{20}$$

then (15) is feasible and any optimal solution $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is such that $\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}) \le \eta_2\|\mathscr{T}(K_\star)\|_2^2$.

Proof. Condition (2) ensures that problem (11) is feasible and $K_\star = U_0 Q_\star P_\star^{-1}$ where $(\gamma_\star, Q_\star, P_\star, L_\star)$ is any optimal solution to (11). As before, the idea is to show that although $(\gamma_\star, Q_\star, P_\star, L_\star)$ need not be feasible for (15), under (20) a feasible solution to (15) is given by $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star)$. To see this, note that since

$(\gamma_\star, Q_\star, P_\star, L_\star)$ is feasible for (11) then $(\gamma_\star, Q_\star, P_\star, L_\star)$ satisfies $\Theta_\star + \Psi_\star + I \preceq 0$, so

$$\begin{aligned}&\eta_2\Theta_* + I = \\ &\eta_2(\Theta_\star + \Psi_\star) - \eta_2\Psi_\star + \eta_2 I + (1-\eta_2)I = \\ &\eta_2(\Theta_\star + \Psi_\star + I) - \eta_2\Psi_\star + (1-\eta_2)I \preceq 0\end{aligned} \tag{21}$$

where the inequality follows from $\eta_2(\Theta_\star + \Psi_\star + I) \preceq 0$ and (20). Hence, $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star)$ satisfies the first constraint of (15). Further, since $(\gamma_\star, Q_\star, P_\star, L_\star)$ is feasible for (11) then $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star)$ satisfies by construction also all the other constraints of (15). Hence the claim follows because $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is optimal and since the cost associated with the solution $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star)$ is such that $\text{tr}(W_x\eta_2P_\star) + \text{tr}(W_u\eta_2L_\star) = \eta_2 \,\|\mathscr{T}(K_\star)\|_2^2$. ■

We then have the following result.

Theorem 2. *Let U_0, X_0 and X_1 be data generated from an experiment on system (4) possibly with nonzero disturbance vector D_0 and consider problem (15). If (15) is feasible and the solution satisfies (18) then the resulting $\overline{K}$ is stabilizing and such that $\|\mathscr{T}(\overline{K})\|_2^2 \le \eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L}))$. If (2) and (20) also hold then*

$$\frac{\|\mathscr{T}(\overline{K})\|_2^2 - \|\mathscr{T}(K_\star)\|_2^2}{\|\mathscr{T}(K_\star)\|_2^2} \le (\eta_1\eta_2 - 1) \tag{22}$$

with η_1 as in (18) and η_2 as in (20). ■

4.2. Limitations of the certainty-equivalence approach

The SDP (15) is a natural variant of (11) for noisy data and is appealing as it gives stability and performance guarantees without assuming specific statistics on the disturbance. In practice, however, it has some limitations.

Specifically, (15) rests on three conditions: (2), (18) and (20). Condition (2), which is needed also for (11), does not bring major issues: it can be verified from data and may hold even when the disturbance has very large magnitude. In fact, if we write the dynamics as

$$x(k+1) = Ax(k) + \begin{bmatrix} B & I \end{bmatrix} \begin{bmatrix} u(k) \\ d(k) \end{bmatrix} \tag{23}$$

it follows from Lemma 1 that if the augmented input (u, d) is persistently exciting of order $n+1$ then

$$\text{rank} \begin{bmatrix} W_0 \\ D_0 \end{bmatrix} = 2n + m \tag{24}$$

and so condition (2) holds. (In fact, this simply means that the disturbance can be cooperative to enforce excitation.) The main issue is related to (18) and (20), in particular to condition (18). In fact, the term $\overline{\Psi}$ that appears in (18) is the gap, associated with the solution, between the stability condition of the ideal formulation (11) and the condition that we use in the certainty-equivalence approach (15) where we neglect the noise term D_0. In order to fulfil (18), we need a solution such that $\overline{\Psi} = D_0\overline{M}D_0' - X_1\overline{M}D_0' - D_0\overline{M}X_1'$ has small norm, hence such that $\overline{M}$ has small norm. In (15), however, there is no such constraint with the consequence that even a small level of noise may lead to the violation of (18), which is also confirmed in simulation.

In the section that follows we will discuss how to incorporate robustness constraints into (15), and study the properties of this new formulation also in comparison with (15).

5. Noise robustness through soft constraints

As just discussed, we would like to incorporate constraints on M into (15). The solution we propose is based on adding a penalty in the cost function. In particular, we consider the following SDP:

$$\min_{(\gamma,Q,P,L,V)} \gamma$$
subject to
$$\begin{cases} X_1QP^{-1}Q'X_1' - P + I \preceq 0 \\ P \succeq I \\ L - U_0QP^{-1}Q'U_0' \succeq 0 \\ V - QP^{-1}Q' \succeq 0 \\ X_0Q = P \\ \text{tr}\,(W_xP) + \text{tr}\,(W_uL) + \alpha\text{tr}\,(V) \le \gamma \end{cases} \tag{25}$$

where $\alpha > 0$ is a parameter that interpolates performance and robustness, where large values of α favour solutions with $M = QP^{-1}Q'$ having small trace, thus small 2-norm. In particular, since $\|A\| \le \sqrt{\text{tr}(A'A)}$ for any matrix A and because $\text{tr}(AB) \le \text{tr}(A)\text{tr}(B)$ for any matrices $A, B \succeq 0$ we have $\|M\| \le \text{tr}(M)$. Thus, since a sufficient condition for (18) is

$$\|D_0\|^2\|M\| + 2\|D_0\|\|X_1\|\|M\| \le 1 - \frac{1}{\eta_1}$$

it follows that large values of α favour stabilizing solutions. The idea to add constraints on M in the form of *soft* penalty comes from the following consideration. As an alternative to (25) one could think of adding a (hard) constraint of the type $M \preceq \tau I$ with $\tau > 0$. This solution has been studied in De Persis and Tesi (2020a) in the context of robust stabilization and in a preliminary version of this paper (De Persis & Tesi, 2020b) in the context of LQR. This approach, however, favours too much robust solutions and typically leads to large gaps from optimality. We refer the reader to De Persis and Tesi (2020b) for a discussion and for numerical simulations that illustrate this point.

The remainder of this section is as follows. We first derive a counterpart of Theorem 2 for (25) which rests again on (18) and (20). In Section 5.2 we then analyse how the disturbance and α affect (18) and (20). Finally, we discuss how stability and performance can be verified from data.

5.1. Stability and performance bounds

We proceed as before. Assume that a solution $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L}, \overline{V})$ to (25) is found and let $\overline{K} = U_0\overline{Q}\,\overline{P}^{-1}$ be the corresponding controller. Also let $(\gamma_\star, Q_\star, P_\star, L_\star)$ be any optimal solution to (11) with $K_\star = U_0Q_\star P_\star^{-1}$. We want to establish the following chain of relations:

$$\begin{aligned}\|\mathscr{T}(\overline{K})\|_2^2 \;&\le\; \eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L})) \\ &\le\; \eta_1\eta_2(\text{tr}(W_xP_\star) + \text{tr}(W_uL_\star) + \alpha\text{tr}(V_\star)) \\ &=\; \eta_1\eta_2(\|\mathscr{T}(K_\star)\|_2^2 + \alpha\text{tr}(V_\star))\end{aligned} \tag{26}$$

for some constants $\eta_1, \eta_2 \ge 1$ with

$$V_\star := M_\star = Q_\star P_\star^{-1}Q_\star' \tag{27}$$

The first inequality follows as in Lemma 4.

Lemma 6. *Suppose that (25) is feasible. Let $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L}, \overline{V})$ be any optimal solution and let $\overline{K} = U_0\overline{Q}\,\overline{P}^{-1}$. Let $\eta_1 \ge 1$. If the solution satisfies (18) then $\overline{K}$ is stabilizing and ensures $\|\mathscr{T}(\overline{K})\|_2^2 \le \eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L}))$.*

Proof. The proof is analogous to the one of Lemma 4 and therefore omitted. (Note in particular that the constraint in (25) that involves V does not appear in (11).) ■

We also have a natural counterpart of Lemma 5, which establishes the second inequality in (26).

