Data-driven stabilization and safe control of nonlinear systems

Luppi, Alessandro

DOI:
10.33612/diss.273310133

Publication date:
2023

Data-driven stabilization and safe control of nonlinear systems

Alessandro Luppi

This research has been carried out at the Faculty of Science and Engineering, University of Groningen, The Netherlands as part of the Smart Manufacturing Systems research group within the ENgineering and TEchnology Institute Groningen (ENTEG).

This research has been carried out under the auspices of the centre for Data Science and Systems Complexity (DSSC) at the University of Groningen and is supported by a Marie Skłodowska-Curie COFUND grant, no. 754315.

The research reported in this dissertation is part of the research program of the Dutch Institute of Systems and Control (DISC). The author has successfully completed the educational program of DISC.

The cover is designed with Haikei.app
An electronic version of this dissertation is available at University of Groningen Library

Version 1.6

Data-driven stabilization and safe control of nonlinear systems

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Tuesday 17 January 2023 at 11.00 hours

by

Alessandro Luppi

born on 25 August 1991
in Casorate Primo, Italy

Supervisor
Prof. C. De Persis

Co-supervisors
Dr. P. Tesi
Dr. ir. B. Besselink

Assessment Committee
Prof. B. Jayawardhana
Prof. S. Tarbouriech
Prof. A. Garulli

CONTENTS

1

INTRODUCTION

1.1. DATA DRIVEN APPROACH

Historically, the design of a stabilizing controller relies on the knowledge of a mathematical model to represent the underlying system. This model can be derived from first principles or from measurements using a proper excitation signal. The latter approach, called system identification, infers the model of the system by analyzing the input/output response of the process. The collected data are used to derive a model that will be used afterward for the design of a stabilizing controller. This solution can be considered as an indirect data-driven design. More recently, the necessity of identifying a model was reconsidered in favor of a direct data driven design, where controllers are directly synthesized from experimental data.

Various efforts have been made in this direction, and we refer the interested reader to [1] for a survey on the topic. A direct design has the advantage that it does not suffer from modeling errors, and the overall controller design is simplified by removing the need for the intermediate identification step. This can be especially appealing when dealing with nonlinear systems as deriving a reliable model is challenging and prone to errors. Potentially complex nonlinear behaviors can be learned from data and used to improve the performance.

One of the main open research questions on data-driven control is how to provide provably correct conditions for stability of the closed-loop system with the learned controller. A promising result comes from behavioral system theory, where Willems et al. [2] have proved that a single data trajectory obtained from one experiment can be used to represent all the input-output trajectories of a linear time-invariant system. Since [3], which highlighted the relevance of [2] for the synthesis of data-enabled predictive control, several papers have found inspiration from the paper of Willems et al.

The result has been revisited by using a classic state-space description in [4], where a data-driven parametrization of a closed-loop system was derived and used to learn feedback controllers for unknown systems from data. Using this setting, several recent contributions have further explored the role of data in synthesis problems in multiple areas of control, including: optimal and robust control [4], [5], robust stabilization of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control with noisy data [4], [6], [7], [8], [9], dissipativity properties [10] [11], robust set-invariance [12], and the L_2-gain [13].

The design of a data-driven stabilizing controller for nonlinear systems has been approached using various methods and for different classes of nonlinear systems: in [14] a novel data-driven control is designed relying on a model inversion approach, the use of the Koopman operator is presented in [15] and [16] to reformulate nonlinear systems in a linear framework and in [17] the Koopman operator is combined with an MPC. Discrete-time bilinear systems are considered in [12] and [18], [19] and [20] study nonlinear polynomial systems and formulate semidefinite programs relaxed using sum-of-squares programming. Second-order discrete Volterra systems are discussed in [21], [22] studies nonlinear non-affine MIMO systems, [9] discuss how to exploit prior knowledge in a data-driven design for linear systems with nonlinear uncertainties, noisy Lur'e systems are considered in [23], and [24] presents a data-driven robust design based on approximate nonlinearity cancellations.

In this thesis, we will not limit our study to controllers that achieve stabilization, but we will also consider data-driven controller design for safety. Namely, for safety-critical systems, it is not sufficient to stabilize the process around a desired operating point. Instead, the (data-driven) controller is also required to be able to steer the system state away from any dangerous regions. The complement of this dangerous set is the so-called safe set and the design of control schemes to respect safety specifications goes by the name of safe control [25–28]. The notion of safety is intuitively related to the notion of invariance, namely, the dynamical propriety that the state belongs to a certain set for all subsequent times after being initialized in there. Historically, a seminal result for invariance was Nagumo's theorem [29], which has been of fundamental importance in the characterization of invariant sets for continuous-time systems. Given that almost every system in practice is subject to some type of constraints on its states or outputs, the notion of invariance is very relevant in control applications to include safety constraints in the design. Indeed, problems related to safety and viability [30] can be addressed by computing sets possessing invariance or closely-related properties.

In the control community, invariance for linear systems was extensively studied in the 1980-1990's, which resulted in several works surveyed in [31]. For nonlinear systems, there has recently been a revived interest with the introduction of control Lyapunov-like functions tailored to enforce invariance, called control barrier functions [32–35]. In this setting, the controller depends critically on the model used for design and recent studies [36–38] focus on finding robust safe controllers to account for possibly inaccurate models. In [39], unmodeled dynamics are taken into account by adding a bounded disturbance on the input to find a robust control barrier function, in an input-to-state-safety fashion.

In an effort to reduce this critical dependence from the model, in this thesis we will present a purely data-driven solution to obtain a safe controller.

1.2. NOTATION

The notation in this thesis is as follows.

- We denote by $\mathbb{R}$ the set of real numbers.
- Let $\mathbb{Z}$ denote the set of integers and with $\mathbb{Z}_{\geq 0}$ we indicate the set of nonnegative integers.
- Given a vector a, $|a|$ denotes its 2-norm.
- An identity matrix is denoted by I.
- For a matrix A, $||A||$ denotes its induced 2-norm, which is equivalent to the largest singular value of A. Moreover, for a scalar $a > 0$, $||A|| \leq a$ if and only if $A^\top A \preceq a^2 I$ where I is the identity matrix.
- For a matrix A, $A^\top$ denotes its transpose.
- A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is positive semidefinite, i.e., $A \succeq 0$, if $x^\top A x \geq 0$ for all $x \in \mathbb{R}^n$.
- Given two matrices A and B, $A \succeq B$ means $A - B \succeq 0$.
- For a positive semidefinite $A = A^\top$, $A^{1/2}$ denotes the unique positive semidefinite root of A.
- For matrices M and $N = N^\top$, we sometimes abbreviate $MNM^\top$ as $MNM^\top = M \cdot N[\star]^\top$.
- For symmetric matrices A and C, we sometimes abbreviate the symmetric matrix $\left[\begin{smallmatrix} A & B^\top \\ B & C \end{smallmatrix}\right]$ as $\left[\begin{smallmatrix} A & \star \\ B & C \end{smallmatrix}\right]$ or $\left[\begin{smallmatrix} A & B^\top \\ \star & C \end{smallmatrix}\right]$.

1.3. THESIS OUTLINE

In this thesis, we investigate direct data-driven designs of stabilizing controllers for nonlinear systems. The first part of the thesis considers nonlinear systems with nonlinearities that satisfy quadratic constraints. In Chapter 2 the problem of absolute stability is introduced and the data-driven framework used is explained. Chapter 3 continues the discussion on the absolute stabilization problem by considering the disturbance directly in the design of the stabilizing controller.

The second part of the thesis is focused on another class of nonlinear systems (polynomial systems) and studies how to design a controller from data that not only allows to stabilize the system but it is also able to satisfy safety constraints. Chapter 4 introduces sum-of-squares (SOS) programming and the setup considered. The main result in Chapter 4 is an algorithm that is able to provide a polynomial controller and the corresponding invariance set for the closed-loop system. Starting from these results, Chapter 5 extends the discussion by considering the case of robust invariance to arrive at a new algorithm that can provide a controller that is guaranteed to work despite the presence of disturbances.

We briefly outline the contents of each chapter.

- In Chapter 2, we present a data-driven solution to design a state feedback controller to guarantee absolute stability for a class of nonlinear system. In this chapter, we consider nonlinear systems where the nonlinear component satisfies a quadratic constraint. Moreover, we assume that the system can be represented with an isolated nonlinear block in closed-loop with the linear component of the dynamics and that it is possible to measure the state of the linear and nonlinear parts of the system separately. First we will assume that some information about the non-linearity of the system is known to formulate our first result, then we will drop this assumption to show how to derive a stabilizing controller without prior knowledge about the system. To simplify the presentation for this chapter, we will neglect the effect of noise in the data.

- In Chapter 3, starting from the results of the previous chapter, we study a more realistic case by considering the presence of noise in the data used for the controller design. We will show how the main results of Chapter 2 can be reformulated to account for imperfect data. To handle the noise we will only assume that it is bounded, and we do not need to assume that a statistical distribution for the disturbance is known.

- In Chapter 4, we introduce the concept of safety as an additional requirement that the controller must guarantees besides closed-loop stability.

Safety requirements are express as a list of state constraints that must never be violated. We do not differentiate between soft and hard safety constraints and all specifications are considered as hard constraints. For the design of a safe controller, we consider polynomial systems. We will not assume to know the degree of the system polynomials, but as we will discuss, it is possible to include prior-knowledge about the system to improve noise robustness. The main result presented in this chapter is a data-driven design procedure that uses a sum of squares relaxation to solve a single semi-definite problem to obtain not only a safe state-feedback controller but also the corresponding expression of the invariant set where the system state is guaranteed to be contained.

- In Chapter 5, we continue the study of safe controllers for polynomial systems by considering the presence of noise during the execution of the control task and not only in the open-loop data collected from the system. The main difference from the formulation of the previous chapter is in the definition of invariant set used to derive the main result. So we will start by defining the concept of robustly invariant set to account for the presence of the noise during the execution of the control task to reformulate a safe controller design with guaranteed noise robustness. We will show with a numerical example that the performance obtained from the robust safe controller is similar to the one obtained with the safe controller design of Chapter 4.
- In Chapter 6, we summarize the main results presented in this thesis and discuss interesting directions for future work.

The detailed contributions of each chapter will be presented therein.

1.4. PUBLICATIONS DURING PHD

Journal Papers

- Luppi, A., De Persis, C., & Tesi, P. "On data-driven stabilization of systems with nonlinearities satisfying quadratic constraints." Systems & Control Letters 163 (2022): 105206.
- Luppi, A., Bisoffi, A., De Persis, C., & Tesi, P. "Data-driven design of safe control for polynomial systems.", submitted, available at arXiv preprint arXiv:2112.12664 (2021).

2

STABILIZATION OF SYSTEMS WITH NONLINEARITIES SATISFYING QUADRATIC CONSTRAINTS

ABSTRACT

In this chapter, we directly design a state feedback controller that stabilizes a class of uncertain nonlinear systems solely based on input-state data collected from a finite-length experiment. Necessary and sufficient conditions are derived to guarantee that the system is absolutely stabilizable and a controller is designed. Results derived under some relaxed prior information about the system, strengthened data assumptions are also discussed. All the results are based on semi-definite programs that depend on input-state data only, which – once solved – directly return controllers. As such they represent end-to-end solutions to the problem of learning control from data for an important class of nonlinear systems. Numerical examples illustrate the method with different levels of prior information.

This chapter has been published in "On data-driven stabilization of systems with nonlinearities satisfying quadratic constraints."Luppi, Alessandro, Claudio De Persis, and Pietro Tesi; Systems & Control Letters 163 (2022): 105206.

2.1. INTRODUCTION

In this chapter, we focus on finding a data-driven solution to the problem of absolute stabilizability. Absolute stabilizability is the problem of enforcing the stability of the origin via feedback for a class of systems comprising a linear part and nonlinearities that satisfy a given condition. In our work, we impose quadratic restrictions on the nonlinearities. The absolute stability problem was originally formulated by A.I. Lurie in [40] and solved with a Lyapunov based approach. Later in [41] V.M. Popov proposed a solution in the frequency domain. These two approaches were later connected by the Kalman-Yakubovich-Popov Lemma [42] that relates an analytic property of a square transfer matrix in the frequency domain to a set of algebraic equations involving parameters of a minimal realization in time domain.

To solve the absolute stabilizability problem in the case in which the system is unknown, we start from the data driven parametrization in [4] for a closed-loop system with the addition of a nonlinear term. Then we propose a data-dependent Lyapunov-based control design assuming that a finite number of samples measuring the nonlinear term are available. One of the advantages of our formulation, compared to a model based one, is that it holds for both continuous-time and discrete-time systems providing a unified analysis and design framework for both classes of systems. We also discuss how our results can be viewed in a frequency domain stability analysis where our main result can be interpreted as a data-dependent feedback Kalman-Yakubovitch Popov Lemma [43, Section 2.7.4]. Finally, we discuss how to deal with different levels of prior knowledge by strengthening the assumption on the collected data. The design of controllers using data perturbed by disturbances will be discussed in Chapter 3.

Our main contributions are necessary and sufficient conditions for which a solution to the data-driven absolute stabilization problem exists. The proposed conditions can be verified directly from data by means of efficient linear programs. These conditions provide a new data based solution to the problem of absolute stabilizability for a class of uncertain nonlinear systems.

This chapter is organized as follows. The notation and problem setup is presented in Section 2.2. In Section 2.3, we present necessary and sufficient conditions under which the data-driven absolute stabilizability problem is solvable and provide the explicit expression for the controller. Finally, in Section 2.4, we relax some prior knowledge about the system and derive new conditions for the problem solution. All the main results are also illustrated with practical examples. Concluding remarks are given in Section 2.5.

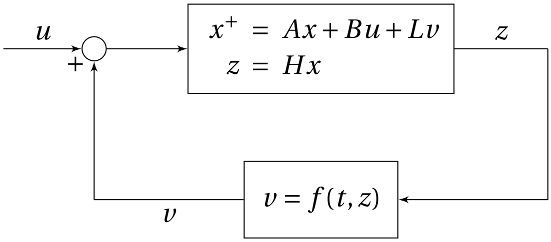

Figure 2.1 | Schematic diagram of system (2.2). Data containing the measurements of the state x, input u and signal v are used to design a feedback controller $u = Kx$ for the system.