Lemma 7. *Suppose that condition (2) is satisfied, and let* $(\gamma_\star, Q_\star, P_\star, L_\star)$ *be any optimal solution to (11). Let* $\eta_2 \geq 1$. *If the solution satisfies (20) then (25) is feasible, and any optimal solution* $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L}, \overline{V})$ *is such that* $\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}) \leq \eta_2(\|\mathscr{T}(K_\star)\|_2^2 + \alpha \mathrm{tr}(V_\star))$.

Proof. The proof is analogous to the proof of Lemma 5. By proceeding as in Lemma 5 it is immediate to verify that $\eta_2(\gamma_\star, Q_\star, P_\star, L_\star, V_\star)$ is a feasible solution to (25) where the constraint that involves $V_\star$ is satisfied by the choice $V_\star = Q_\star P_\star^{-1} Q_\star'$, which implies $\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}) + \alpha\mathrm{tr}(\overline{V}) \leq \eta_2\,(\mathrm{tr}(W_xP_\star) + \mathrm{tr}(W_uL_\star) + \alpha \mathrm{tr}(V_\star))$. ■

We then have the following result.

Theorem 3. *Let* U_0, X_0 *and* X_1 *be data generated from an experiment on system (4) possibly with nonzero disturbance vector* D_0, *and consider problem (25) with* $\alpha > 0$ *arbitrary. If (25) is feasible and the solution satisfies (18) then the resulting controller* $\overline{K}$ *is stabilizing and such that* $\|\mathscr{T}(\overline{K})\|_2^2 \leq \eta_1(\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}))$. *If (2) and (20) also hold then*

$$\frac{\|\mathscr{T}(\overline{K})\|_2^2 - \|\mathscr{T}(K_\star)\|_2^2}{\|\mathscr{T}(K_\star)\|_2^2} \leq (\eta_1\eta_2 - 1) + \eta_3 \tag{28}$$

with η_1 *as in (18),* η_2 *as in (20), and where*

$$\eta_3 := \alpha \eta_1 \eta_2 \frac{\mathrm{tr}(V_\star)}{\|\mathscr{T}(K_\star)\|_2^2} \tag{29}$$

with $V_\star$ *as in (27).* ■

Theorem 3 is similar to Theorem 2. Like for (15), also here conditions (2) and (20) are not needed to find a stabilizing controller (Lemma 6), but they provide guarantees on the performance gap from $K_\star$. The disadvantage with respect to (15) is that we now have an extra term η_3 which degrades the performance as α increases.

In Theorem 4 we make one step further and provide a characterization of the amount of noise that our method can tolerate. The result is stated in an *epsilon-delta* manner, the precise bound is given in (41). Later in Section 5.2 we will discuss on this result.

Theorem 4. *Let* U_0, X_0 *and* X_1 *be data generated from an experiment on system (4) with persistently exciting input of order* $n+1$. *Consider problem (25) with* $\alpha > 0$ *arbitrary. Then, for every* $\eta_2 \geq 1$ *and* $\eta_1 \geq \eta_2$ *there exists a value* $\delta \geq 0$ *such that if* $\|D_0\| \leq \delta$ *then (2) and (20) are satisfied (implying that (25) is feasible), and any solution to (25) satisfies (18).*

This result relies of the following lemma.

Lemma 8. *Under the same assumptions as in Theorem 4, for every* $\epsilon \geq 0$ *there exists* $\delta \geq 0$ *such that if* $\|D_0\| \leq \delta$ *then* W_0 *has full row rank and* $\|D_0\|\|W_0^\dagger\| \leq \epsilon$, *where* $A^\dagger$ *is the right inverse of a full row rank matrix* A.

Proof. Let $\xi := (x(0), u)$, and let $X_{0,\xi}$ and $X_{0,d}$ be the state data generated by ξ and d, that is

$$[X_{0,\xi}]_{k+1} := A^k x(0) + \sum_{i=0}^{k-1} A^{k-1-i} B u(i)$$

$$[X_{0,d}]_{k+1} := \sum_{i=0}^{k-1} A^{k-1-i} d(i)$$

for $k = 0, \ldots, T-1$, where $[A]_i$ denotes the ith column of a matrix A. Hence, $W_0 = W_{0,\xi} + W_{0,d}$ where

$$W_{0,\xi} := \begin{bmatrix} U_0 \\ X_{0,\xi} \end{bmatrix}, \; W_{0,d} := \begin{bmatrix} 0 \\ X_{0,d} \end{bmatrix} \tag{30}$$

Since by assumption the input signal is persistently exciting of order $n+1$ then $W_{0,\xi}$ has full row rank, which implies $\sigma_{min}(W_{0,\xi}) > 0$ where σ_{min} denotes the smallest singular value. Thus, by standard results on matrix perturbation, W_0 has full row rank if $\|W_{0,d}\| < \sigma_{min}(W_{0,\xi})$. Now recall that, for any matrix A, $\|A\| \leq \|A\|_F \leq \sqrt{r}\|A\|$ where $\|\cdot\|_F$ denotes Frobenius norm and r is the rank of A. Also recall that $\|A\|_F = \|vec(A)\|$ where vec denotes vectorization. It is readily seen that $vec(X_{0,d}) = \Omega\, vec(D_0)$ where

$$\Omega := \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ A & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A^{T-2} & A^{T-3} & A^{T-4} & \cdots & 0 \end{bmatrix}$$

Thus $\|W_{0,d}\| \leq \sqrt{n}\|\Omega\|\|D_0\|$, so W_0 has full row rank if

$$\|D_0\| \leq \frac{1}{2\sqrt{n}\|\Omega\|\|W_{0,\xi}^\dagger\|} \tag{31}$$

where we used the relation $\sigma_{min}(W_{0,\xi}) = \|W_{0,\xi}^\dagger\|^{-1}$, which holds because $W_{0,\xi}$ has full row rank by hypothesis. This inequality is satisfied when $\|D_0\|$ is sufficiently small since Ω and $W_{0,\xi}$ are independent of the disturbance. This shows the first part of the claim. Now recall that $\sigma_{min}(A+B) \geq \sigma_{min}(A) - \|B\|$ for any two matrices A and B. By letting $A = W_{0,\xi}$ and $B = W_{0,d}$, condition (31) gives

$$\|W_0^\dagger\|^{-1} \geq \|W_{0,\xi}^\dagger\|^{-1} - \|W_{0,d}\| \tag{32}$$

Note that condition (31) also implies $2\|W_{0,d}\| \leq \|W_{0,\xi}^\dagger\|^{-1}$. Hence, condition (31) ensures

$$\|W_0^\dagger\| \leq 2\|W_{0,\xi}^\dagger\| \tag{33}$$

This proves the second part of the claim because $\|W_{0,\xi}^\dagger\|$ is independent of the disturbance. ■

Proof of Theorem 4. Let $\|D_0\|$ be such that (31) is satisfied. Then (2) is satisfied and (11) is feasible by Theorem 1. We now determine a particular optimal solution to (11). Let $K_\star$ be the optimal controller as in (8), and let $P_\star \succeq I$ be the unique solution to $(A+BK_\star)P_\star(A+BK_\star)' - P_\star + I = 0$. Define

$$\Phi_\star := \begin{bmatrix} K_\star \\ I \end{bmatrix} \tag{34}$$

$$G_\star := W_0^\dagger \Phi_\star, \quad Q_\star := G_\star P_\star \tag{35}$$

with $W_0^\dagger$ well-defined. Let $\gamma_\star := \mathrm{tr}\,(W_xP_\star) + \mathrm{tr}\,(W_uL_\star)$ with $L_\star := U_0Q_\star P_\star^{-1} Q_\star' U_0'$. Thus $(\gamma_\star, Q_\star, P_\star, L_\star)$ satisfies all the constraints in (11). Further, since $K_\star$ can be written as $K_\star = U_0 Q_\star P_\star^{-1}$ we have $L_\star = K_\star P_\star K_\star'$. By definition of $\mathcal{H}_2$-norm, $\|\mathscr{T}(K_\star)\|_2^2 = \mathrm{tr}\,(W_xP_\star) + \mathrm{tr}\,(W_uL_\star)$ so this solution is optimal for (11) as it achieves the same cost of any other optimal solution (Theorem 1). Thus, a sufficient condition for (20) is

$$\eta_2\|D_0\|^2 \mathrm{tr}(M_\star) + 2\eta_2\|D_0\|\|X_1\|\mathrm{tr}(M_\star) \leq 1 - \frac{1}{\eta_2} \tag{36}$$

with $M_\star = G_\star P_\star G_\star'$. (Note that we could have derived a less conservative condition using $\|M_\star\|$ instead of $\eta_2\mathrm{tr}(M_\star)$ in (36).

We want a condition on the noise enforcing (36) because it is useful for the subsequent part of the proof.) Recall that $X_1 = [B \;\; A]W_0 + D_0$. Since $M_\star = W_0^\dagger \Phi_\star P_\star \Phi_\star' (W_0^\dagger)'$ we thus have $\mathrm{tr}(M_\star) = \|W_0^\dagger \Phi_\star P_\star^{1/2}\|_F^2 \leq \omega \|W_0^\dagger\|^2$, where we have set $\omega := n\|P_\star\|\|\Phi_\star\|^2$. As a final step, recall that $\|W_0\| \leq \|W_{0,\xi}\| + \sqrt{n}\|\Omega\|\|D_0\|$ and that (31) implies $\|W_0^\dagger\| \leq 2\|W_{0,\xi}^\dagger\|$. Hence, under (31), a sufficient condition for (36), hence for (20), is

$$12\omega\|D_0\|^2\|W_{0,\xi}^\dagger\|^2 + 8\omega\sqrt{n}\|\Xi\|\|\Omega\|\|D_0\|^2\|W_{0,\xi}^\dagger\|^2 + 8\omega\|\Xi\|\|W_{0,\xi}\|\|D_0\|\|W_{0,\xi}^\dagger\|^2 \leq \frac{1}{\eta_2}\left(1 - \frac{1}{\eta_2}\right)$$

where $\Xi := [B \;\; A]$. Hence, a condition ensuring both (31) and (36), hence both (2) and (20) is

$$\|D_0\| \leq \frac{1}{\|W_{0,\xi}^\dagger\|} \min\left\{\frac{1}{2\sqrt{n}\|\Omega\|}, \frac{\varepsilon_2}{\sqrt{36\omega}}, \frac{\varepsilon_2}{24\omega\|\Xi\|\kappa(W_{0,\xi})}, \frac{\varepsilon_2}{\sqrt{24\omega\sqrt{n}\|\Xi\|\|\Omega\|}}\right\} =: \delta_1 \tag{37}$$

where $\varepsilon_2 := \eta_2^{-1}(1-\eta_2^{-1})$ and $\kappa(A) := \|A\|\|A^\dagger\|$ is the condition number of A. Condition (37) is satisfied for $\|D_0\|$ sufficiently small because all the quantities defining δ_1 are independent of the noise, and this implies that (2) and (20) hold for $\|D_0\|$ sufficiently small. We next consider (18). Let $(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L}, \overline{V})$ be any optimal solution to (25), which exists under (37) by Lemma 7. We consider two sub-cases. Assume that $\mathrm{tr}(\overline{M}) \leq \eta_2 \mathrm{tr}(M_\star)$. Under (37), since (37) implies (36) and because $\eta_1 \geq \eta_2$ we have

$$\|D_0\|^2\mathrm{tr}(\overline{M}) + 2\|D_0\|\|X_1\|\mathrm{tr}(\overline{M}) \leq 1 - \frac{1}{\eta_1} \tag{38}$$

which ensures (18). Assume next that $\mathrm{tr}(\overline{M}) > \eta_2 \mathrm{tr}(M_\star)$. This is the more difficult case which can occur when α is chosen small. Note that $X_1\overline{M} = (A + B\overline{K})\overline{P}\overline{G}' + D_0\overline{M}$, thus condition (18) can be written as

$$-D_0\overline{M}D_0' - (A + B\overline{K})\overline{P}\overline{G}'D_0' - D_0\overline{G}\overline{P}(A + B\overline{K})' \preceq \left(1 - \frac{1}{\eta_1}\right) I \tag{39}$$

We want to bound in norm the second and the third term on the left side of the above inequality. Let for brevity $\eta_\star := \eta_2\|\mathscr{T}(K_\star)\|_2^2$, $\rho := \min\{\sigma_{min}(W_x), \sigma_{min}(W_u)\}$, and define $\overline{\Phi}$ as in (34) with respect to $\overline{K}$. By Lemma 7, $\mathrm{tr}(W_x\overline{P}) + \mathrm{tr}(W_u\overline{L}) + \alpha\mathrm{tr}(\overline{M}) \leq \eta_\star + \alpha\eta_2\mathrm{tr}(V_\star)$ where $V_\star = M_\star = G_\star P_\star G_\star'$. Since $\mathrm{tr}(\overline{M}) > \eta_2\mathrm{tr}(M_\star)$ we have $\|\overline{P}\| \leq \eta_\star/\rho$ and $\|\overline{L}\| \leq \eta_\star/\rho$. Further, since $\overline{L} = \overline{K}\overline{P}\overline{K}'$ and $\overline{P} \succeq I$ we have $\|\overline{K}\|^2 \leq \eta_\star/\rho$. Accordingly, we have $\|A + B\overline{K}\| = \|\Xi\overline{\Phi}\| \leq \|\Xi\|(1 + \eta_\star/\rho)$. It remains to bound the norm of $\overline{G}$. Notice that $\alpha\mathrm{tr}(\overline{M}) \leq \eta_\star + \alpha\eta_2\mathrm{tr}(M_\star)$. Hence, by recalling that $\mathrm{tr}(M_\star) \leq \omega\|W_{0,\xi}^\dagger\|^2$ we then have $\|\overline{M}\| \leq \beta$ where we set

$$\beta := \frac{\eta_\star}{\alpha} + \omega\eta_2\|W_{0,\xi}^\dagger\|^2 \tag{40}$$

Hence, since $\overline{M} = \overline{G}\overline{P}\overline{G}'$ and $\overline{P} \succeq I$ we have $\|\overline{G}\| \leq \sqrt{\beta}$. It follows that under (37) a sufficient condition for (39) is given by $2\sqrt{\beta}\|D_0\|\|\Xi\|(\rho + \eta_\star)\eta_\star \leq \rho^2(1 - \eta_1^{-1})$. Hence, a condition ensuring (37) and (39), hence (2), (18) and (20) is given by

$$\|D_0\| \leq \min\left\{\delta_1, \frac{\rho^2\varepsilon_1}{2\sqrt{\beta}\|\Xi\|(\rho + \eta_\star)\eta_\star}\right\} \tag{41}$$

where we set $\varepsilon_1 := 1 - \eta_1^{-1}$. This condition is satisfied if $\|D_0\|$ is sufficiently small since, except for D_0, all the other quantities are independent of the noise. ■

5.2. Discussion

Theorems 3 and 4 highlight some interesting facts which we now discuss.