2.2. FRAMEWORK AND PROBLEM FORMULATION

We study the stabilization problem of a nonlinear system of the form

$$x^+ = Ax + Bu + \hat{f}(t,x) \tag{2.1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, and where $\hat{f}(t,x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is a memoryless, possibly time-varying nonlinearity. The matrices A, B and the map $\hat{f}$ are unknown. If some prior information is available regarding $\hat{f}$, we will model it in the form $\hat{f}(t,x) = Lv$, $v = f(t,Hx)$, with L and H known matrices and $f : \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}^q$. The system then becomes (see Figure 2.1)

$$\begin{aligned} x^+ &= Ax + Bu + Lv \\ z &= Hx \\ v &= f(t,z). \end{aligned} \tag{2.2}$$

We will consider certain constraints on the admissible nonlinearities. Specifically, we will assume that the inequality

$$\begin{bmatrix} z \\ f(t,z) \end{bmatrix}^\top \begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^\top & \hat{R} \end{bmatrix} \begin{bmatrix} z \\ f(t,z) \end{bmatrix} \geq 0 \tag{2.3}$$

holds for all the pairs $(t,z) \in \mathbb{R} \times \mathbb{R}^p$ with $z \in \operatorname{im} H$, where $\hat{Q} = \hat{Q}^\top \in \mathbb{R}^{p \times p}$, $\hat{S} \in \mathbb{R}^{p \times q}$ and $\hat{R} \prec 0 \in \mathbb{R}^{q \times q}$ are known matrices (definite matrices are implicitly defined as symmetric matrices). Since $\hat{R} \prec 0$ the inequality (2.3) implies $f(t,0) = 0$ for all $t \in \mathbb{R}$. In the remainder, we will sometimes ask a constraint of the form (2.3) to be *regular*, by which we mean that there exists a pair $(\overline{z}, f(\overline{t},\overline{z}))$ such that the inequality (2.3) evaluated at $(\overline{z}, f(\overline{t},\overline{z}))$ strictly holds.

For this class of systems, a notion of stability widely studied in the literature is the so-called *absolute stability* which is now introduced (*cf.* [44, Definition 7.1]).

2

Definition 1 *System* (2.2) *is said to be absolutely stabilizable via linear state-feedback* $u = Kx$ *if there exists a matrix* K *such that the origin of the closed-loop system*

$$\begin{aligned} x^+ &= (A+BK)x + Lv \\ z &= Hx \\ v &= f(t,z) \end{aligned} \tag{2.4}$$

is globally uniformly asymptotically stable for any function f *that satisfies the inequality* (2.3). ∎

This framework covers a notable class of nonlinear systems that has appeared in several studies. In the following, we report few notable examples:

1. *(Lipschitz or norm bounded nonlinearities)* These are nonlinearities that satisfy
$$f(t,z)^\top f(t,z) \le \ell^2 z^\top z,$$
where ℓ is a bound on the Lipschitz constant, which are considered in LMI-based robust stabilization of nonlinear systems [45]. In the absolute stability theory literature, e.g. [46, Section 3], these nonlinearities are referred to as norm-bounded nonlinearities. In this case, (2.3) holds with $\hat{Q} = \ell^2 I_p$, $\hat{S} = 0$ and $\hat{R} = -I_q$.

2. *(Bounds on partial gradients)* Large Lipschitz constants might result in unfeasible conditions, and for this reason much literature is devoted to deriving less conservative Lipschitz characterizations ([47] and references therein). Here, we recall a result from [48] that considers the nonlinear time-invariant term $\hat{f}(x)$ instead of $Lf(t,z)$. Under continuous differentiability of $\hat{f}$, and assuming that bounds $\underline{f}_{ij}, \overline{f}_{ij}$ on the partial derivatives are known, namely
$$\underline{f}_{ij} \le \frac{\partial \hat{f}_i}{\partial x_j} \le \overline{f}_{ij}$$
it holds that $\hat{f}(x) = Lf(z)$, where $L = I_n \otimes \mathbb{1}_n^\top$, $H = I_n$ and (2.3) holds with the matrices $\hat{Q}, \hat{S}$ and $\hat{R}$ depending on the vectors of bounds $\overline{f}, \underline{f}$. For instance, in the case $n = 2$, the matrix in (2.3) takes the form
$$\left[\begin{array}{cc|cccc} \sum_{i=1}^{2}(\overline{c}_{i1} - c_{i1}) & 0 & c_{11} & 0 & c_{21} & 0 \\ 0 & \sum_{i=1}^{2}(\overline{c}_{i2} - c_{i2}) & 0 & c_{12} & 0 & c_{22} \\ \hline \star & \star & -1 & 0 & 0 & 0 \\ \star & \star & 0 & -1 & 0 & 0 \\ \star & \star & 0 & 0 & -1 & 0 \\ \star & \star & 0 & 0 & 0 & -1 \end{array}\right] \tag{2.5}$$

where $c_{ij} = (\overline{f}_{ij} + \underline{f}_{ij})/2$ and $\overline{c}_{ij} = (\overline{f}_{ij} - \underline{f}_{ij})/2$. For additional degree of freedom, one can introduce a vector λ of non-negative multipliers, see [48] for details.

3. *(Strongly convex functions with Lipschitz gradient)* Let $\hat{f}(x) = \nabla g(x)$, with $\hat{f}(0) = 0$ and $g : \mathbb{R}^n \to \mathbb{R}$ a continuously differentiable strongly convex function with parameter m and having Lipschitz gradient with parameter ℓ, with $0 < m < \ell$. Then condition (2.3) holds with $\hat{Q} = -2m\ell I_n$, $\hat{S} = (\ell + m)I_n$ and $\hat{R} = -2I_n$, see [49, Proposition 5, (3.13d)].

4. (*Sector bounded nonlinearities*) Another notable case is when the nonlinear function f satisfy the sector condition $(f(t,z) - K_1 z)^\top (K_2 z - f(t,z)) \geq 0$ for $(t,z) \in \mathbb{R} \times \mathbb{R}^p$, with $K = K_2 - K_1 \succ 0$ [44, Definition 6.2]. In this case, inequality (2.3) holds with $\hat{Q} = -K_2^\top K_1 - K_1^\top K_2$, $\hat{S} = K_1^\top + K_2^\top$ and $\hat{R} = -2I_n$.

5. *(Fully recurrent neural network (RNN))* Recurrent neural networks are specialized in the analysis of sequence of data and they differ from traditional neural network by their "memory". While traditional deep neural networks assume that inputs and outputs are independent of each other, the output of recurrent neural networks depend on the prior elements within the sequence. Systems like (2.2) with $B = 0$ are also used to represent RNN, in which case $f(z) = (f(z_1) \ldots f(z_p))^\top$, $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable monotone nonlinear function whose derivative f' is bounded from above on the domain $\mathbb{R}$, e.g. $f(z_i) = \tanh(z_i)$, and $z_i = H_i^\top x + b_i$, where $H_i^\top$ is the row i of the matrix of weights H and b_i is the entry i of the vector of biases b. It can be shown [50, Section IV] that when $b = 0$ (for the case $b \neq 0$ see [50, Section V]) the condition (2.3) holds with $\hat{Q} = 0$, $\hat{S} = \Gamma$, $\hat{R} = -2\Gamma$, where Γ is any symmetric matrix such that for any $i = 1, 2, \ldots, p$, $\gamma_{ij} < 0$ for every $j \neq i$ and $\sum_{j=1}^{p} \gamma_{ij} > 0$ for any i. Hence, without loss of generality, $\hat{R} \prec 0$.

6. *(Norm-bound linear difference inclusion)* System (2.2) falls into the class of the so-called norm-bound linear difference inclusion [51, Chapter 5].

2.2.1. PROBLEM FORMULATION

The problem of interest is to design a control law ensuring absolute stability for the closed-loop system in the event that information about the system is in the form of data samples. In this respect, we assume to collect data of the system through open-loop offline experiments. We use the notation U_0, X_0, X_1 and F_0 to denote the data matrices

$$U_0 := \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix} \tag{2.6a}$$

$$X_0 := \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \tag{2.6b}$$

$$X_1 := \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix} \tag{2.6c}$$

$$F_0 := \begin{bmatrix} f(0,z(0)) & f(1,z(1)) & \cdots & f(T-1,z(T-1)) \end{bmatrix}. \tag{2.6d}$$

We consider the following assumptions:

Assumption 1 *The matrices U_0, X_0, X_1 and F_0 are known.* ■

Assumption 2 *The matrix*

$$W_0 := \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} \tag{2.7}$$

is full row-rank. ■

Before proceeding, we make some remarks.

Assumption 1 means that we can collect input-state samples of the system. We note in particular that F_0 can be measured when the nonlinear block $f(t,Hx)$ is physically detached from the dynamical block $x^+ = Ax + Bu + Lv$, as schematized in Figure 2.1. Besides that, assuming that F_0 is known permits us to establish a clean data-based analogue of absolute stability, as well as a data-based analogue of some related results available for model-based control including the circle criterion and the feedback Kalman-Yakubovitch-Popov Lemma.

Assumption 2 deals instead with the question of richness of data. As discussed next, when this assumption holds then it is possible to express the behavior of (2.2) under a control law $u = Kx$, with K arbitrary, purely in terms of the data matrices in (2.6). It is known that for linear controllable systems this assumption actually reduces to a design condition that can be enforced by suitably choosing U_0, see [2]. The question of how to design experiments so as to enforce this condition for nonlinear systems has been recently addressed in [52]. Ways to relax this assumption will be discussed in Section 2.4.

2.3. LEARNING CONTROL FROM DATA

In this section, we derive data-based conditions for absolute stability. The first step is to provide a data-based representation of the closed-loop system. Under Assumption 2, for any matrix $K \in \mathbb{R}^{m\times n}$ there exists a matrix $G \in \mathbb{R}^{T\times n}$ satisfying

$$\begin{bmatrix} K \\ I_n \end{bmatrix} = W_0 G \tag{2.8}$$

Accordingly, the behavior of (2.2) under a control law $u = Kx$ can be equivalently expressed as

$$\begin{aligned} x^+ &= (X_1 - LF_0)Gx + Lv \\ z &= Hx \\ v &= f(t,z) \end{aligned} \tag{2.9}$$

which follows from the chain of equalities

$$A + BK = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} K \\ I_n \end{bmatrix} = \begin{bmatrix} B & A \end{bmatrix} W_0 G = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} G \tag{2.10}$$

and from $X_1 = AX_0 + BU_0 + LF_0$. In the following, we will use a shorter notation by defining $X_L := X_1 - LF_0$.

We address the absolute stabilizability problem considering quadratic Lyapunov functions $V(x) = x^\top Px$, in which case the problem becomes the one of finding two matrices G and $P \succ 0$ such that

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top X_L^\top P X_L G - P & G^\top X_L^\top PL \\ \star & L^\top PL \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \tag{2.11}$$

holds for all $x \neq 0$ and for all $v = f(t,z)$ that satisfy

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} \geq 0 \tag{2.12}$$

having defined

$$\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} := \begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix}^\top \begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^\top & \hat{R} \end{bmatrix} \begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} \tag{2.13}$$

The following result then holds.

Theorem 1 ***(Data-driven absolute stabilizability)***
Consider the nonlinear system (2.2) *and let the constraint* (2.12) *be regular. Suppose that Assumption 1 and 2 hold.*

Then, there exist two matrices G and $P \succ 0$ such that (2.11) *holds for all $(x, v) \neq 0$ that satisfy* (2.12)

1. *($Q \succeq 0$) if and only if there exists a $T \times n$ matrix Y such that the matrix inequality*

$$\begin{bmatrix} -X_0 Y & X_0 YS & Y^\top X_L^\top & X_0 Y Q^{1/2} \\ \star & R & L^\top & 0 \\ \star & \star & -X_0 Y & 0 \\ \star & \star & \star & -I \end{bmatrix} \prec 0 \tag{2.14}$$

holds;

2 2. *($Q = 0$) if and only if there exists a $T \times n$ matrix Y such that the matrix inequality*

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top X_L^\top \\ \star & R & L^\top \\ \star & \star & -X_0 Y \end{bmatrix} \prec 0 \tag{2.15}$$

holds;

3. *($Q \preceq 0$) if there exists a $T \times n$ matrix Y such that the matrix inequality* (2.15) *holds. In this case, the regularity of* (2.12) *is not needed.*

In all the three cases, a state-feedback matrix K that ensures absolute stability for the closed-loop system can be computed as $K = U_0 Y (X_0 Y)^{-1}$.

Proof.

1. We want

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top X_L^\top P X_L G - P & G^\top X_L^\top P L \\ \star & L^\top P L \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0$$

to hold for all $(x, v) \neq 0$ satisfying condition (2.12), and G must additionally obey $X_0 G = I_n$. We start by focusing on the relation between (2.11) and (2.12). Note that $v = f(t,z)$ with $z = Hx$, so v depends on x. However, this dependence can be neglected and we can equivalently ask that (2.11) holds for all nonzero $(x, v) \in \mathbb{R}^n \times \mathbb{R}^q$ satisfying (2.12), where now v is viewed as a free vector. This is because, as noted in [53, Section 2.1.2], for any vector v satisfying (2.12) there is a function $f(t,z)$ that satisfies (2.12) and passes through that point. Thus, we arrived at a stability condition for which the lossless S-procedure applies. In particular, by [53, Theorem 2.19] a necessary and sufficient condition to have (2.11) fulfilled for all nonzero (x, v) satisfying (2.12) is that there exists a scalar $\tau \geq 0$ such that

$$\begin{bmatrix} G^\top X_L^\top P X_L G - P + \tau Q & G^\top X_L P L + \tau S \\ \star & L^\top P L + \tau R \end{bmatrix} \prec 0 \tag{2.16}$$

holds for some G and $P \succ 0$, where G must additionally obey $X_0 G = I_n$. Without loss of generality[1], let $\tau > 0$, normalize the matrix P ($P/\tau \to P$), and obtain

$$\begin{bmatrix} G^\top X_L^\top P X_L G - P + Q & G^\top X_L^\top P L + S \\ \star & L^\top P L + R \end{bmatrix} \prec 0 \tag{2.17}$$

[1] If $\tau = 0$ the matrix inequality (2.11) never holds since $L^\top P L \succeq 0$.