1. Role of α and quality of the data. To have (41) fulfilled one needs $\|D_0\| \leq \delta_1$, where δ_1 is defined in (37). Now, the term δ_1 is independent of the noise: it only depends on the system parameters and on the matrix $W_{0,\xi}$ defined in (30), which is the noiseless input-state response of the system to $\xi = (x(0), u)$. This indicates that the fulfilment of the condition $\|D_0\| \leq \delta_1$ will depend on the quality of the data, namely on the ratio between the magnitude of ξ (useful information) and the magnitude of the noise (useless information). In fact, notice that in δ_1 there are two terms affected by ξ, the condition number $\kappa(W_{0,\xi})$ and $\|W_{0,\xi}^\dagger\|$. The condition number is invariant to a scaling of ξ. On the other hand, the term $\|W_{0,\xi}^\dagger\|$ decreases as ξ scales up since, given $\zeta := \gamma\xi$ with $\gamma \in \mathbb{R}$, it holds that $W_{0,\zeta} = \gamma W_{0,\xi}$. Hence the term $\chi := \|D_0\|\|W_{0,\xi}^\dagger\|$ decreases as the ratio between the magnitude of ξ and the magnitude of the noise increases, ensuring the fulfilment of the condition $\|D_0\| \leq \delta_1$. Under such circumstances, the fulfilment of (41) will thus depend on how big $\|D_0\|$ is compared with $\sqrt{\beta}$, with β defined in (40) (all the other terms in the fraction in (41) are again constants). Since β depends on α and $\|W_{0,\xi}^\dagger\|$, and we are considering the case in which χ is small, then (41) will then essentially depend on the ratio $\|D_0\|/\sqrt{\alpha}$, so a large enough α will suffice to ensure (41). Thus, one can expect that for data of a reasonable quality (χ sufficiently small), and α large enough, the SDP (25) will be effective against noise, which is indeed confirmed in simulation. Clearly, picking α large may degrade the performance as it may lead to large values of η_3 in (28). However, this is not always the case. To see this, notice that the term η_3 in (29) can be upper bounded as

$$\eta_3 \leq \alpha\eta_1\eta_2 \frac{n\|P_\star\|\|\Phi_\star\|^2\|W_0^\dagger\|^2}{\|\mathscr{T}(K_\star)\|_2^2} \tag{42}$$

where we used the bound $\mathrm{tr}(V_\star) \leq n\|P_\star\|\|\Phi_\star\|^2\|W_0^\dagger\|^2$ established in the proof of Theorem 4 after (36). Except for $W_0^\dagger$, all the terms in (42) are parameters. By (33), $\|W_0^\dagger\| \leq 2\|W_{0,\xi}^\dagger\|$ when χ is sufficiently small and this implies $\alpha\|W_0^\dagger\|^2 \leq 4\chi^2\alpha/\|D_0\|^2$. This shows that even large ratios $\sqrt{\alpha}/\|D_0\|$ can lead to small values of η_3 if χ is small. Also this observation is confirmed in simulation. We finally note that $\alpha > 0$ is needed to have (41) fulfilled. If we let $\alpha = 0$ the proof of Theorem 4 breaks down. This stresses the difficulties to get robustness guarantees with the certainty-equivalence solution (15).

2. Role of the weight matrices. The bound in (41) depends on the weight matrices via $\rho = \min\{\sigma_{min}(W_x), \sigma_{min}(W_u)\}$. Simulations show that it is more difficult to get stability when W_u is taken large relatively to W_x and α. This makes sense: when W_u is large the cost function will mostly depend on L which must obey $L \succeq U_0QP^{-1}Q'U_0'$. To keep L small the algorithm will then favour solutions Q with $U_0Q \approx 0$. However, this does not ensure solutions with small $\mathrm{tr}(V)$, meaning that robustness might be lost. This is particularly risky for systems unstable in open-loop. In fact, $U_0Q \approx 0$ implies to restrict the search to low-gain controllers, and this may render stabilization an impossible task. In contrast, when W_u is small the algorithm will favour solutions with small $\mathrm{tr}(V)$ because the other term $\mathrm{tr}(W_xP)$ that appears in the cost function cannot be lowered arbitrarily in view of the constraint $P \succeq I$. Decreasing W_u, however, may affect performance and this can be inferred from (42): as ρ decreases, $\|\mathscr{T}(K_\star)\|_2^2$ gets smaller, potentially yielding larger performance errors.

3. Sample complexity. A final point is related to the number T of data. Like (11), also the SDP (25) is cheap in terms of sample complexity since (41) can be satisfied even with $T = (m+1)n + m$ (the ideal bound of the noiseless case). We cannot, however, provide the typical complexity result stating that the conditions for stability and performance become easier to satisfy as T gets larger. This is however not surprising since (41) makes no assumptions on the noise statistics. We will discuss a case where having large datasets is beneficial in Section 6.2.

5.3. Verifying stability and performance bounds from data

We close this section by discussing how conditions (18) and (20) can be verified from data. Suppose we have computed a solution to (25). If we know that $\|D_0\| \le \delta$ for some $\delta > 0$ then (18) can be tested via the following condition:

$$\delta^2\|\overline{M}\| + 2\delta\|X_1\overline{M}\| \le 1 - \frac{1}{\eta_1} \tag{43}$$

which thus provides a data-dependent test for stability.

If condition (43) is verified then we immediately have also an estimate of the performance gap with respect to the optimal controller $K_\star$. In fact, by Theorem 3, condition (43) ensures $\|\mathscr{T}(\overline{K})\|_2^2 \le \eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L}))$, thus

$$\frac{\|\mathscr{T}(\overline{K})\|_2^2 - \|\mathscr{T}(K_\star)\|_2^2}{\|\mathscr{T}(K_\star)\|_2^2} \le \frac{\eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L}))}{\rho n} - 1 \tag{44}$$

(We used the fact that $\|\mathscr{T}(K_\star)\|_2^2 \ge \rho n$ which follows since $\|\mathscr{T}(K_\star)\|_2^2 \ge \text{tr}(W_xP_\star)$ and $P_\star \succeq I$.). Alternatively, we can get an estimate of the performance loss via (28). Assume that we computed a solution to (25) and that (2) and (43) are satisfied. By Lemma 6, $\eta_1(\overline{\gamma}, \overline{Q}, \overline{P}, \overline{L})$ is feasible for (11), so $\text{tr}(W_xP_\star) + \text{tr}(W_uL_\star) \le \eta_1(\text{tr}(W_x\overline{P}) + \text{tr}(W_u\overline{L})) =: \tilde{\eta}$ where $(\gamma_\star, Q_\star, P_\star, L_\star)$ denotes any optimal solution to (11). Pick in particular the solution given in (34) and (35). Then we have $\|P_\star\| \le \tilde{\eta}/\rho$ and $\|K_\star\|^2 \le \tilde{\eta}/\rho$. Further, we have $M_\star = G_\star P_\star G_\star'$ where $G_\star = W_0^\dagger \Phi_\star$ and where $\Phi_\star$ is as in (34). This implies that $\|M_\star\| \le \|W_0^\dagger\|^2\tilde{\eta}(\rho+\tilde{\eta})/\rho^2$ and $\|X_1M_\star\| \le \|X_1W_0^\dagger\|\|W_0^\dagger\|\tilde{\eta}(\rho+\tilde{\eta})/\rho^2$. Thus, a sufficient condition for (20) is

$$\frac{\tilde{\eta}(\rho+\tilde{\eta})}{\rho^2}\left(\delta^2\|W_0^\dagger\|^2 + 2\delta\|X_1W_0^\dagger\|\|W_0^\dagger\|\right) \le 1 - \frac{1}{\eta_2} \tag{45}$$

which can be tested from data. Hence, if (45) is verified for some $\eta_2 \ge 1$ then (28) applies. In particular, we can upper bound η_3 as $\eta_3 \le \eta_4 := \alpha\eta_1\eta_2\tilde{\eta}(\rho+\tilde{\eta})\|W_0^\dagger\|^2/\rho^3$ which follows from (42), $\|P_\star\|\|\Phi_\star\|^2 \le \tilde{\eta}(\rho+\tilde{\eta})/\rho^2$, and from $\|\mathscr{T}(K_\star)\|_2^2 \ge \rho n$. This leads to

$$\frac{\|\mathscr{T}(\overline{K})\|_2^2 - \|\mathscr{T}(K_\star)\|_2^2}{\|\mathscr{T}(K_\star)\|_2^2} \le (\eta_1\eta_2 - 1) + \eta_4 \tag{46}$$

Compared with (44) this upper bound requires the additional condition (45). However, when applicable it leads to better estimates, in fact estimates potentially close to zero (when we pick α small and $\delta\|W_0^\dagger\|$ is small), which is not possible in general with (44).