By Schur complement, the latter is equivalent to

$$\begin{bmatrix} -P+Q & S & G^\top X_L^\top \\ S^\top & R & L^\top \\ X_L G & L & -P^{-1} \end{bmatrix} \prec 0 \tag{2.18}$$

To prevent the simultaneous presence of P and P^{-1}, we factorize $Q = Q^{1/2}Q^{1/2}$, apply the Schur complement another time, left- and right-multiply by block.diag(P^{-1}, I, I, I), so as to obtain (2.14), where $Y = GP^{-1}$. Note in particular that, in view of this change of variable, the constraint $X_0 G = I_n$ has become $X_0 Y = P^{-1}$. In turn, this implies that we can substitute P^{-1} with $X_0 Y$, which is the reason why the LMI (2.14) only depends on Y.

Finally, the relation $K = U_0 G$ gives $K = U_0 Y P = U_0 Y (X_0 Y)^{-1}$.

2. The previous arguments continue to hold until the matrix inequality (2.18), which now holds without the matrix Q (since $Q = 0$). This allows us to directly arrive at the matrix inequality of reduced order (2.15).

3. Since $Q \preceq 0$, it is straightforward to realize that (2.18) is implied by

$$\begin{bmatrix} -P & S & G^\top X_L^\top \\ S^\top & R & L^\top \\ X_L G & L & -P^{-1} \end{bmatrix} \prec 0 \tag{2.19}$$

Hence, (2.15) is a sufficient condition for (2.11) to hold for all $(x, v) \neq 0$ that satisfy (2.12).

■

The matrix inequalities (2.14) or (2.15), are actually *linear* matrix inequalities (LMI) in the decision variable Y. In fact, all the conditions of this chapter are given in terms of linear matrix inequalities and equalities, which are semi-definite programs. These can be solved by standard numerical solvers, such as `cvx` [54].

Remark 1 (Relaxing Assumption 2)
Theorem 1 rests on the assumption that the matrix W_0 is full row rank. It is immediate to see that having X_0 full row rank is actually necessary since, otherwise, (2.14) *cannot have a solution because $X_0 Y$ cannot be positive definite. In contrast,* (2.14) *might have a solution even when U_0 is not full row rank, in line with what has been shown in [55] for linear systems. This happens when there exists a controller K ensuring absolute stability that lies in the column space of U_0. In this sense, Assumption 2 guarantees that all possible controllers are evaluated.* ■

2

2.3.1. DISCUSSION

A few points worth of discussion are in order:

Regularity of (2.3) **for sector bounded nonlinearities.**
The regularity of the constraint (2.3) is satisfied in some notable cases. In case the nonlinearity $f(t,z)$ is decoupled, namely,

$$f(t,z) = \mathrm{col}(f_1(t,z_1),\dots,f_p(t,z_p))$$

and each component satisfy a sector bound constraint, then K_1, K_2 are diagonal matrices and regularity of (2.3) is guaranteed by the condition $K_2 - K_1 \succ 0$, that is the interior of the sectors is non empty.

Exponential stabilizability.
To guarantee exponential convergence of the state to the origin with decay rate $0 < \rho < 1$, it is enough to replace (2.14) with a weak matrix inequality in which the block $(1,1)$ of the matrix on the left-hand side is replaced by $-\rho X_0 Y$. In this case, the search for a solution Y must be preceded by a line search on ρ.

Data-dependent feedback Kalman-Yakubovitch-Popov Lemma.
Theorem 1 can be viewed as a data-dependent feedback Kalman-Yakubovitch-Popov Lemma [43, Section 2.7.4], meaning that it results in a data-dependent feedback design guaranteeing the well-known frequency domain condition of the closed-loop system. In fact, in the proof of Theorem 1 we have shown that the condition (2.14) is equivalent to the existence of $P \succ 0$ such that (2.17) holds. As $Q \succeq 0$, from the block $(1,1)$ of (2.17) we deduce that the matrix $(X_1 - LF_0)G = A + BK$ is Schur stable, hence $\det(e^{i\omega}I - A - BK) \neq 0$ for all $\omega \in \mathbb{R}$ and by the Kalman-Yakubovich-Popov lemma for discrete-time systems [56, Theorem 2], (2.17) implies the frequency domain condition

$$\begin{bmatrix} (e^{i\omega}I - A - BK)^{-1}L \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} (e^{i\omega}I - A - BK)^{-1}L \\ I \end{bmatrix} \prec 0 \tag{2.20}$$

for all $\omega \in \mathbb{R}$, where * denotes the conjugate transpose. Hence, condition (2.14) leads to the existence of a gain matrix K such that the frequency condition (2.20) holds. Conversely, if $\det(e^{i\omega}I - A - BK) \neq 0$ for all $\omega \in \mathbb{R}$ and the matrix inequality holds, then by [56, Theorem 2] there exists a matrix $P = P^\top$ and a real number $\tau \geq 0$ such that (2.17) holds. Note however that there is no guarantee, except in special cases, that $P \succ 0$, and therefore (2.14), which would require a positive definite matrix $X_0 Y = P^{-1}$, cannot be concluded.

Passive nonlinearities.

The analysis of the special case of passive nonlinearities, i.e. $z^\top f(z) \geq 0$ for all $z \in \mathbb{R}^p$, in which case $Q = 0$, $S = H^\top$, $R = 0$, is deferred to Section 2.3.2.

2.3.2. CONTINUOUS-TIME SYSTEMS

One of the features of the data-dependent representation introduced in [4] and here adopted to deal with nonlinear systems, is that it holds for both continuous-time and discrete-time systems thus allowing for a unified analysis and design framework for both classes of systems. In this subsection, we see how Theorem 1 becomes in the case of continuous-time systems. Besides being of interest on its own sake, our motivation is to have a result to be used for some illustrative examples, which are more commonly found for continuous-time systems in the literature.

We start with the data-dependent representation for continuous-time systems, given by

$$\begin{aligned} \dot{x} = &\quad (X_1 - LF_0)Gx + Lv \\ z = &\quad Hx \\ v = &\quad f(t,z) \end{aligned} \tag{2.21}$$

and

$$X_1 = \begin{bmatrix} \dot{x}(t_0) & \dot{x}(t_1) & \dots & \dot{x}(t_{T-1}) \end{bmatrix} \tag{2.22}$$

with t_k, $k = 0, 1, \dots, T-1$, the sampling times at which measurements are taken during the off-line experiment. We assume that f satisfies the standard conditions for the existence and uniqueness of the solution to the feedback interconnection, namely piece-wise continuity in t and local Lipschitz property in z.

Theorem 2 ***(Data-driven absolute stabilizability of continuous-time systems)*** *Consider the nonlinear continuous-time system*

$$\dot{x} = Ax + Bu + Lf(t,z), \quad z = Hx \tag{2.23}$$

Let Assumptions 1 and 2 hold. Let the constraint (2.12) *be regular. There exists two matrices G and* $P \succ 0$ *such that* (2.24) *holds for all* $(x, v) \neq 0$ *that satisfy* (2.12)

1. *(*$Q \succeq 0$*) if and only if there exists a* $T \times n$ *matrix* Y *such that the matrix inequality*

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y & L + X_0 Y S & X_0 Y Q^{1/2} \\ \star & R & 0 \\ \star & \star & -I \end{bmatrix} \prec 0$$

holds.

2. *($Q = 0$) if and only if there exists a $T \times n$ matrix Y such that the matrix inequality*

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y & L + X_0 Y S \\ \star & R \end{bmatrix} \prec 0$$

holds.

3. *($Q \preceq 0$) if there exists a $T \times n$ matrix Y such that the matrix inequality* (2.27) *holds. In this case, the regularity of* (2.12) *is not needed.*

In all the three cases, the matrix K that solves the problem is given by $K = U_0 Y (X_0 Y)^{-1}$.

■

Proof. The proof follows closely the one of Theorem 1. Similar to the discrete case in the previous section, we focus on the existence of a matrix $P \succ 0$ such that

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top X_L^\top P + P X_L G & PL \\ \star & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \tag{2.24}$$

holds for all $x \neq 0$ and for all $v = f(t, z)$ that satisfy (2.12) and obtain a necessary and sufficient condition given by the existence of $P \succ 0$ such that

$$\begin{bmatrix} G^\top X_L^\top P + P X_L G + Q & PL + S \\ \star & R \end{bmatrix} \prec 0 \tag{2.25}$$

holds.

1. In case $Q \succeq 0$, starting from (2.25) using similar manipulations as in Theorem 1 return the inequality

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y & L + X_0 Y S & X_0 Y Q^{1/2} \\ \star & R & 0 \\ \star & \star & -I \end{bmatrix} \prec 0 \tag{2.26}$$

where $Y = GP^{-1}$. The relation $K = U_0 G$ gives the control gain $K = U_0 Y (X_0 Y)^{-1}$.

2. In the case $Q = 0$, we obtain from (2.26) the simpler condition

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y & L + X_0 Y S \\ \star & R \end{bmatrix} \prec 0, \tag{2.27}$$

3. The condition (2.27) is also a sufficient condition for the data-dependent absolute stabilizability of the continuous-time system when $Q \prec 0$.

■

An important special case is that of passive nonlinearities, namely $z^\top f(t,z) \geq 0$ for all z, which corresponds to the case where f belongs to the sector $[0,\infty]$ [44, Definition 6.2]. Passive nonlinearities can be written in the form (2.12) letting $Q=0$, $S=H^\top$ and $R=0$. Since $R=0$, this case does not directly fall in the previous analysis. However, it is an easy matter to see that, in this case, a sufficient data-dependent condition for the absolute stabilizability via linear feedback $u=Kx$ of (2.21) amounts to the existence of a matrix Y such that

$$\begin{aligned} & X_0 Y \succ 0 \\ & Y^\top X_L^\top + X_L Y \prec 0 \\ & L + X_0 Y H^\top = 0 \end{aligned} \tag{2.28}$$

2

If a solution to (2.28) exists then the matrix K that solves the problem is given by $K = U_0 Y (X_0 Y)^{-1}$. In fact, recalling that $A+BK = (X_1 - LF_0)G = X_L G$, condition (2.28) can be recognized as a data-dependent condition for the *strict positive realness* [44, Lemma 6.3] of the closed-loop system $(H, A+BK, L)$, where the constraint $(A+BK)^\top P + P(A+BK) \prec 0$, $P \succ 0$, is written in the equivalent form

$$Y^\top X_L^\top + X_L Y \prec 0$$

introducing the change of variable $Y = GP^{-1}$, which implies the identity $P^{-1} = X_0 Y$ because of the constraint $X_0 G = I$. Condition (2.28), in turn, is a sufficient condition for absolute stability under passive nonlinearities [44, Theorem 7.1], the so-called *multivariable circle criterion.*

Remark 2 (Inferring open-loop properties from data-driven design)
Condition (2.28) *is also the data-dependent version of a well-known passifiability condition [43, Theorem 2.12]: if L is full column rank, there exists a feedback controller $u = Kx$ which makes the triple $(H, A+BK, L)$ state strictly passive* if and only if *the system defined by the triple (H, A, B) is minimum phase and the matrix $HL \prec 0$. Since H, L are part of our prior knowledge, the condition $HL \prec 0$ can be checked. Hence, if the inequality* (2.28) *is feasible, we infer the property of the open-loop triple (H, A, B) being minimum phase without explicitly knowing the matrices A, B but rather relying on the data X_0, X_1, F_0. Using conditions for data-driven control to infer properties of an open-loop system deserves further attention in future work.* ■

Example 1 *We introduce an example to illustrate the application of the results in this section. In particular, we focus on the condition* (2.28). *We consider a*

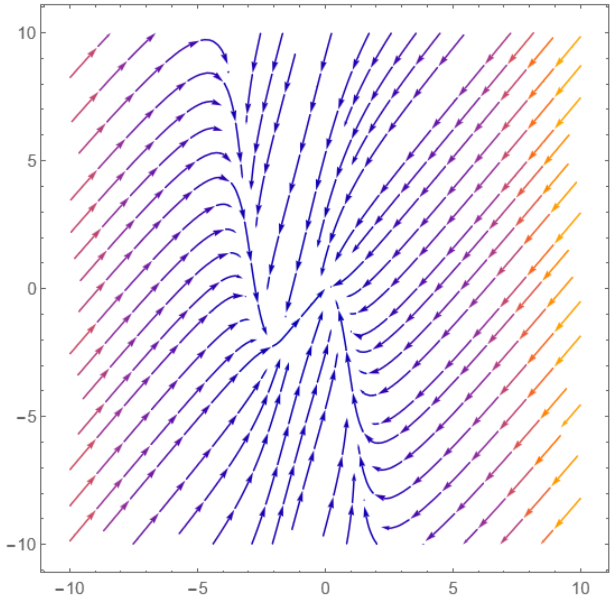

Figure 2.2 | Phase portrait of system (2.29) in Example 1 under the designed feedback control $u = [4.3339\ \ -3.7435]\, x$. The solutions asymptotically converge to the origin.

pre-compensated surge subsystem of an axial compressor model, see e.g. [57]

$$\dot{x} = \begin{bmatrix} \frac{9}{8} & -1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} -1 \\ -\beta \end{bmatrix} \varphi(x_1) \tag{2.29}$$

with $\beta > 9/8$ *a parameter and* φ *a passive nonlinearity such that* $z\varphi(z) \geq 0$. *Specifically,* $\varphi(z) = \frac{1}{2}z^3 + \frac{3}{2}z^2 + \frac{9}{8}z$. *Hence, for this example, we observe that a precise knowledge of* L *is not required, any estimate* $\hat{L} = \alpha \begin{bmatrix} -1 & -\beta \end{bmatrix}^\top$, *with* $\alpha > 0$, *used in* (2.28) *does not affect the outcome of the design. We perform an open-loop experiment from the initial condition* $x(0) = \begin{bmatrix} 2 & -1 \end{bmatrix}^\top$, *with* $\alpha = 1$, $\beta = 1.2$, *under the input* $u(t) = \sin t$ *over the time horizon* $[0, 1]$ *using* $T = 5$ *evenly spaced sampling times, and collect the measurements in the matrices* U_0, X_0, X_1, F_0:

$$U_0 = \begin{bmatrix} 0 & 0.2474 & 0.4794 & 0.6816 & 0.8415 \end{bmatrix}$$

$$X_0 = \begin{bmatrix} 2 & 1.269 & 1.3208 & 1.5113 & 1.7451 \\ -1 & -2.993 & -4.3724 & -6.0225 & -8.2189 \end{bmatrix}$$

$$X_1 = \begin{bmatrix} -21.25 & -5.309 & -4.6511 & -5.9817 & -8.1951 \\ -29.4 & -11.428 & -12.1319 & -15.7636 & -21.2112 \end{bmatrix}$$

$$F_0 = \begin{bmatrix} 12.25 & 4.8648 & 5.2547 & 6.8522 & 9.1886 \end{bmatrix}$$

Assumption 2 holds. We replace the data matrices in (2.28) *along with* $\hat{L} = \alpha \begin{bmatrix} -1 & -\beta \end{bmatrix}^\top$, *having set* $\alpha = 2$ *and* $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$. *We remark that the parameter* α *used in* (2.28)

is different from the value used during the experiment to stress that the precise knowledge of L is not needed.
We solve (2.28) *with* `cvx` *[54] for* Y*, and obtain*

$$Y = \begin{bmatrix} 1.2922 & 1.6018 \\ -0.1923 & 1.0528 \\ 0.5113 & -0.5863 \\ 0.5192 & -1.1827 \\ -1.0316 & 0.2419 \end{bmatrix}$$

from which

$$K = U_0 Y (X_0 Y)^{-1} = \begin{bmatrix} 4.3339 & -3.7435 \end{bmatrix}$$

We observe that

$$Y^\top X_L^\top + X_L Y = \begin{bmatrix} -23.6176 & -30.3340 \\ -30.3340 & -39.1227 \end{bmatrix} \prec 0$$

and the entries of $\alpha L + X_0 Y H^\top$*, with* $\alpha = 2$*, are of order* 10^{-12}*, thus* $2x^\top P((A + BK)x + Lf(z)) < 0$ *for all* x*, which guarantees asymptotic stability uniformly with respect to any passive nonlinearity* f*. A phase portrait of the closed-loop system is shown in Fig. 2.2. Finally, we observe that should the nonlinearity* f *be time-varying, i.e.* $f(t,z)$ *during the experiment and different from the one appearing in the dynamics when the control is applied, the same result of uniform asymptotic stability will continue to hold as long as* $z^\top f(t,z) \geq 0$ *during the experiment and in the closed-loop system.* ■

Before concluding the section, we observe that there has been some recent interest on promoting stability of systems with energy preserving quadratic nonlinearities [58], also with data-driven methods [59]. These systems are of the form (2.1) with $B = 0$ and $\hat{f}(t,x)$ independent of t and such that

$$\hat{f}(x) = \begin{bmatrix} x^\top Q_1 x \\ \vdots \\ x^\top Q_n x \end{bmatrix} \text{ with } x^\top \hat{f}(x) = 0$$

and Q_i, for $i = 1,\dots,n$, constant matrices. This class of systems falls into the category of systems with passive nonlinearities and $-L = H = I_n$, provided that the term Bu is added. However, the problem considered in [58], [59], namely unveiling a locally asymptotically stable point by shifting the state variables, is different from the one considered here of making the system globally asymptotically stable by feedback.

Giving up the knowledge about L, H is a difficult task. In the next section, we examine one possibility based on strengthening the requirement on the collected data.

2.4. RELAXING SOME PRIOR KNOWLEDGE BY STRENGTHENED DATA ASSUMPTIONS

The last example has shown the difficulty to relax the knowledge about the matrices L, H, which influence how the nonlinearity affects the dynamics and which state variables appear in the nonlinear function. The situation dramatically changes as far as L is concerned if we consider a stronger assumption on the set of available data. We also examine how to use nonlinear feedback. Specifically, we use the term $f(t,z)$, measured for all time t, in the design of the feedback control.

As remarked in Section 2.2, real time knowledge of $f(t,z)$ is justified in those case in which the term $f(t,z)$ appears as a physically detached block whose output can be measured. In model-based absolute stability theory, the case in which the nonlinearity f is unknown but the signal $f(t,z)$ is available for on-line measurements has been considered in [57]. Alternatively, the term $Lf(t,z(t))$ can originate from modeling the nonlinearity via a vector of known regressors f and a matrix of unknown coefficients L, as classically done in nonlinear adaptive control [60]. This is also the point of view taken in recent papers that combine sparsity-promoting techniques and machine learning [61]. Here, however, since we are not interested in estimating the dynamics but directly controlling it, we do not need to assume to know the analytic expression of f.

If we can measure in real-time $f(t,z)$, then we can use it also in the feedback policy, along with the state $x(t)$. Hence, here we consider the case in which the system (2.2) is controlled via the feedback

$$u(t) = Kx(t) + Mf(t,z(t)), \quad z(t) = Hx(t) \tag{2.30}$$

where K, M are matrices to design. Again we stress that the feedback gains K, M are to be designed without knowing the analytic expression of f nor A, B, L but only the real time measurements of the vector $x(t)$ and $f(t,z)$ The matrix H must be known since it appears in the matrix Q (see (2.13)), which in turn appears in the LMI conditions that we give below.

Since the matrix F_0 in (2.6d) is known, along with X_0, X_1, U_0, we take advantage of this knowledge by revising Assumption 2 as follows:

Assumption 3 *The matrix*

$$\Psi_0 := \begin{bmatrix} X_0 \\ F_0 \\ U_0 \end{bmatrix}$$

is full-row rank.

For any matrix $\begin{bmatrix} K & M \end{bmatrix} \in \mathbb{R}^{m\times(n+q)}$, we let the matrix $G = \begin{bmatrix} G_1 & G_2 \end{bmatrix} \in \mathbb{R}^{T\times(n+q)}$, where G_1 has n columns and G_2 q columns, satisfy

$$\begin{bmatrix} I_n & 0_{n\times q} \\ 0_{q\times n} & I_q \\ K & M \end{bmatrix} = \begin{bmatrix} X_0 \\ F_0 \\ U_0 \end{bmatrix} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \tag{2.31}$$

Then we obtain the relation

$$\begin{aligned} \begin{bmatrix} A & L \end{bmatrix} + B\begin{bmatrix} K & M \end{bmatrix} = & \left[\begin{array}{cc|c} A & L & B \end{array}\right] \begin{bmatrix} I_n & 0_{n\times q} \\ 0_{q\times n} & I_q \\ K & M \end{bmatrix} \\ = & \left[\begin{array}{cc|c} A & L & B \end{array}\right] \left[\begin{array}{c} X_0 \\ F_0 \\ \hline U_0 \end{array}\right] \begin{bmatrix} G_1 & G_2 \end{bmatrix} \\ = & X_1 \begin{bmatrix} G_1 & G_2 \end{bmatrix} \end{aligned}$$

where we have exploited the identity $X_1 = AX_0 + BU_0 + LF_0$. We conclude that system (2.2) in closed-loop with the nonlinear feedback (2.30) is equivalent to the nonlinear data-dependent system

$$\begin{aligned} x^+ = & X_1 \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} x \\ f(t,z) \end{bmatrix} \\ = & X_1 G_1 x + X_1 G_2 f(t,z) \\ z = & Hx \end{aligned} \tag{2.32}$$

with matrices G_1, G_2 that satisfy (2.31). We now study the absolute stability of such data-dependent system under the quadratic constraint assumption. We only state the result in the case $Q \succeq 0$ since the other cases are immediately obtained. As before, we address the problem considering quadratic Lyapunov functions $V(x) = x^\top P x$, so that the problem becomes the one of the existence of a symmetric positive definite matrix P such that (2.33) holds (cf. (2.11)).

Theorem 3 ***(Data-driven absolute stabilizability II)***
Consider the nonlinear system (2.2) *and let the constraint* (2.12) *be regular. Let Assumptions 1 and 3 hold and let* $Q \succeq 0$. *There exists three matrices* G_1, G_2 *and* $P \succ 0$ *such that*

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G_1^\top X_1^\top P X_1 G_1 - P & G_1^\top X_1^\top P X_1 G_2 \\ \star & G_2^\top X_1^\top P X_1 G_2 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \tag{2.33}$$

2

holds for all $(x, v) \neq 0$ *that satisfy* (2.12) *if and only if there exist* $T \times n$, $T \times q$ *and* $n \times n$ *matrices* Y_1, Y_2, W *such that the conditions*

$$\begin{bmatrix} -W & WS & Y_1^\top X_1^\top & WQ^{1/2} \\ \star & R & Y_2^\top X_1^\top & 0 \\ \star & \star & -W & 0 \\ \star & \star & \star & -I_n \end{bmatrix} \prec 0$$
$$\begin{bmatrix} X_0 Y_1 - W & X_0 Y_2 \\ F_0 Y_1 & F_0 Y_2 - I_q \end{bmatrix} = 0 \tag{2.34}$$

hold. In this case, the matrices K, M *that solve the problem are given by* $K = U_0 Y_1 W^{-1}$ *and* $M = U_0 Y_2$.

Proof. Repeating the same analysis as in the proof of Theorem 1 but this time for the representation (2.32), we obtain the counterpart of (2.16), which is

$$\begin{bmatrix} G_1^\top X_1^\top P X_1 G_1 - P + Q & G_1^\top X_1^\top P X_1 G_2 + S \\ \star & G_2^\top X_1^\top P X_1 G_2 + R \end{bmatrix} \prec 0 \tag{2.35}$$

where $P \succ 0$ is to be determined, and we have carried out the normalization $\frac{P}{\tau} \to P$. The same manipulations that followed (2.16) lead in this case to

$$\begin{bmatrix} -P & S & G_1^\top X_1^\top & Q^{1/2} \\ S^\top & R & G_2^\top X_1^\top & 0 \\ \star & \star & -P^{-1} & 0 \\ \star & \star & \star & -I_n \end{bmatrix} \prec 0$$

By pre- and post-multiplying the matrix above by the matrix block.diag(P^{-1}, I, I, I) we obtain (2.34) having set $W := P^{-1}$, $Y_1 := G_1 P^{-1}$, $Y_2 := G_2$. Isolating the equation

$$\begin{bmatrix} I_n & 0_{n\times q} \\ 0_{q\times n} & I_q \end{bmatrix} = \begin{bmatrix} X_0 \\ F_0 \end{bmatrix} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \tag{2.36}$$

in (2.31), taking its transpose and multiplying it on the left by block.diag(P^{-1}, I_p), we obtain

$$\begin{aligned} \begin{bmatrix} P^{-1} & 0 \\ 0 & I_q \end{bmatrix} &= \begin{bmatrix} P^{-1} G_1^\top X_0^\top & P^{-1} G_1^\top F_0^\top \\ G_2^\top X_0^\top & G_2^\top F_0^\top \end{bmatrix} \\ &= \begin{bmatrix} Y_1^\top X_0^\top & Y_1^\top F_0^\top \\ Y_2^\top X_0^\top & Y_2^\top F_0^\top \end{bmatrix} \end{aligned}$$

that is, the constraints (2.31) expressed in the variables Y_1, Y_2, W. In particular, since $P^{-1} = X_0 Y_1$, we have $X_0 Y_1 \succ 0$. Moreover, by $\begin{bmatrix} K & M \end{bmatrix} = U_0 \begin{bmatrix} G_1 & G_2 \end{bmatrix}$, we obtain $K = U_0 G_1 = U_0 Y_1 P = U_0 Y_1 (X_0 Y_1)^{-1}$ and $M = U_0 Y_2$. ∎

Example 2 *We consider a slightly revised version of Example 1 given by*

$$\dot{x} = \begin{bmatrix} \frac{9}{8} & -1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \varphi(x_1)$$

where the nonlinearity $\varphi(x_1)$ is defined as before.Compared with (2.29)*, the term $-\beta\varphi(x_1)$ is missing from the dynamics. In fact, it will be shown that the feasibility of* (2.34) *in Theorem 3 leads to correctly select from data the parameters in the nonlinear controller* (2.30) *to provide such a term. The condition* (2.34) *in the case of passive nonlinearities for continuous-time systems is obtained via straightforward modifications of* (2.28)*, and returns the following condition: there exist $T \times n$ and $T \times q$ matrices Y_1, Y_2 such that*

$$\begin{aligned} & Y_1^\top X_1^\top + X_1 Y_1 \prec 0 \\ & X_1 Y_2 + X_0 Y_1 H^\top = 0 \\ & X_0 Y_1 \succ 0 \\ & X_0 Y_2 = 0 \\ & F_0 Y_2 = I_p \\ & F_0 Y_1 = 0 \end{aligned} \tag{2.37}$$