The choice of the weight matrices clearly affects also these tests. In particular, according to the previous considerations, decreasing W_u can help to have (43) fulfilled but it may lead to larger performance errors.

6. Extensions: nonlinear systems and de-noising

6.1. Nonlinear systems

The previous analysis extends to the problem of finding the LQR law for a nonlinear system around an equilibrium using data collected from the nonlinear system. In fact, around an equilibrium a nonlinear system can be expressed via its first order approximation plus a reminder, which acts as a process disturbance for the linearized dynamics. Consider a smooth nonlinear system

$$x(k+1) = f(x(k), u(k)) + \xi(k) \tag{47}$$

where ξ is a process disturbance, and let $(\overline{x}, \overline{u})$ be a *known* equilibrium pair, that is such that $\overline{x} = f(\overline{x}, \overline{u})$. Thus, we can rewrite the dynamics as

$$\delta x(k+1) = A\delta x(k) + B\delta u(k) + d(k) \tag{48}$$

where $\delta x := x - \overline{x}$, $\delta u := u - \overline{u}$, A and B are the state and input matrices of the linearized dynamics, $d := \xi + r$ where r accounts for higher-order terms and it has the property that is goes to zero faster than δx and δu, namely

$$r = R(\delta x, \delta u)\begin{bmatrix}\delta x\\ \delta u\end{bmatrix} \tag{49}$$

where $R(\delta x, \delta u)$ is a matrix of smooth functions with the property that $R(\delta x, \delta u)$ goes to zero as $[\delta x' \; \delta u']'$ goes to zero. Now, if the pair (A, B) defining the linearized system is stabilizable then a controller K rendering $A + BK$ stable also exponentially stabilizes the equilibrium $(\overline{x}, \overline{u})$ for the original nonlinear system. Thus, the analysis in Theorem 3 carries over directly to this case.

Corollary 1. *Consider a nonlinear system as in (47), along with a known equilibrium pair $(\overline{x}, \overline{u})$. and let $K_\star$ be the optimal LQR controller of the system linearized around $(\overline{x}, \overline{u})$. Then, Theorem 3 holds with (4) replaced by (48).* ∎

6.2. Gaussian disturbances and large datasets

For random disturbances with special distribution it might be convenient to have large datasets. This is the case when the disturbance has *normal* (Gaussian) distribution. As shown next, the reason is that for this distribution we can decrease the noise effect through de-noising strategies which become effective as the number of data points gets large.

We discuss a simple de-noising strategy based on averaging of *ensembles* (Wang & Uhlenbeck, 1945). Roughly, the idea is that the effect of the disturbances can be filtered out by taking an average of several signal *cycles*. Suppose we can make N experiments on system (4), each of length T, and let $(U_0^{(n)}, D_0^{(n)}, X_0^{(n)}, X_1^{(n)})$, $n = 1, \ldots, N$, be the dataset resulting from the nth experiment. Given N matrices $S^{(n)}$, $n = 1, \ldots, N$, let

$$\underline{S} := \frac{1}{N}\sum_{n=1}^{N} S^{(n)} \tag{50}$$

denote their average. Accordingly, if we average N datasets we obtain the relation

$$\underline{X}_1 = A\underline{X}_0 + B\underline{U}_0 + \underline{D}_0 \tag{51}$$

Hence, the average signals still provide a valid trajectory of the system, meaning that all previous results apply without any modifications. In the rest of this section, to simplify the notation, we let $(X_1, X_0, U_0, D_0) := (\underline{X}_1, \underline{X}_0, \underline{U}_0, \underline{D}_0)$ with the understanding that all the matrices now represent average quantities.

We now specialize the analysis to disturbances with normal distribution. Following Lindgren (1993, Definition p. 435), a vector/collection of random variables is called *multivariate* normal if every linear combination of its entries has normal distribution. For a multivariate normal vector d we will write $d \sim \mathcal{N}(\mu, \Sigma)$, where μ is the mean vector while $\Sigma \succeq 0$ is the variance–covariance matrix. According to this definition, assume that the disturbance vectors $d(k)$ are (stochastically) independent with distribution $\mathcal{N}(0,\sigma^2 I)$. It is immediate to see that in this case the average matrix D_0 has independent and identically distributed (i.i.d.) entries with distribution $\mathcal{N}(0,\sigma^2/N)$. In fact, each entry of D_0 is the sample mean of independent normals having distribution $\mathcal{N}(0,\sigma^2)$, thus it is normal with distribution $\mathcal{N}(0,\sigma^2/N)$ (Lindgren, 1993, Theorems 3 and 4 pp. 180–181). Further, since the vectors $d(k)$ have zero covariances their entries are independent (Lindgren, 1993, Corollary p. 436). This implies that the entries of D_0 form a multivariate normal collection because any of their linear combinations is itself a linear combination of independent normals (Lindgren, 1993, Theorems 3 and 4 pp. 180–181). Finally, any two distinct entries of D_0 have zero covariance since obtained from independent variables. Thus the entries of D_0 are independent, which follows again from Lindgren (1993, Corollary p. 436).

Hence, D_0 has i.i.d. entries with distribution $\mathcal{N}(0,\sigma^2/N)$, which is actually nothing but a consequence that averaging N i.i.d. variables reduces the variance by N. This implies that $E[D_0D_0'] = \sigma^2\frac{T}{N}I$, and hence $E[\|D_0\|] = \sigma\sqrt{T/N}$. Now, the interesting point is that finite sample guarantees on $\|D_0\|$ also hold.

Lemma 9. *Consider N experiments, each of length T, on system (4). Let the disturbance vectors be independent with distribution $\mathcal{N}(0,\sigma^2 I)$. Then, for any $\mu > 0$, the average matrix $D_0 \in \mathbb{R}^{n\times T}$ satisfies*

$$\|D_0\| \le \sigma\sqrt{\frac{T}{N}}\left(1+\mu+\sqrt{\frac{n}{T}}\right) \tag{52}$$

with probability at least $1-e^{-T\mu^2/2}$.

Proof. By the previous discussion, D_0 has i.i.d. entries with distribution $\mathcal{N}(0,\sigma^2/N)$. Then, the result follows directly from Wainwright (2019, Theorem 6.1). ■

Building on Lemma 9, we have the following result which descends directly from Theorems 3 and 4.

Theorem 5. *Consider N experiments, each of length T, on system (4) and let $(X_1,X_0,U_0,D_0) := (\underline{X}_1,\underline{X}_0,\underline{U}_0,\underline{D}_0)$ be the average data matrices. Assume that the experiments are carried out from the same initial state and with the same input signal u, and that u is persistently exciting of order $n+1$. Also assume that the disturbance vectors are i.i.d. with distribution $\mathcal{N}(0,\sigma^2 I)$. Consider problem (25) with $\alpha > 0$ arbitrary. Then, for every $\mu > 0$, $\eta_2 > 1$ and $\eta_1 \ge \eta_2$ there exists a value $\underline{N}$ such that if $N \ge \underline{N}$, with probability at least $1-e^{-T\mu^2/2}$, (25) is feasible and its solution is such that $\overline{K}$ is stabilizing and satisfies (28).* ■

The explicit value of $\underline{N}$ can be directly computed from (41) and, in fact, Theorem 5 is similar in spirit to Theorem 4. The main difference with respect to Theorem 4 is that now the disturbance is also allowed to have large magnitude since its effect can be compensated by N. The assumption that the experiments are carried out from the same initial state and with the same input is used to guarantee that $W_{0,\xi}$ in (41) has full row rank and is independent of N. This assumption can be relaxed by asking that there exists a constant $\underline{w} > 0$ such that $\sigma_{min}(W_{0,\xi}) \ge \underline{w}$ for every N. However, in this case, it is difficult to characterize initial states and inputs for which $\sigma_{min}(W_{0,\xi})$ is independent of N. The conditions to verify stability and performance from data are the same as in Section 5.3.