We consider the same experiment as in Example 1: initial condition $x(0) = \begin{bmatrix} 2 & -1 \end{bmatrix}^\top$, $\alpha = 2$, and input $u(t) = \sin t$ over the time horizon $[0,1]$. We take $T = 10$ evenly spaced sampling times. We collect the measurements in the matrices U_0, X_0, X_1, F_0, which we do not report here for the sake of brevity. It can be checked that Assumption 3 is satisfied. We obtain the solution

$$\begin{bmatrix} Y_1 & \vline & Y_2 \end{bmatrix} = \left[\begin{array}{rr|r} 0.9823 & -3.5073 & -5.6005 \\ -2.0064 & 8.5180 & 12.8729 \\ -1.3370 & 7.1478 & 10.2375 \\ 0.41465 & 3.1658 & 2.6801 \\ 2.2302 & -0.2915 & -4.4256 \\ 3.5496 & -3.1425 & -10.2866 \\ 3.7054 & -4.8273 & -12.6223 \\ 2.2325 & -4.4031 & -9.1124 \\ -0.7529 & -1.5849 & -0.4900 \\ -6.3569 & 3.4286 & 16.4407 \end{array}\right]$$

from which we compute the feedback gains

$$K = \begin{bmatrix} 7.0779 & -3.9230 \end{bmatrix}, \quad M = -3.5130$$

and the Lyapunov matrix

$$P = (X_0 Y_1)^{-1} = \begin{bmatrix} 4.1628 & -2.0853 \\ -2.0853 & 1.1872 \end{bmatrix}$$

2

which satisfies the Lyapunov inequality

$$Y_1^\top X_1^\top + X_1 Y_1 = \begin{bmatrix} -2.5259 & -2.6865 \\ -2.6865 & -5.2943 \end{bmatrix} \prec 0$$

and the condition $P(L+BM) = (X_0 Y_1)^{-1} X_1 Y_2 = -H^\top$. *We observe that the program* (2.37) *is able to correctly compute from data that the gain* M *satisfies* $M < -9/8$, *which is a necessary condition for feedback* (2.30) *to render the closed-loop system strictly positive real [57, Example 1].* ■

Remark 3 *Identities* (2.31) *suggest a way to renounce to the knowledge of L without resorting to a nonlinear feedback involving* $f(t,z)$. *This can be achieved by imposing* $M = 0$ *in* (2.31), *which amounts to adding the constraint* $0 = U_0 G_2$ *to* (2.34). *Under such conditions, we conclude that Theorem 3 holds when the feedback is the* linear *one* $u = Kx$. *With respect to the case where L is known, the price to pay is that we need Assumption 3 instead of Assumption 2, which is less stringent.* ■

Example 3 *To illustrate the previous remark, we consider Example 1 again,*[2] *this time however without assuming that the matrix L in* (2.29) *is known. In fact, differently from Example 1 where we employed* (2.28), *here we solve* (2.37) *with the addition of* $0 = U_0 G_2$. *We use the same data* X_0, X_1, U_0, F_0 *as in Example 1. We observe that* $\begin{bmatrix} X_0^\top & F_0^\top & U_0^\top \end{bmatrix}$ *is full row rank. We obtain*

$$K = U_0 Y (X_0 Y)^{-1} = \begin{bmatrix} 35.8066 & -2.1645 \end{bmatrix}$$

which makes the closed-loop matrix $A + BK$ *Hurwitz, with Lyapunov matrix*

$$P = (X_0 Y_1)^{-1} = \begin{bmatrix} 0.5217 & -0.0181 \\ -0.0181 & 0.015 \end{bmatrix}$$

which satisfies $PL + H^\top = 0$. ■

2.5. CONCLUSIONS

We have presented a purely data-driven solution to derive a state feedback controller to stabilize systems with quadratic nonlinearities, providing necessary and sufficient conditions for the absolute stabilizability of the closed-loop system. We have discussed several variants of the results under different feedback (linear and nonlinear) and strengthened conditions on the data used for the design. To focus on the impact of nonlinearities satisfying a quadratic constraint in the data-dependent control design, we first considered in this chapter noiseless data. The

[2]We do not use the system in Example 2 because it cannot be stabilized by a linear feedback [57].

addition of process disturbances during the acquisition of data will be discussed in Chapter 3. The proposed conditions consist of semi-definite programs that depend on input-state data only and once solved they directly return controllers. As such, they represent end-to-end solutions to learning control from data for an important class of nonlinear systems.

3

STABILIZATION OF SYSTEMS WITH NONLINEARITIES SATISFYING QUADRATIC CONSTRAINTS WITH DISTURBED DATA

ABSTRACT

In the previous chapter, we presented a complete data-driven design for a feedback controller that solve the problem of absolute stability for nonlinear system with quadratic constrains. To simplify the presentation, we have disregarded the presence of noise in the data used for the design, noise that in real applications is often not negligible. This chapter continues the discussion on the absolute stability problem by considering the data noise in the design of the controller. Necessary and sufficient conditions are derived to guarantee absolute stability and noise robustness for both discrete-time and continuous-time systems.

This chapter has been published in "On data-driven stabilization of systems with nonlinearities satisfying quadratic constraints."Luppi, Alessandro, Claudio De Persis, and Pietro Tesi; Systems & Control Letters 163 (2022): 105206.

3

3.1. DATA-DRIVEN ABSOLUTE STABILIZABILITY WITH DISTURBED DATA

We now examine the case in which the system is affected by disturbances during the process of collecting data, namely we focus on the following perturbed system

$$\begin{aligned} x^+ &= Ax + Bu + Lv + Ed \\ z &= Hx \\ v &= f(t,z) \end{aligned} \tag{3.1}$$

In this system, $d \in \mathbb{R}^s$ is an unknown signal representing process disturbances affecting the dynamics during the data collection phase and $E \in \mathbb{R}^{n\times s}$ is a known matrix describing the way d affects the different states of the system. If such an information is not available, then $E = I_n$ and $s = n$.

In the presence of d, in addition to the matrices U_0, X_0, X_1 and F_0 in (2.6), one introduces the matrix D_0 collecting the disturbance samples, i.e.

$$D_0 := \begin{bmatrix} d(0) & d(1) & \cdots & d(T-1) \end{bmatrix} \tag{3.2}$$

which, in contrast to U_0, X_0, X_1, F_0, is unknown. These matrices of data satisfy the identity

$$X_1 = AX_0 + BU_0 + LF_0 + D_0.$$

Based on this identity, under Assumption 2, and for any matrices $K \in \mathbb{R}^{m\times n}$ and $G \in \mathbb{R}^{T\times n}$ satisfying (2.8), one obtains that system (3.1) in closed-loop with the control law $u = Kx$ can be equivalently expressed as

$$\begin{aligned} x^+ &= (X_1 - LF_0 - ED_0)Gx + Lv \\ z &= Hx \\ v &= f(t,z) \end{aligned} \tag{3.24}$$

Note that we assume to neglect the effect of the disturbances during the execution of the control task, as this would require to make the nonlinear system stable with respect to external perturbations, which is considerably more difficult than the absolute stabilizability problem considered here. We note, however, that global asymptotic stability implies input-to-state stability with “small disturbances" [62, Theorem 2].

To address the absolute stabilizability problem for system (3.24), we follow an analogous line of arguments as in the case of disturbance-free measurements treated in Chapter 2, marking the differences due to the presence of disturbances in due course. Hence, we consider a quadratic Lyapunov function $V(x) = x^\top P x$ and look for matrices $P \succ 0$ and G such that

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top (X_L - ED_0)^\top P(X_L - ED_0)G - P & G^\top (X_L - ED_0)^\top PL \\ \star & L^\top PL \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \quad (3.25)$$

holds for all $x \neq 0$ and for all $v = f(t,z)$ that satisfy

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} \geq 0. \quad (3.26)$$

Notice that the difference of (3.25) in comparison with the noiseless case

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top X_L^\top P X_L G - P & G^\top X_L^\top PL \\ \star & L^\top PL \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0$$

lies in the presence of the unknown matrix D_0. Retracing the same steps as in the proof of Theorem 1, we arrive at the equivalent condition that there exists a $T \times n$ matrix Y such that

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top (X_L - ED_0)^\top & X_0 Y Q^{1/2} \\ \star & R & L^\top & 0 \\ \star & \star & -X_0 Y & 0 \\ \star & \star & \star & -I \end{bmatrix} \prec 0 \quad (3.27)$$

where $X_L := X_1 - LF_0$.
In contrast to the noiseless case (from Theorem 1)

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top X_L^\top & X_0 Y Q^{1/2} \\ \star & R & L^\top & 0 \\ \star & \star & -X_0 Y & 0 \\ \star & \star & \star & -I \end{bmatrix} \prec 0,$$

condition (3.27) is not implementable due to the presence of the unknown term ED_0. Without any condition on the disturbance affecting the measurements it is hard to give an implementable condition of absolute stabilizability. We will consider the disturbance as unknown-but-bounded, a solution that has a long history in the literature, as the main idea dates back to the works of [63] [64]. Hence, we introduce an energy bound with the following condition on the disturbance matrix D_0

$$D_0 \in \mathscr{D} := \{D \in \mathbb{R}^{n \times T} : DD^\top \preceq \Delta\Delta^\top\}, \quad (3.28)$$

where Δ is some known matrix. This condition allows one to get rid of the dependence on D_0 via the following matrix inequality, which is a consequence of a completion-of-the-squares argument. Given matrices Γ, Θ, for any D_0 satisfying (3.28) and any $\varepsilon > 0$, it holds that

$$\Gamma D_0 \Theta + \Theta^\top D_0^\top \Gamma^\top \preceq \varepsilon \Gamma \Delta \Delta^\top \Gamma^\top + \varepsilon^{-1} \Theta^\top \Theta.$$

Using this inequality, we obtain inequality

$$\begin{bmatrix} 0 & 0 & -Y^\top D_0^\top E^\top & 0 \\ 0 & 0 & 0 & 0 \\ -ED_0 Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} \varepsilon^{-1} Y^\top Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon E \Delta \Delta^\top E^\top & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.29}$$

Hence, (3.27) is implied by the existence of a $T \times n$ matrix Y and a constant $\varepsilon > 0$ such that

$$\begin{bmatrix} -X_0 Y + \varepsilon^{-1} Y^\top Y & X_0 Y S & Y^\top X_L^\top & X_0 Y Q^{1/2} \\ \star & R & L^\top & 0 \\ \star & \star & -X_0 Y + \varepsilon E \Delta \Delta^\top E^\top & 0 \\ \star & \star & \star & -I \end{bmatrix} \prec 0 \tag{3.30}$$

holds. Yet another application of the Schur complement allows us to get rid of the product $Y^\top Y$ in block $(1,1)$ and obtain

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top X_L^\top & X_0 Y Q^{1/2} & Y^\top \\ \star & R & L^\top & 0 & 0 \\ \star & \star & -X_0 Y + \varepsilon E \Delta \Delta^\top E^\top & 0 & 0 \\ \star & \star & \star & -I & 0 \\ \star & \star & \star & \star & -\varepsilon I \end{bmatrix} \prec 0. \tag{3.31}$$

We formalize the argument as follows.

Theorem 4 ***(Data-driven absolute stabilizability under noisy measurements)*** *Consider the nonlinear system* (3.1)*. Suppose that Assumption 1 and 2 hold. Then, there exist two matrices G and $P \succ 0$ such that* (3.25) *holds for all $(x, v) \neq 0$ that satisfy* (3.26) *and D_0 that satisfy* (3.28)

1. *($Q \succeq 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality* (3.31) *holds;*

2. *($Q = 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality*

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top X_L^\top & Y^\top \\ \star & R & L^\top & 0 \\ \star & \star & -X_0 Y + \varepsilon E \Delta \Delta^\top E^\top & 0 \\ \star & \star & \star & -\varepsilon I \end{bmatrix} \prec 0 \tag{3.32}$$

holds;

3. *($Q \preceq 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality* (3.32) *holds.*

In all the three cases, a state-feedback matrix K that ensures absolute stability for the closed-loop system can be computed as $K = U_0 Y (X_0 Y)^{-1}$.

Proof.

1. It has been already proven, see steps from (3.25) to (3.31).

2. Again, tracing the steps of the proof of Theorem 1, we arrive at the following equivalent condition to (3.25), namely there exist matrices $P \succ 0$ and G such that

$$\begin{bmatrix} -P & S & G^\top (X_L - ED_0)^\top \\ \star & R & L^\top \\ \star & \star & -P^{-1} \end{bmatrix} \prec 0 \tag{3.33}$$

which is a modified version of (2.18). From (3.33), by left- and right-multiplying by block.diag(P^{-1}, I, I), we obtain

$$\begin{bmatrix} -X_0 Y & X_0 Y S & Y^\top (X_L - ED_0)^\top \\ \star & R & L^\top \\ \star & \star & -X_0 Y \end{bmatrix} \prec 0 \tag{3.34}$$

where $Y = GP^{-1}$ as before. Bearing in mind the argument that led to (3.29), we obtain that (3.33) is implied by

$$\begin{bmatrix} -X_0 Y + \varepsilon^{-1} Y^\top Y & X_0 Y S & Y^\top X_L^\top \\ \star & R & L^\top \\ \star & \star & -X_0 Y + \varepsilon E \Delta \Delta^\top E^\top \end{bmatrix} \prec 0 \tag{3.35}$$

which is equivalent to (3.32) by an application of Schur complement.

3. Since $Q \preceq 0$, condition (3.33) implies (3.25), hence, in view of the arguments used to prove the previous point, condition (3.32) is sufficient for (3.25) to hold.

■

Note that we have removed the requirement on the regularity of the constraint (3.26) since the conditions are sufficient.

3.1.1. CONTINUOUS-TIME SYSTEMS WITH DISTURBED DATA

In this section, we preset the case of a continuous-time system with an additive noise in the measurements. Consider the nonlinear continuous-time system affected by perturbations in the form of an additive process disturbance d, i.e.

$$\dot{x} = Ax + Bu + Lf(t,z) + Ed, \qquad z = Hx \tag{3.36}$$

Then the data-dependent closed-loop system representation is

$$\dot{x} = (X_1 - LF_0 - ED_0)Gx + Lf(t,z) + Ed, \qquad z = Hx \tag{3.37}$$

with $D_0 := \begin{bmatrix} d(t_0) & d(t_1) & \cdots & d(t_{T-1}) \end{bmatrix}$ as in (3.2).

Theorem 5 ***(Data-driven absolute stabilizability of continuous-time systems with disturbed data)***
Consider the nonlinear continuous-time system (3.36). *Let Assumptions 1 and 2 hold. There exists two matrices G and $P > 0$ such that*

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G^\top (X_L - ED_0)^\top P + P(X_L - ED_0)G & PL \\ \star & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \tag{3.38}$$

holds for all $(x, v) \neq 0$ that satisfy (3.26) *and D_0 that satisfy* (3.28)

1. *($Q \succeq 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality*

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y + \varepsilon E \Delta \Delta^\top E^\top & L + X_0 Y S & X_0 Y Q^{1/2} & Y^\top \\ \star & R & 0 & 0 \\ \star & \star & -I & 0 \\ \star & \star & \star & -\varepsilon I \end{bmatrix} \prec 0 \tag{3.39}$$

   holds.

2. *($Q = 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality*

$$\begin{bmatrix} Y^\top X_L^\top + X_L Y & L + X_0 Y S & Y^\top \\ \star & R & 0 \\ \star & \star & -\varepsilon I \end{bmatrix} \prec 0 \tag{3.48}$$

   holds.

3. *($Q \preceq 0$) if there exists a $T \times n$ matrix Y and a scalar $\varepsilon > 0$ such that the matrix inequality* (3.48) *holds.*

In all the three cases, the matrix K that solves the problem is given by $K = U_0 Y (X_0 Y)^{-1}$.

■

The proof descends from minor modification of previous arguments (Theorem 2) and is therefore omitted.

3.1.2. PASSIVE NONLINEARITIES

In the case of passive nonlinearities, a sufficient data-dependent condition for the absolute stabilizability via linear feedback $u = Kx$ amounts to the existence of a matrix Y and a scalar $\varepsilon > 0$ such that

$$\begin{aligned} & X_0 Y \succ 0 \\ & \begin{bmatrix} Y^\top X_L^\top + X_L Y + \varepsilon E \Delta \Delta^\top E^\top & Y^\top \\ Y & -\varepsilon I \end{bmatrix} \prec 0 \\ & L + X_0 Y H^\top = 0 \end{aligned} \tag{3.49}$$

If a solution to (3.49) exists then the matrix K that solves the problem is given by $K = U_0 Y (X_0 Y)^{-1}$. In the following, we apply this result to a numerical example.

Example 4 *We revisit Example 1 from Chapter 2 and consider disturbances on the system's dynamics, hence obtaining*

$$\dot{x} = \begin{bmatrix} \frac{9}{8} & -1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} -1 \\ -\beta \end{bmatrix} \varphi(x_1) + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{3.50}$$

where β, φ are as before and $d := \begin{bmatrix} d_1 & d_2 \end{bmatrix}^\top$ is the vector of disturbances. Note that for this example $E = I_2$. The entire setup of Example 1 is unchanged except for the disturbance addition. In particular, an experiment is performed from the same initial condition and applying the same input.

As for d, at the T sampling times we generate values for d_1, d_2 uniformly distributed in the interval $[-\delta/\sqrt{n}, \delta/\sqrt{n}]$, with $\delta = 0.01$, $n = 2$, which are interpolated to emulate a continuous-time disturbance signal. Below is the occurrence of D_0 used in the simulations of this example

$$D_0 = \begin{bmatrix} -0.0012 & 0.0031 & -0.0071 & -0.0028 & -0.0050 \\ -0.0058 & -0.0044 & -0.0022 & -0.0015 & 0.0005 \end{bmatrix}$$

D_0 is unknown to the designer. We set $\Delta := \delta\sqrt{T} I_2 = 0.0158\, I_2$. With such a choice, $D_0 D_0^\top \preceq \Delta\Delta^\top$. We collect the measurements in the matrices U_0, X_0, X_1, F_0, with U_0 as in Example 1 and

$$\begin{aligned} X_0 &= \begin{bmatrix} 2 & 1.2693 & 1.3208 & 1.5113 & 1.7452 \\ -1 & 2.9946 & -4.3760 & -6.0270 & -8.2248 \end{bmatrix} \\ X_1 &= \begin{bmatrix} -21.2512 & -5.3085 & -4.6550 & -5.9798 & -8.1966 \\ -29.4058 & -11.4380 & -12.1346 & -15.7649 & -21.2137 \end{bmatrix} \\ F_0 &= \begin{bmatrix} 12.2500 & 4.8671 & 5.2549 & 6.8521 & 9.1899 \end{bmatrix} \end{aligned}$$

Assumption 2 holds. We solve (3.49) *with* cvx *[54] for ε and Y, and obtain $\varepsilon =$* 90.9064

$$Y = \begin{bmatrix} 1.3812 & 1.5803 \\ -0.1661 & 3.7288 \\ 0.4856 & -2.8528 \\ -0.1946 & -4.5080 \\ -0.5150 & 2.9152 \end{bmatrix}$$

from which

$$K = U_0 Y (X_0 Y)^{-1} = \begin{bmatrix} 14.3773 & -12.1370 \end{bmatrix}$$

It can be checked that the closed-loop matrix $(X_L - ED_0)G$, with $G = Y(X_0 Y)^{-1}$, is Hurwitz (eigenvalues equal to $-5.2460, -12.0643$*). The entries of $\alpha L + X_0 Y H^\top$, with $\alpha = 2$, are of order 10^{-10}. Hence, the controller guarantees asymptotic stability of the closed-loop system uniformly with respect to any passive nonlinearity f.* ■

3.2. RELAXING SOME PRIOR KNOWLEDGE

A version of Theorem 3 from Chapter 2 that applies to the case in which disturbances perturb the data collection can also be given based on the same assumptions: the matrices U_0, X_0, X_1 and F_0 are known (Assumption 1) and the matrix

$$\Psi_0 := \begin{bmatrix} X_0 \\ F_0 \\ U_0 \end{bmatrix}$$

is full-row rank (Assumption 3).

Theorem 6 ***(Data-driven absolute stabilizability II with disturbed data)*** *Consider the nonlinear system* (3.1)*. Let Assumptions 1 and 3 hold and let $Q \succeq 0$. There exist three matrices G_1, G_2 and $P \succ 0$ such that*

$$\begin{bmatrix} x \\ v \end{bmatrix}^\top \begin{bmatrix} G_1^\top (X_1 - ED_0)^\top P (X_1 - ED_0) G_1 - P & G_1^\top (X_1 - ED_0)^\top P (X_1 - ED_0) G_2 \\ \star & G_2^\top (X_1 - ED_0)^\top P (X_1 - ED_0) G_2 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} < 0 \tag{3.51}$$

holds for all $(x, v) \neq 0$ that satisfy (3.26) *and D_0 that satisfy* (3.28) *if there exist $T \times n$, $T \times q$ and $n \times n$ matrices Y_1, Y_2, W and a scalar $\varepsilon > 0$ such that the conditions*

$$\begin{bmatrix} -W & WS & Y_1^\top X_1^\top & WQ^{1/2} & Y_1^\top \\ \star & R & Y_2^\top X_1^\top & 0 & Y_2^\top \\ \star & \star & -W + \varepsilon E \Delta \Delta^\top E^\top & 0 & 0 \\ \star & \star & \star & -I_n & 0 \\ \star & \star & \star & \star & -\varepsilon I_T \end{bmatrix} \prec 0 \tag{3.52}$$

$$\begin{bmatrix} X_0 Y_1 - W & X_0 Y_2 \\ F_0 Y_1 & F_0 Y_2 - I_q \end{bmatrix} = 0$$

hold. In this case, the matrices K, M that solve the problem are given by $K = U_0 Y_1 W^{-1}$ and $M = U_0 Y_2$.

The proof is similar to the one of Theorem 3 with minor modifications and is not repeated here.

3.2.1. PASSIVE NONLINEARITIES WITH ADDITIVE DISTURBANCES

As for the case of passive nonlinearities examined in Chapter 2, Example 2, we make the following observations.

The directional derivative $\dot{V}(x)$ of the Lyapunov function $V(x) = x^\top P x$ along the data-dependent vector field of system (3.1) in closed-loop with $u = Kx$, namely $\dot{x} = (X_1 - ED_0)G_1 x + (X_1 - ED_0)G_2 f$, is given by

$$\begin{aligned} \dot{V}(x) = \quad & x^\top (P(X_1 - ED_0)G_1 + G_1^\top (X_1 - ED_0)^\top P)x \\ & +2x^\top P(X_1 - ED_0)G_2 f(t,z) \end{aligned}$$

Due to the presence of the disturbance-induced matrix D_0, it is immediately recognized that the conditions (2.37) (reported below)

$$\begin{aligned} & Y_1^\top X_1^\top + X_1 Y_1 \prec 0 \\ & X_1 Y_2 + X_0 Y_1 H^\top = 0 \\ & X_0 Y_1 \succ 0 \\ & X_0 Y_2 = 0 \\ & F_0 Y_2 = I_p \\ & F_0 Y_1 = 0 \end{aligned}$$

are no longer sufficient to guarantee that $\dot{V}(x) < 0$ for all $x \neq 0$ and for all $f(t,z)$ that satisfy $z^\top f(t,z) \geq 0$ for all z. However, (2.37) suggests modified conditions to guarantee such a property for $\dot{V}(x)$. We preliminarily observe that the conditions $X_0 Y_1 \succ 0$, $X_0 Y_2 = 0$, $F_0 Y_2 = I_p$, $F_0 Y_1 = 0$, which descend from (2.31) (reported below),

$$\begin{bmatrix} I_n & 0_{n\times q} \\ 0_{q\times n} & I_q \\ K & M \end{bmatrix} = \begin{bmatrix} X_0 \\ F_0 \\ U_0 \end{bmatrix} \begin{bmatrix} G_1 & G_2 \end{bmatrix}$$

remain unchanged.

The first modified condition, a robustified version of the condition $Y_1^\top X_1^\top + X_1 Y_1 \prec 0$, implies the Hurwitz property of the matrix $(X_1 - ED_0)G_1$ for all D_0 satisfying (3.28), and is obtained by the completion-of-squares argument recalled in Section 3.1. The condition takes the form

$$\begin{bmatrix} Y_1^\top X_1^\top + X_1 Y_1 + \varepsilon E D_0 D_0^\top E^\top & Y_1^\top \\ Y_1 & -\varepsilon I_T \end{bmatrix} \prec 0$$

where, in addition to Y, we have the positive scalar ε as decision variable.

The condition $X_1Y_2 + X_0Y_1H^\top = 0$ in (2.37) aims at assigning to the nonlinear term the value $-z^\top f(t,z) \leq 0$ to $x^\top P(X_1 - ED_0)G_2 f(t,z)$. The presence of D_0, however, only allows the designer to obtain the bound $\dot{V}(x) < -2x^\top PED_0G_2 f(t,Hx)$ for $x \neq 0$, where the sign of $-2x^\top P\, ED_0G_2 f(t,Hx)$ is not determined because of D_0. In this case, to infer absolute stabilizability properties, we need to strengthen the assumption on the nonlinearity and require that $\frac{|f(t,z)|}{|z|} \to 0$ as $|z| \to 0$ uniformly with respect to t, i.e., for any $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that $|f(t,z)| < \epsilon|z|$ for all $|z| < \delta_\epsilon$ and all $t \geq 0$, which guarantees local uniform asymptotic stability of the closed-loop system. An estimate of the region of attraction of the closed-loop system could be carried out using the methods in [24].

3.3. CONCLUSIONS

With this chapter, we have completed the discussion on the design of an absolute stabilizing controller directly from measurements. As in the noiseless case, we have provided necessary and sufficient conditions for the absolute stabilizability of the closed-loop system for both the discrete-time and continuous-time case. The derived algorithm consists of a semi-definite program that from noisy system measurements (input-state data) returns a stabilizing controller without any additional step.

In the next chapter, we will introduce the concept of safety as an additional requirement in the controller design in addition to closed-loop stability. Moreover, instead of considering nonlinear systems with quadratic nonlinearities, we will focus on another class of nonlinear systems called polynomial systems.

4

DATA-DRIVEN DESIGN OF SAFE CONTROL FOR POLYNOMIAL SYSTEMS

ABSTRACT

We consider the problem of designing an invariant set using only a finite set of input-state data collected from an unknown polynomial system in continuous time. We consider noisy data, i.e., corrupted by an unknown-but-bounded disturbance. We derive a data-dependent sum-of-squares program that enforces invariance of a set and also optimizes the size of the invariant set while keeping it within a set of user-defined safety constraints; the solution of this program directly provides a polynomial invariant set and a state-feedback controller. We numerically test the design on a system of two platooning cars.

4.1. INTRODUCTION

This chapter presents the work done on enforcing invariance with a controller designed directly from data extending the results in [65] in multiple ways. First, polynomial nonlinear systems are considered in this chapter. Polynomial systems

---

This chapter has been published in "Data-driven design of safe control for polynomial systems." Luppi A., Bisoffi A., De Persis C.,and Tesi P.; arXiv preprint arXiv:2112.12664 (2021).

are a notable class of nonlinear systems widely used to model processes in engineering applications such as fluid dynamics [66, 67] and robotics [68]. Second, contrary to [65], we do not assume knowledge of the set that we would like to make invariant. Finally, we consider safety constraints to obtain an invariant set that also includes the safety set, all states satisfying the safety requirements. The resulting controller is called a safe controller.

To obtain a safe controller usually the nonlinear control design is formulated as an optimization problem but it is generally computationally intractable to verify whether a multivariate (matrix) polynomial is nonnegative. For the case of polynomial systems a computationally viable approach is to adopt a relaxation and verify instead whether such a polynomial is a sum-of-squares (SOS) since SOS optimization can be solved through semidefinite programming (SDP) [69, 70]. In a model-based setting, SOS programs have been used to design stabilizing controllers for hybrid systems [71], for disturbance analysis in linear systems [72] and to obtain inner-approximations of the basin of attraction [73] [74], to name a few applications.

Still, SOS programs suffer from some limitations that have been or are being addressed in the literature. A first one is scalability of SOS programs: tools are available [75] that automatically reduce the problem size without major computational costs, and recent works on large-scale semidefinite and polynomial optimization [76] improve scalability of SOS programs significantly. Secondly, it is often the case that the obtained SOS program is bilinear in the decision variables. This occurs in model-based SOS programs as in [73] and also in our case. An iterative approach to solve these bilinear SOS programs is commonly used [72]. Alternatively, there exist tools to solve these bilinear SOS programs directly, such as PENBMI and BMIBNB. For these reasons, the limitations of SOS programs seem largely outweighed by their positive features.