7. Monte Carlo simulations

7.1. Random linear systems

We consider 100 systems as in (4) with $n = 3$ and $m = 1$, under three types of noise: *white Gaussian noise* (WGN), constant bias and sinusoidal disturbances. In all the cases, we also consider different levels of noise. Simulations were performed in Matlab. For every type (and level) of noise we test (25) with $\alpha = 0.1$, 1 and 10. For each experiment, we choose the entries of A and B and of the initial state from $\mathcal{N}(0,1)$ (command `randn`). For each experiment, the controller is designed using $T = 20$ samples generated by applying an input $u \sim \mathcal{N}(0,1)$ (by Lemma 1 condition (2) requires a minimum of 7 samples). WGN is generated by taking $d \sim \mathcal{N}(0,\sigma^2 I)$. We vary σ by considering different scenarios of the signal-to-noise (SNR), computed (command `snr`) by comparing the data matrices $[x(0)\ \ BU_0]$ and D_0 (*cf.* Section 5.2.1). We included B to measure the effective magnitude of the input signal that enters the system relative to the noise. Constant bias is obtained by applying to each input channel a value κ from a uniform distribution $(-\overline{\kappa},\overline{\kappa})$. Sinusoidal disturbance is obtained by applying to each input channel a signal $\kappa\sin(k)$ with κ as above.

We have solved (25) using CVX (Grant & Boyd, 2014), considering as weight matrices $W_x = I$ and $W_u = 1$. We denote by $\mathcal{S}$ the percentage of times we find a stabilizing controller, and by $\mathcal{V}$ the percentage of times we are able to infer stability via (43) using a bound δ for the norm of D_0. When we consider WGN, we select δ equal to the right side of (52) with $\mu = 0.1$. This choice, which corresponds to a 50% overestimate of σ, gives a correct bound for $\|D_0\|$ in all the experiments. As for constant and sinusoidal disturbances, we consider a worst-case estimate $\hat{D}_0 = \overline{\kappa}\mathbf{1}_{n\times T}$ where $\mathbf{1}_{n\times T}$ is the $n\times T$ matrix of all ones, yielding $\delta = \sqrt{Tn}\overline{\kappa}$. For each type (and level) of noise, we let

$$\mathcal{E}_k := \frac{\|\mathscr{T}(\overline{K}^{(k)})\|_2^2 - \|\mathscr{T}(K_\star^{(k)})\|_2^2}{\|\mathscr{T}(K_\star^{(k)})\|_2^2}$$

be the performance gap in the kth experiment. We denote by $\mathcal{M}$ the median of $\mathcal{E}_k$ through all the experiments that return a stabilizing controller, and by $\mathcal{P}$ the median of the estimate of $\mathcal{E}_k$ obtained via (44) or (46), choosing the best outcome. Each type (and level) of noise is tested with the same set of plant matrices and inputs. The results, reported in Table 1, can be summarized as follows:

1. Stability and stability verification. As expected, the best results in terms of stability are obtained with $\alpha = 10$ which gives very nice results also with very low SNR, while $\alpha = 1$ is a bit more fragile but still effective. In contrast, $\alpha = 0.1$ proves to be much more fragile. This fact could have been inferred by noting that $\delta > 0.33$ already for $\sigma = 0.05$, which gives a value $\delta/\sqrt{\alpha} > 1$ that can be regarded too high for the SDP (25) to be effective against noise (*cf.* Section 5.2.1). This also confirms the fragility of the certainty-equivalent solution (15) corresponding to $\alpha = 0$, which, for WGN, achieves a score $\mathcal{S} = 76\%$ already for $\sigma = 0.1$. The gap between $\mathcal{S}$ and $\mathcal{V}$ was expected and becomes more and more evident as the SNR gets larger. This gap is due to replacing the noise matrix D_0 in (18) with its norm bound δ, which we use in (43) to derive a stability test verifiable from data. Condition (18) is in fact much less conservative than (43). For instance, for WGN with $\sigma = 0.1$, (18) is verified in 77% of the cases with $\alpha = 1$ indicating that the theoretical condition for stability is not so conservative. Nonetheless, also (43) gives acceptable results for a medium–high SNR (≥ 20). Simulations also confirm that robustness increases if we decrease W_u (*cf.* Section 5.2.2). For instance, for WGN with $\sigma = 0.1$, a choice $W_x = I$ and $W_u = 0.01$ returns $\mathcal{S} = 96\%$ and $\mathcal{V} = 48\%$ using $\alpha = 1$.

2. Performance and performance verification. In this case the best results are obtained for $\alpha = 1$ which thus provides a good balance between robustness and performance with median error $\mathcal{M}$ below 5% even for $\sigma = 0.3$. As expected, $\alpha = 10$ results in lower performance for high SNR, a case where large values of α are indeed not needed. We also see that $\alpha = 0.1$ performs

Table 1
Simulation results for 100 random systems using (25) with $W_x = I$ and $W_u = 1$ (if $\mathcal{V} = 0\%$ we write $\times$to denote that the performance cannot be estimated). For WGN, the last three rows report simulation results with averaging.

	WGN $\sigma = 0.01$ SNR $= 33.4$ (dB)	WGN $\sigma = 0.03$ SNR $= 23.9$	WGN $\sigma = 0.05$ SNR $= 19.4$	WGN $\sigma = 0.1$ SNR $= 13.4$	WGN $\sigma = 0.3$ SNR $= 3.9$	Constant bias $\bar{\kappa} = 0.05$ SNR $= 24.7$	Constant bias $\bar{\kappa} = 0.1$ SNR $= 18.7$	Sine wave $\bar{\kappa} = 0.05$ SNR $= 30.2$	Sine wave $\bar{\kappa} = 0.1$ SNR $= 24.1$
$\mathcal{S}$ for $\alpha = 0.1$	96%	91%	89%	82%	77%	95%	91%	96%	94%
$\alpha = 1$	**100**%	97%	95%	91%	83%	97%	96%	98%	96%
$\alpha = 10$	**100**%	**98**%	**97**%	**96**%	**90**%	**98**%	**97**%	**100**%	**100**%
$\mathcal{V}$ for $\alpha = 0.1$	86%	43%	17%	0%	0%	14%	1%	17%	2%
$\alpha = 1$	**92**%	76%	52%	11%	0%	36%	8%	38%	7%
$\alpha = 10$	**92**%	**77**%	**53**%	**16**%	0%	**42**%	**13**%	**42**%	**13**%
$\mathcal{M}$ for $\alpha = 0.1$	**1e−4**	0.0069	0.0183	0.0516	0.1231	**0.0019**	0.0083	**0.0012**	0.0042
$\alpha = 1$	0.0011	**0.0022**	**0.0052**	**0.0137**	**0.0469**	0.0024	**0.0055**	0.0017	**0.0025**
$\alpha = 10$	0.0274	0.0303	0.0333	0.0380	0.0533	0.0254	0.0282	0.0272	0.0303
$\mathcal{P}$ for $\alpha = 0.1$	**1.3954**	2.5980	12.4329	×	×	11.6236	12.3443	4.5470	5.5054
$\alpha = 1$	1.4631	**2.2209**	3.9421	4.8283	×	**4.6175**	9.0106	**4.4997**	**5.0295**
$\alpha = 10$	1.6143	2.5319	**3.4542**	**4.3980**	×	5.0817	**8.6347**	4.8935	7.6219
$\mathcal{S}_{ave}$ for $\alpha = 1$	100%	100%	100%	100%	96%				
$\mathcal{V}_{ave}$ for $\alpha = 1$	100%	99%	97%	94%	70%				
$\mathcal{M}_{ave}$ for $\alpha = 1$	0.0012	0.0013	0.0013	0.0014	0.0034				
$\mathcal{P}_{ave}$ for $\alpha = 1$	1.1345	1.2801	1.3380	1.6097	2.6603				

worse than $\alpha = 1$ for low SNR. This is because for low SNR there are many instances where the solution with $\alpha = 0.1$ is close to instability, resulting in poor performance. The gap between $\mathcal{M}$ and $\mathcal{P}$ is similar to the one between $\mathcal{S}$ and $\mathcal{V}$. We note that for low SNR the best results for $\mathcal{P}$ are often obtained with $\alpha = 10$. The explanation is that for low SNR the best value for $\mathcal{P}$ is often the one coming from (44) which depends on η_1, and η_1 is smaller when α is large. In fact, large values of α lead to small values of $\text{tr}(M)$, making (43) easier to fulfil even with small values of η_1.