From the side of (direct) data-driven control, this work is positioned in the literature thread [77–79], to name a few, which exploits the so-called fundamental lemma in [2, Th. 1]. Within this thread, [65, 80–82] are the most closely related works to this one. As mentioned before, [65] addresses also an invariance problem, but for systems with *linear* dynamics and where the set to be rendered invariant is *given*. The works [80, 81] and [82, Section 5] consider nonlinear input-affine polynomial systems as here, but the goal is data-driven almost global [80] or global [81, 82] stabilization; invariance is a weaker dynamical property (e.g., solutions do not need to converge to an attractor within the invariant set), hence the conditions here are less conservative and yet significant to enforce safety.

In this chapter, we show how to enforce invariance in absence of a model of the system to be controlled, but using only a set of input-state data points collected from it in an experiment. We consider an input-affine nonlinear system with polynomial dynamics and a polynomial controller. This allows us to make the data-based invariance conditions tractable by using an SOS relaxation and alternately solving two SOS programs. In this work, we consider the realistic setting when invariance needs to be guaranteed despite the presence of an unknown-but-bounded [83] additive noise in data. Moreover, we show that also in the data-driven case it is possible to optimize the size of the invariant set while respecting safety constraints. Finally, we provide numerical evidence to show the effectiveness of the approach, in particular on the practical example of two platooning cars.

The chapter is structured as follows. In Section 4.2 recalls the theory enabling our results. In Section 4.3, we set up the invariance problem for polynomial systems. In Section 4.5, we derive a data-driven controller that solves the invariance problem, and also includes safety constraints. In Section 4.7, we exemplify our data-based solution for two platooning cars.

4.2. PRELIMINARIES

In this section, we present a fundamental result from real algebraic geometry, the Positivstellensatz. We first report from [70, 72] the notions needed to state that result.

We start from sum-of-squares (SOS) polynomials and SOS matrix polynomials. A function $h:\mathbb{R}^n \to \mathbb{R}$ is a *monomial* of degree d in $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ if

$$h(x) = a x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}$$

with $a \in \mathbb{R}$, $q_1, \dots, q_n \in \mathbb{Z}_{\geq 0}$ and $d = \sum_{i=1}^n q_i$. A function $h:\mathbb{R}^n \to \mathbb{R}$ is a *polynomial* if it is a sum of (a finite number of) monomials $h_1, h_2, \dots : \mathbb{R}^n \to \mathbb{R}$ with finite degree, and the largest degree of the h_i's is the degree of h. Π denotes the set of polynomials.

Definition 2 (SOS polynomial) *$h \in \Pi$ is an* SOS *polynomial if there exist $h_1, \dots, h_k \in \Pi$ such that $h(x) = \sum_{i=1}^k h_i(x)^2$. The set of SOS polynomials $h \in \Pi$ is denoted as Σ.*

A function $H:\mathbb{R}^n \to \mathbb{R}^{r_1 \times r_2}$ is a *matrix polynomial* if the entries of H satisfy $h_{ij} \in \Pi$ for all $i = 1, \dots, r_1$ and $j = 1, \dots, r_2$, and the largest degree of the entries of H is the degree of H. The set of matrix polynomials $H:\mathbb{R}^n \to \mathbb{R}^{r_1 \times r_2}$ is denoted by Π_{r_1, r_2}. A function $H:\mathbb{R}^n \to \mathbb{R}^{r_1 \times r_2}$ is a *square matrix polynomial* if $r_1 = r_2$. The set of square matrix polynomials $H:\mathbb{R}^n \to \mathbb{R}^{r \times r}$ is denoted by Π_r.

Definition 3 (SOS matrix polynomial [70]) *$H \in \Pi_r$ is an* SOS matrix polynomial *if there exist $H_1, \dots, H_k \in \Pi_r$ such that $H(x) = \sum_{i=1}^k H_i(x)^\top H_i(x)$. The set of SOS matrix polynomials $H \in \Pi_r$ is denoted by Σ_r.*

SOS polynomials are instrumental to define three sets of polynomials appearing in the Positivstellensatz.

Definition 4 (Multiplicative monoid [72, Def. 3]) *Given* $g_1, \dots, g_t \in \Pi$*, the* multiplicative monoid *generated by* g_j*'s is the set of all finite products of* g_j*'s, including* 1 *(i.e., the empty product). It is denoted by* $\mathscr{M}(g_1, \dots, g_t)$*, with* $\mathscr{M}(\emptyset) := 1$ *for completeness.*

An example is $\mathscr{M}(g_1, g_2) = \{g_1^{k_1} g_2^{k_2} : k_1, k_2 \in \mathbb{Z}_{\geq 0}\}$.

Definition 5 (Cone [72, Def. 4]) *Given* $f_1, \dots, f_r \in \Pi$*, the* cone *generated by* f_i*'s is* $\mathscr{C}(f_1, \dots, f_r) := \left\{s_0 + \sum_{i=1}^{l} s_i b_i : l \in \mathbb{Z}_{\geq 0}, s_i \in \Sigma, b_i \in \mathscr{M}(f_1, \dots, f_r)\right\}$.

4

If $s \in \Sigma$ and $f \in \Pi$, then $f^2 s \in \Sigma$. Then, a cone of $\{f_1, \dots, f_r\}$ can be expressed as a sum of 2^r terms without loss of generality. An example is $\mathscr{C}(f_1, f_2) = \{s_0 + s_1 f_1 + s_2 f_2 + s_3 f_1 f_2 : s_0, \dots, s_3 \in \Sigma\}$ where terms like $s_4 f_1^2$ or $s_5 f_2^2$ with s_4, $s_5 \in \Sigma$ are not needed since they are captured by s_0 anyway.

Definition 6 (Ideal [72, Def. 5]) *Given* $h_1, \dots, h_u \in \Pi$*, the* ideal *generated by* h_k*'s is* $\mathscr{I}(h_1, \dots, h_u) := \{\sum_{k=1}^{u} h_k p_k : p_k \in \Pi\}$.

An example is $\mathscr{I}(h_1, h_2) = \{h_1 p_1 + h_2 p_2 : p_1, p_2 \in \Pi\}$. With Definitions 4-6, we finally recall the version in [72] of the Positivstellensatz (P-Satz), in the next fact.

Fact 1 (Positivstellensatz (P-Satz) [72, Th. 1]) *Given* $f_1, \dots, f_r \in \Pi$, $g_1, \dots, g_t \in \Pi$, *and* $h_1, \dots, h_u \in \Pi$*, the following are equivalent.*

1. *The set*
$$\left\{ x \in \mathbb{R}^n : \begin{array}{c} f_1(x) \geq 0, \dots, f_r(x) \geq 0 \\ g_1(x) \neq 0, \dots, g_t(x) \neq 0 \\ h_1(x) = 0, \dots, h_u(x) = 0 \end{array} \right\} = \emptyset.$$

2. *There exist polynomials* $f \in \mathscr{C}(f_1, \dots, f_r)$, $g \in \mathscr{M}(g_1, \dots, g_t)$, $h \in \mathscr{I}(h_1, \dots, h_u)$ *such that*
$$f + g^2 + h = 0.$$

4.2.1. NOTIONS ON MATRIX ELLIPSOIDS

To have a self-contained and complete explanation of the notions that we used to formulate the main results of this chapter, we need to discuss the concepts of matrix ellipsoid explained in [86]. For symmetric matrices $\mathrm{P} \in \mathbb{R}^{p \times p}$, $\mathrm{Q} \in \mathbb{R}^{q \times q}$, $\mathrm{A} \in \mathbb{R}^{p \times p}$, $\mathrm{C} \in \mathbb{R}^{q \times q}$ and matrices $Z_c \in \mathbb{R}^{p \times q}$, $\mathrm{B} \in \mathbb{R}^{p \times q}$, we define as *matrix ellipsoid* a set in one of the next two forms:

$$\mathcal{E}_{\text{mat}} := \left\{ Z \in \mathbb{R}^{p\times q} : (Z - Z_c)^\top P^{-2} (Z - Z_c) \preceq Q \right\} \tag{4.1}$$

$$\mathcal{E}'_{\text{mat}} := \left\{ Z \in \mathbb{R}^{p\times q} : Z^\top A Z + Z^\top B + B^\top Z + C \preceq 0 \right\} \tag{4.2}$$

where $P \succ 0, Q \succ 0$ and $A \succ 0$, $B^\top A^{-1} B - C \succ 0$. The constraints $Q \succ 0$ and $B^\top A^{-1} B - C \succ 0$ ensure that $\mathcal{E}_{\text{mat}}$ and $\mathcal{E}'_{\text{mat}}$ are not empty or do not reduce to a singleton; the constraints $P \succ 0$ and $A \succ 0$ ensure then that the matrix ellipsoid is a bounded set. We stress that many sets considered in the sequel have to be expressed in terms of these matrix ellipsoids, and that $\mathcal{E}_{\text{mat}}$ and $\mathcal{E}'_{\text{mat}}$ are natural extensions of the classical ellipsoids in the Euclidean space. Standard computations reformulate (4.2) as

$$\mathcal{E}'_{\text{mat}} = \{ Z \in \mathbb{R}^{p\times q} : (Z + A^{-1} B)^\top A (Z + A^{-1} B) - (B^\top A^{-1} B - C) \preceq 0 \}$$

Hence, $\mathcal{E}_{\text{mat}}$ and $\mathcal{E}'_{\text{mat}}$ are the same set for

$$Z_c = -A^{-1} B,\ P^{-2} = A,\ Q = B^\top A^{-1} B - C \tag{4.3}$$

Establishing the correspondence in (4.3) is useful since in the following (Section 4.5) matrix ellipsoids appear more naturally in the form (4.2) than (4.1). On the other hand when we need to define a size for matrix ellipsoids, it is easier using $\mathcal{E}_{\text{mat}}$ in (4.1). By $Q \succ 0$ (4.1) is equivalently written as

$$\begin{aligned} \mathcal{E}_{\text{mat}} &= \left\{ Z \in \mathbb{R}^{p\times q} : Q^{-1/2} (Z - Z_c)^\top P^{-2} (Z - Z_c) Q^{-1/2} \preceq I \right\} \\ &= \left\{ Z_c + P Y O^{1/2} : Y \in \mathbb{R}^{p\times q}, Y^\top Y \preceq I \right\} \end{aligned}$$

where $Y^\top Y \preceq I$ is equivalent to $|Y| \leq 1$.

4.3. COLLECTING DATA AND ENFORCING INVARIANCE

We consider polynomial systems of the form

$$\dot{x} = f(x) + g(x)u \tag{4.4}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, f and g are polynomial vector fields. The specific expressions of f and g are unknown. The polynomial system (4.4) can be written into the linear-like form by re-writing the polynomials as a multiplication of coefficients and monomials

$$\dot{x} = \mathsf{A} Z(x) + \mathsf{B} W(x) u \tag{4.5}$$

where the coefficient matrices $\mathsf{A} \in \mathbb{R}^{n\times N_\mathsf{A}}$ and $\mathsf{B} \in \mathbb{R}^{n\times N_\mathsf{B}}$ are *unknown*, the *known* $N_\mathsf{A} \times 1$ vector $Z(x)$ contains as entries the distinct monomials in x that may appear in f, and the *known* $N_\mathsf{B} \times m$ matrix $W(x)$ contains as entries the monomials that may appear in g. The conditions we will propose to design an invariant set are the same regardless of the choice of the monomials in Z and W. On the other hand, different choices of Z and W affect feasibility and quality of the solution arising from those conditions, as is generally the case with model structure selection. In Section 4.7, we will present guidelines for the choice of monomials in Z and W.

We consider the control law $u = K(x)$ where $K \in \Pi_{m,1}$ is to be designed. The closed-loop dynamics results in

4

$$\dot{x} = \mathsf{A}Z(x) + \mathsf{B}W(x)K(x) = \begin{bmatrix} \mathsf{A} & \mathsf{B} \end{bmatrix} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}. \tag{4.6}$$

Data are generated through an experiment in the presence of an additive disturbance d as

$$\dot{x} = \mathsf{A}Z(x) + \mathsf{B}W(x)u + d. \tag{4.7}$$

We apply an input sequence of T elements, and measure the state and state derivative sequences generated by (4.7); we sample uniformly these sequences at times $0, \tau_\mathrm{s}, \ldots, (T-1)\tau_\mathrm{s}$ for sampling period $\tau_\mathrm{s} > 0$; this results in data points, for $j = 0, \ldots, T-1$,

$$\dot{x}^j := \dot{x}(j\tau_\mathrm{s}),\ z^j := Z(x(j\tau_\mathrm{s})),\ v^j := W(x(j\tau_\mathrm{s}))u(j\tau_\mathrm{s}). \tag{4.8}$$

A disturbance sequence given for $j = 0, \ldots, T-1$ by

$$d^j := d(j\tau_\mathrm{s})$$

acts during the experiment but is *unknown*. Hence, the data generation mechanism is described by

$$\dot{x}^j = \mathsf{A}z^j + \mathsf{B}v^j + d^j,\ j = 0, 1, \ldots, T-1. \tag{4.9}$$

4.4. INVARIANT SET

Our goal is to use the collected data points to obtain an invariant set for (4.6) as specified in the next definition.