3. *Averaging*. For WGN, robustness can be increased by averaging data from multiple experiments (*cf.* Section 6.2). This is advantageous with respect to increasing α because no performance loss is introduced. To emphasize this point, the last three rows of Table 1 report the results for $\alpha = 1$ with $N = 100$ experiments for each system. As expected, the improvement is significant. We note that $\mathcal{M}$ does not converge to zero. This is consistent with the fact that, due to $\alpha > 0$, the solution does not match the optimal one even when $D_0 = 0$. To improve performance, we must decrease α which does not bring issues in this case. For instance, for $\sigma = 0.01$ a choice $\alpha = 0.01$ returns $\mathcal{S} = \mathcal{V} = 100\%$ along with $\mathcal{M} = 1e{-}6$ and $\mathcal{P} = 0.8803$. Very small values of $\mathcal{P}$ are possible in this case since the best values for $\mathcal{P}$ are now those resulting from (46).

7.2. Case study of Dean et al. (2020)

We tested our method on the Laplacian system considered in Dean et al. (2020) under the same setting of input and noise in $\mathcal{N}(0, I)$ (SNR= 0). There, the authors consider an indirect method, called *Coarse-ID control*, that estimates a model from multiple experiments (called *rollouts*) and then designs a controller using both the model and uncertainty estimate. In Dean et al. (2020, Section 6), this method is tested for different choices of the number N of rollouts and the rollout length T using a bootstrap error estimate, which is validated empirically. A total of ≈ 300 samples are shown to be sufficient to have a good trade-off between robustness and performance, namely $\mathcal{S} = 100\%$ and $\mathcal{M} \approx 1$ over 100 trials with $T \approx 60$ and $N = 6$ (Dean et al., 2020, Figure 8). The value $\mathcal{M} \approx 1$ is good since the LQR weight matrices are selected as $W_x = 10^{-3} I$ and $W_u = I$, which gives a small optimal cost ≈ 0.1373. The performance loss further decreases up to $\mathcal{M} \approx 0.6$ for $T = 100$. Similar results but with slightly worse performance are obtained by increasing N and decreasing T, namely $\mathcal{S} = 100\%$ and $\mathcal{M} \approx 2$ using $T = 6$ and $N = 100$ (Dean et al., 2020, Figure 2).

As for our method, we consider $T = 20$ (by Lemma 1 condition (2) requires at least 15 samples). The results are in line with those in Section 7.1, namely large values of α give more robustness but lower performance, while small values of α have opposite effect. We report the results for $\alpha = 1$. With this choice we obtain $\mathcal{S} = 100\%$ and $\mathcal{M} = 0.5766$ already with $N = 8$ over 100 trials obtained by randomly varying input and noise, hence with less than 200 samples ($N > 1$ is needed here given the extremely low SNR). Thus we can conclude that, although originally devised to tackle deterministic noise, our method is equally suited for the stochastic case. As noted in Section 7.1.3, for large values of N one can further improve the performance by decreasing α. For instance, $N = 100$ returns $\mathcal{S} = 100\%$ and $\mathcal{M} = 0.2787$ using $\alpha = 0.1$. We finally point out that, at least for our method, $W_x = 10^{-3} I$ and $W_u = I$ is not the best choice in terms of robustness (*cf.* Section 5.2.2). For instance, for $\alpha = 1$, a choice of the LQR weights equal to $W_x = I$ and $W_u = 10^{-3} I$ returns $\mathcal{S} = 100\%$ and $\mathcal{M} = 0.6270$ already with $N = 1$.

7.3. Nonlinear inverted pendulum

Consider the Euler discretization of an inverted pendulum. The system is as in (47) with

$$f(x,u) = \begin{bmatrix} x_1 + \Delta x_2 \\ \dfrac{\Delta g}{\ell} \sin x_1 + \left(1 - \dfrac{\Delta \nu}{m\ell^2}\right) x_2 + \dfrac{\Delta}{m\ell^2} u \end{bmatrix}$$

where Δ is the sampling time, m is the mass, ℓ is the distance from the base to the centre of mass of the balanced body, ν is the coefficient of rotational friction, and g is the acceleration due to gravity. The states x_1, x_2 are the angular position and velocity, while u is the applied torque. The system has an unstable equilibrium in $(\bar{x}, \bar{u}) = (0, 0)$ corresponding to the pendulum upright position, hence $\delta x = x$ and $\delta u = u$. We assume that the parameters are $\Delta = 0.01$, $m = \ell = 1$, $\nu = 0.01$, and $g = 9.8$.

We made 100 experiments by considering initial conditions in $\mathcal{N}(0, 0.1)$, corresponding to an initial displacement from the equilibrium of about $\pm 10^\circ$, and $u \sim \mathcal{N}(0, 1)$. The results are in line with the previous ones. In particular, when $\xi = 0$ (the only disturbance source is the nonlinearity) we obtain $\mathcal{S} = 100\%$ with $\mathcal{M} = 0.035$ using (25) with trajectories of length $T = 20$. We also considered the case of WGN noise affecting the velocity dynamics, namely u replaced by $u + \xi$ with $\xi \sim \mathcal{N}(0, \sigma)$. In this case, we obtain $\mathcal{S} = 100\%$ for $\sigma \le 0.1$ (SNR ≥ 20) up to $\mathcal{S} = 12\%$ for $\sigma = 1$

(SNR ≈ 0). Just like for linear systems, at the expense of reducing the performance, robustness can be enhanced by decreasing W_u or by increasing α (for instance, with $\alpha = 10$ we have $\mathcal{S} = 64\%$ for WGN with $\sigma = 1$).

8. Concluding remarks

The design of (optimal) controllers from noisy data is a very challenging and largely unsolved problem. In this paper we took some steps in this direction for the LQR problem. By resorting to a convex SDP formulation of the LQR problem, we proposed two novel methods that explicitly account for noise through an augmented cost function which favours noise-robust solutions. Both methods provide finite sample stability guarantees, and do not require specific noise models such as the noise being white.

A great leap forward would come from extending the ideas of this paper to incorporate state and input *safety* constraints (Wabersich & Zeilinger, 2018). At the moment of writing, we aim at tackling this challenge using concepts and tools from *set-invariance* control. For stabilization problems with no optimality requirements, recent results have shown that data-based formulations of set-invariance properties can be efficiently cast as linear programs, and they can handle noisy data (Bisoffi, De Persis, & Tesi, 2019). Another important research venue is the case of partial state information. We have addressed this problem in the context of stabilization (De Persis & Tesi, 2020a), and the extension to optimality constraints seems doable although it requires consideration of several technical aspects.

Appendix

Proof of Lemma 2. The proof follows the same logical steps as Scherer and Weiland (2019, Proposition 3.13) given for the model-based approach. Here, we consider a data-based version. Since $X_0 Q = P$ and $K = U_0 Q P^{-1}$ we have

$$\begin{bmatrix} K \\ I \end{bmatrix} = \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} Q P^{-1} = W_0 Q P^{-1}$$

This implies $A + BK = (X_1 - D_0) Q P^{-1}$. Hence, the first constraint in (11) is equivalent to $S \preceq 0$ where

$$S := (A + BK) P (A + BK)' - P + I \tag{A.1}$$

Hence K is stabilizing. As for the second part of the claim, since $S \preceq 0$ there exists a matrix Ξ such that $S + \Xi \Xi' = 0$. Thus, P coincides with the controllability Gramian of the extended system

$$\begin{bmatrix} x(k+1) \\ z(k) \end{bmatrix} = \left[\begin{array}{c|c} A + BK & \begin{bmatrix} I & \Xi \end{bmatrix} \\ \hline \begin{bmatrix} W_x^{1/2} \\ W_u^{1/2} K \end{bmatrix} & 0 \end{array}\right] \begin{bmatrix} x(k) \\ d(k) \\ \xi(k) \end{bmatrix} \tag{A.2}$$

with ξ an additional input. Let us call $\mathscr{T}_e(K)$ the transfer function of (A.2). By definition of $\mathcal{H}_2$-norm,

$$\begin{aligned} \|\mathscr{T}_e(K)\|_2^2 &= \operatorname{tr}(W_x P) + \operatorname{tr}\left(W_u K P K'\right) \\ &= \operatorname{tr}(W_x P) + \operatorname{tr}\left(W_u U_0 Q P^{-1} Q' U_0'\right) \\ &\le \operatorname{tr}(W_x P) + \operatorname{tr}(W_u L) \end{aligned}$$

In turn, this implies $\|\mathscr{T}(K)\|_2^2 \le \operatorname{tr}(W_x P) + \operatorname{tr}(W_u L)$ because $\mathscr{T}_e(K) = [\mathscr{T}(K) \;\; \mathscr{T}_\Xi(K)]$, where $\mathscr{T}_\Xi(K)$ is the transfer function of (A.2) from ξ to z. ■

Proof of Lemma 3. Consider any stabilizing controller and denote by P the controllability Gramian associated with the closed-loop system (5), which solves $S = 0$ with S as in (A.1). Consider the system of equations (12) in the unknown G, which has a solution under (2). In particular, pick

$$G_+ := W_0^\dagger \begin{bmatrix} K \\ I \end{bmatrix}$$

where $A^\dagger$ denotes the right inverse of a full row rank matrix A. Now, define $Q := G_+ P$, $L := U_0 Q P^{-1} Q' U_0'$ and $\gamma := \operatorname{tr}(W_x P) + \operatorname{tr}(W_u L)$. Thus (γ, Q, P, L) is feasible for (11), and $K = U_0 Q P^{-1}$ which shows the first part of the claim. The second part of the claim follows from

$$\begin{aligned} \|\mathscr{T}(K)\|_2^2 &= \operatorname{tr}(W_x P) + \operatorname{tr}\left(W_u K P K'\right) \\ &= \operatorname{tr}(W_x P) + \operatorname{tr}(W_u L) \end{aligned}$$

which concludes the proof. ■

References

Abbasi-Yadkori, Y., Lazíc, N., & Szepesvári, C. (2018). Model-free linear quadratic control via reduction to expert prediction. https://arxiv.org/pdf/1804.06021.
Alpago, D., Lygeros, J., & Dörfler, F. (2020). An extended Kalman filter for data-enabled predictive control. https://arxiv.org/pdf/2003.08269.
Åström, K., & Wittenmark, B. (1989). *Adaptive control*. Addison-Wesley.
Baggio, G., Katewa, V., & Pasqualetti, F. (2019). Data-driven minimum-energy controls for linear systems. *IEEE Control Systems Letters*, *3*(3), 589–594.
Berberich, J., Koch, A., Scherer, C., & Allgöwer, F. (2019). Robust data-driven state-feedback design. https://arxiv.org/pdf/1909.04314.
Bisoffi, A., De Persis, C., & Tesi, P. (2019). Data-based guarantees of set invariance properties. https://arxiv.org/pdf/1911.12293.
Bradtke, S., Ydstie, B., & Barto, A. (1994). Adaptive linear quadratic control using policy iteration. In *1994 American control conference*. IEEE.
Campi, M., & Weyer, E. (2002). Finite sample properties of system identification methods. *IEEE Transactions on Automatic Control*, *47*(13), 29–34.
Chen, T., & Francis, B. (1995). *Optimal sampled-data control systems*. London: Springer.
Cohen, A., Koren, T., & Mansour, Y. (2019). Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. https://arxiv.org/pdf/1902.06223.
Coulson, J., Lygeros, J., & Dörfler, F. (2019). Data-enabled predictive control: In the shallows of the deePC. In *2019 18th European control conference (ECC)* (pp. 307–312).
Dai, T., & Sznaier, M. (2018). A moments based approach to designing MIMO data driven controllers for switched systems. In *57th IEEE conference on decision and control (CDC)*. IEEE.
De Persis, C., & Tesi, P. (2019). On persistency of excitation and formulas for data-driven control. In *58th IEEE conference on Decision and Control (CDC)* (pp. 873–878). IEEE.
De Persis, C., & Tesi, P. (2020a). Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, *63*(3), 909–924.
De Persis, C., & Tesi, P. (2020b). Low-complexity learning of linear quadratic regulators from noisy data. https://arxiv.org/pdf/2005.01082.
Dean, S., Mania, H., Matni, N., Recht, B., & Tu, S. (2020). On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, *20*, 633–679.
Fazel, M., Ge, R., Kakade, S. M., & Mesbahi, M. (2018). Global convergence of policy gradient methods for the linear quadratic regulator. In *35th international conference on machine learning, ICML 2018, Vol. 4* (pp. 2385–2413).
Feron, E., Balakrishnan, V., Boyd, S., & El Ghaoui, L. (1992). Numerical methods for H_2 related problems. In *1992 American control conference* (pp. 2921–2922). IEEE.
Fiechter, C. (1997). PAC Adaptive control of linear systems. In *1997 conference on learning theory*.
Gonçalves da Silva, G. R., Bazanella, A., Lorenzini, C., & Campestrini, L. (2019). Data-driven LQR control design. *IEEE Control Systems Letters*, *3*(1), 180–185.
Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx.
Guo, M., De Persis, C., & Tesi, P. (2020). Learning control for polynomial systems using sum of squares. In *59th IEEE conference on Decision and Control (CDC)* (pp. 2436–2441). IEEE.
Hou, Z.-S., & Wang, Z. (2013). From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, *235*, 3–35.
Lindgren, B. (1993). *Statistical theory* (fourth ed.). Champman-Hall, CRC.
Mania, H., Tu, S., & Recht, B. (2019). Certainty equivalence is efficient for linear quadratic control. https://arxiv.org/pdf/1902.07826.
Novara, C., Formentin, S., Savaresi, S., & Milanese, M. (2016). Data-driven design of two degree-of-freedom nonlinear controllers: the D^2-IBC approach. *Automatica*, *72*, 19–27.

Recht, B. (2019). A tour of reinforcement learning: the view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, *3*, 253–279.

Salvador, J., Muñoz de la Peña, D., Alamo, T., & Bemporad, A. (2018). Data-based predictive control via direct weight optimization. In *6th IFAC Conference on Nonlinear Model Predictive Control* (pp. 356–361).

Scherer, C., & Weiland, S. (2019). *Lecture notes, dutch institute for systems and control, Linear matrix inequalities in control*. The Netherlands: Delft.

van Waarde, H., Eising, J., Trentelman, H., & Camlibel, K. (2020). Data informativity: a new perspective on data-driven analysis and control. *IEEE Transactions on Automatic Control*, *99*, 1.

Wabersich, K. P., & Zeilinger, M. N. (2018). Scalable synthesis of safety certificates from data with application to learning-based control. In *2018 European control conference, ECC 2018* (pp. 1691–1697).

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*. Cambridge University Press.

Wang, M., & Uhlenbeck, G. (1945). On the theory of the brownian motion II. *Reviews of Modern Physics*, *17*(2–3), 323–342.

Willems, J. C., Rapisarda, P., Markovsky, I., & De Moor, B. (2005). A note on persistency of excitation. *Systems & Control Letters*, *54*(4), 325–329.

Claudio De Persis is a Professor with the Engineering and Technology Institute, Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands. He received the Laurea and Ph.D. degree in engineering in 1996 and 2000 both from the University of Rome "La Sapienza," Italy. His main research interest includes automatic control and its applications, and his recent research focuses on learning for control, networked cyber–physical systems, smart grids, and resilient control.

Pietro Tesi received the Ph.D. degree in computer and control engineering from the University of Florence, Italy, in 2010, where is currently an Associate Professor. He also holds a Guest Research position with the University of Groningen, The Netherlands. His main research interests include adaptive and learning systems, and networked control. Dr. Tesi is an Editor for the International Journal of Robust and Nonlinear Control, and an Associate Editor for the IEEE Control Systems Letters.