Definition 7 (Invariant set) *For $a\colon \mathbb{R}^n \to \mathbb{R}^n$ polynomial, i.e., $a \in \Pi_{n,1}$, and for an arbitrary $x_0 \in \mathbb{R}^n$, denote $t \mapsto \alpha(t, x_0)$ the (unique maximal) solution to $\dot{x} = a(x)$ with initial condition $x_0 = \alpha(0, x_0)$ and defined on the interval $[0, T(x_0))$ (with $T(x_0)$ possibly $+\infty$). A set $\mathcal{I}$ is said to be* invariant *for $\dot{x} = a(x)$ if $x_0 \in \mathcal{I}$ implies that for all $t \in [0, T(x_0))$, $\alpha(t, x_0) \in \mathcal{I}$.*

Consider the set

$$\mathscr{I} := \{x \in \mathbb{R}^n : h(x) \leq 0\} \tag{4.10}$$

with $h \in \Pi$. To impose that $\mathscr{I}$ is an invariant set according to Definition 7, we require the condition

$$\{x \in \mathbb{R}^n : h(x) = 0\} \subseteq \{x \in \mathbb{R}^n : \frac{\partial h}{\partial x}(x)\dot{x} \leq -\epsilon\} \tag{4.11}$$

where $\dot{x}$, used for brevity, takes the expression in (4.6) and the parameter $\epsilon > 0$ is introduced to guarantee some degree of robustness at the boundary of $\mathscr{I}$.

Remark 1 *We emphasize that whereas we use noisy data for control design as per the data generation mechanism in* (4.9)*, we design a controller to enforce* nominal *invariance for $d = 0$ as in* (4.6)*, instead of robust invariance. Nonetheless, our design features some degree of robustness thanks to ϵ, as we now show. Define the nominal closed-loop vector field in* (4.6) *as $f_{\text{n}}(x) := \mathsf{A}Z(x) + \mathsf{B}W(x)K(x)$ and consider now the perturbed dynamics $\dot{x} = f_{\text{n}}(x) + d$ with $|d|^2 \leq \omega$ (as we will have later in* (4.15)*). $\mathscr{I}$ is robustly invariant for this perturbed dynamics as long as*

$$\tfrac{\partial h}{\partial x}(x)(f_{\text{n}}(x) + d) \leq 0 \quad \forall (x,d)\colon\ h(x) = 0, |d|^2 \leq \omega \tag{4.12}$$

By achieving (4.11)*, we have that for all x and d with $h(x) = 0$ and $|d|^2 \leq \omega$,*

$$\tfrac{\partial h}{\partial x}(x)(f_{\text{n}}(x) + d) \leq -\epsilon + \left|\tfrac{\partial h}{\partial x}(x)\right| \sqrt{\omega}.$$

Then, robust invariance in (4.12) *is achieved if $\left|\tfrac{\partial h}{\partial x}(x)\right| \leq \epsilon/\sqrt{\omega}$ for all x such that $h(x) = 0$ or, equivalently, if*

$$\begin{bmatrix} -\epsilon^2/\omega & \tfrac{\partial h}{\partial x}(x) \\ \tfrac{\partial h}{\partial x}(x)^\top & -I \end{bmatrix} \leq 0 \quad \forall x\colon\ h(x) = 0,$$

which can be relaxed into an SOS condition and added to our optimization program (see later Theorem 8). For these reasons, we rather consider for simplicity nominal invariance and introduce $\epsilon > 0$ to guarantee nonetheless some degree of robustness at the boundary of $\mathscr{I}$.

With the goal of applying Fact 1, (4.11) can be cast as

$$\left\{x \in \mathbb{R}^n :\ h(x) = 0, \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon \geq 0,\ \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon \neq 0\right\} = \emptyset.$$

This empty set condition is equivalent, by Fact 1, to the existence of $\ell_1 \in \Pi$, $s_0, s_1 \in \Sigma$ and $k_1 \in \mathbb{Z}_{\geq 0}$ such that

$$\ell_1 h + s_0 + s_1\Big(\frac{\partial h}{\partial x}\dot{x} + \epsilon\Big) + \Big(\frac{\partial h}{\partial x}\dot{x} + \epsilon\Big)^{2k_1} = 0. \tag{4.13}$$

With the final goal of implementing numerically this condition, we simplify (4.13) by setting $s_0 = 0$, $k_1 = 1$ and $\ell_1 = \ell\big(\frac{\partial h}{\partial x}\dot{x} + \epsilon\big)$ and, for $\ell \in \Pi$ and $s_1 \in \Sigma$, obtain

$$\Big(\frac{\partial h}{\partial x}\dot{x} + \epsilon\Big)\Big[\ell h + s_1 + \Big(\frac{\partial h}{\partial x}\dot{x} + \epsilon\Big)\Big] = 0,$$

which is a relaxation of (4.13). So, the original (4.11) is implied by

$$\begin{aligned} &\ell(x)h(x) + \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon + s_1(x) = 0 \\ &\ell(x)h(x) + \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon = -s_1(x) \leq 0 \qquad \forall x,\ s_1 \in \Sigma. \end{aligned} \tag{4.14}$$

The arguments above can be summarized as follows.

Lemma 1 *If there exists $\ell \in \Pi$ such that condition* (4.14) *holds, then $\mathcal{I}$ in* (4.10) *is an invariant set for* (4.7).

Indeed, for all x such that $h(x) = 0$ (corresponding to the boundary of $\mathcal{I}$), the condition imposes $\frac{\partial h}{\partial x}(x)\dot{x} \leq -\epsilon$, i.e., the Lie derivative of h is strictly negative. For all x such that $h(x) < 0$ (corresponding to the interior of $\mathcal{I}$), the condition imposes $\frac{\partial h}{\partial x}(x)\dot{x} \leq -\epsilon - \ell(x)h(x)$. Since ℓ is a polynomial without any sign definiteness requirement and ϵ is a design parameter selected as a small positive number, the term $-\epsilon - \ell(x)h(x)$ does not need to be negative and can actually be positive; hence, the condition may allow even for a positive $\frac{\partial h}{\partial x}(x)\dot{x}$, consistently with set invariance being less restrictive than attractivity.

4.5. DATA-DRIVEN SAFE CONTROL

Our goal is to formulate a condition depending exclusively on noisy data to find an invariant set for the actual system (4.6). We substitute in (4.14) the closed-loop dynamics in (4.6) and obtain

$$\ell(x)h(x) + \epsilon + \frac{\partial h}{\partial x}(x)\begin{bmatrix}\mathsf{A} & \mathsf{B}\end{bmatrix}\begin{bmatrix} Z(x) \\ W(x)K(x)\end{bmatrix} \leq 0 \quad \forall x.$$

Since the model is not available and the true coefficient matrices A and B are unknown, we rather enforce the previous inequality on all matrices (A, B) that are consistent with data (for a given disturbance model), as we now characterize. By *consistent with data*, we mean all matrices (A, B) that could have produced the measured data sequences $\{\dot{x}^j, z^j, v^j\}_{j=0}^{T-1}$ as in (4.8) for an additive disturbance d that is bounded. A realistic bound on disturbance is that the norm of any possible disturbance instance d is upper-bounded, and so are the norms of the unknown $d^0, \dots, d^{T-1}$. This corresponds to an *instantaneous* bound given, for $\omega \geq 0$, by the set

$$\mathscr{D}_{\text{ins}} := \{d \in \mathbb{R}^n : |d|^2 \leq \omega\}. \tag{4.15}$$

Remark 2 *Note that $|d|^2 \leq \omega$ is equivalent to $d^\top d \leq \omega$ or $dd^\top \preceq \omega I$.*

Then, based on the bound in $\mathscr{D}_{\text{ins}}$, the system matrices consistent with a single data point $j \in \{0, \dots, T-1\}$ belong to the set

$$\begin{aligned} \mathscr{C}_{\text{i}}^j :=& \Big\{[A\ B] :\ \dot{x}^j = Az^j + Bv^j + d,\ dd^\top \preceq \omega I\Big\}, \qquad (4.16) \\ =& \Big\{[A\ B] :\ \begin{bmatrix} I & A & B \end{bmatrix} \begin{bmatrix} I & \dot{x}^j \\ 0 & -z^j \\ 0 & -v^j \end{bmatrix} \cdot \begin{bmatrix} \omega I & 0 \\ 0 & -I \end{bmatrix} [\star]^\top \succeq 0\Big\}. \end{aligned}$$

$\mathscr{C}_{\text{i}}^j$ is the set of all matrices for which some disturbance d satisfying the bound in (4.15) could have generated the measured data point j, as in [84, 85]. Here and in the following the subscript i denotes the formulation for the instantaneous bound (4.15), it is not an index.

The set of matrices (A, B) consistent with all data points $j = 0, \dots, T-1$ is then

$$\mathscr{C}_{\text{i}} := \bigcap_{j=0}^{T-1} \mathscr{C}_{\text{i}}^j. \tag{4.17}$$

Unfortunately (4.17) is difficult to obtain exactly. In fact, $\mathscr{C}_{\text{i}}$ is an intersection of matrix ellipsoids, and even for systems with one input and one output, finding the ellipsoid of minimum volume containing $\mathscr{C}_{\text{i}}$ is NP-complete. As consequence for matrices ζ_0, $P = P^\top \succ 0$, $Q = Q^\top \succeq 0$, the set $\mathscr{C}_{\text{i}}$ cannot be expressed as a matrix ellipsoid[1] of the form

$$\{\zeta \in \mathbb{R}^{(n+m)\times n} : (\zeta - \zeta_0)^\top P^{-2} (\zeta - \zeta_0) \preceq Q\}, \tag{4.18}$$

which is instrumental to obtain our main result.

We thus set up a convex optimization problem to obtain a computable over-approximation $\overline{\mathscr{C}_{\text{i}}}$ of $\mathscr{C}_{\text{i}}$.

[1]As the name suggests, a matrix ellipsoid is an extension of the classical (vector) ellipsoid $\{\zeta \in \mathbb{R}^p : (\zeta - \zeta_0)^\top P^{-2} (\zeta - \zeta_0) \leq q\}$ with $P = P^\top \succ 0$ and $q \geq 0$, see [86] for details.

Using $\overline{\mathscr{C}}_{\mathrm{i}}$ it is possible to obtain (4.18) for the matrix ellipsoid defined for $A_{\mathrm{i}} \succ 0$ as

$$\overline{\mathscr{C}}_{\mathrm{i}} := \left\{ [A\ B] = \zeta^\top : \begin{bmatrix} I & \zeta^\top \end{bmatrix} \begin{bmatrix} B_{\mathrm{i}}^\top A_{\mathrm{i}}^{-1} B_{\mathrm{i}} - I & B_{\mathrm{i}}^\top \\ B_{\mathrm{i}} & A_{\mathrm{i}} \end{bmatrix} \begin{bmatrix} I \\ \zeta \end{bmatrix} \preceq 0 \right\}. \tag{4.19}$$

Indeed, for $A_{\mathrm{i}} \succ 0$, the condition in (4.19) is precisely $(\zeta + A_{\mathrm{i}}^{-1} B_{\mathrm{i}})^\top A_{\mathrm{i}} (\zeta + A_{\mathrm{i}}^{-1} B_{\mathrm{i}}) \preceq I$, so that the selections

$$\bar{\zeta}_{\mathrm{i}} := -A_{\mathrm{i}}^{-1} B_{\mathrm{i}}, \quad \overline{P}_{\mathrm{i}}^{-2} := A_{\mathrm{i}}, \quad \overline{Q}_{\mathrm{i}} := I \tag{4.20}$$

rewrite $\overline{\mathscr{C}}_{\mathrm{i}}$ equivalently as

4

$$\overline{\mathscr{C}}_{\mathrm{i}} = \{[A\ B] = \zeta^\top : (\zeta - \bar{\zeta}_{\mathrm{i}})^\top \overline{P}_{\mathrm{i}}^{-2} (\zeta - \bar{\zeta}_{\mathrm{i}}) \preceq \overline{Q}_{\mathrm{i}}\}. \tag{4.21}$$

In summary, because the form (4.21) allows us to obtain our main result, we over-approximate $\mathscr{C}_{\mathrm{i}}$ in (4.17) through $\overline{\mathscr{C}}_{\mathrm{i}}$ in (4.19) where the matrices $A_{\mathrm{i}} \succ 0$ and B_{i} are determined by solving an optimization program, which we recall from [86, Section 5.1].
From data, define for $j = 0, \dots, T-1$

$$\begin{aligned} c_j &:= -\omega I + \dot{x}^j (\dot{x}^j)^\top, \\ b_j &:= -\begin{bmatrix} z^j \\ v^j \end{bmatrix} (\dot{x}^j)^\top, \ a_j := \begin{bmatrix} z^j \\ v^j \end{bmatrix} \begin{bmatrix} z^j \\ v^j \end{bmatrix}^\top. \end{aligned} \tag{4.22}$$

As it is done for classical ellipsoids [51, Section 3.7.2], we impose that the matrix ellipsoid $\overline{\mathscr{C}}_{\mathrm{i}}$, which is well-defined for $A_{\mathrm{i}} \succ 0$, includes $\mathscr{C}_{\mathrm{i}}$ through the (lossy) S-procedure [51, Section 2.6.3] and we then minimize the size of $\overline{\mathscr{C}}_{\mathrm{i}}$. This corresponds to the optimization program

$$\begin{aligned} &\text{minimize } -\log\det A_{\mathrm{i}} \\ &\text{subject to} \\ &\begin{bmatrix} -I - \sum_{j=0}^{T-1} \tau_j c_j & B_{\mathrm{i}}^\top - \sum_{j=0}^{T-1} \tau_j b_j^\top & B_{\mathrm{i}}^\top \\ B_{\mathrm{i}} - \sum_{j=0}^{T-1} \tau_j b_j & A_{\mathrm{i}} - \sum_{j=0}^{T-1} \tau_j a_j & 0 \\ B_{\mathrm{i}} & 0 & -A_{\mathrm{i}} \end{bmatrix} \preceq 0, \\ &A_{\mathrm{i}} \succ 0,\ \tau_j \geq 0 \text{ for } j = 0, \dots, T-1. \end{aligned} \tag{4.23}$$

When this optimization program is solved, we use the returned A_{i} and B_{i} to obtain the matrices $\bar{\zeta}_{\mathrm{i}}$, $\overline{P}_{\mathrm{i}}$ $\overline{Q}_{\mathrm{i}}$ as in (4.20). Before further analyzing the optimization program, we discuss in the next remark an alternative bound on the disturbance that is commonly used.

Remark 3 *When an instantaneous bound on the disturbance is given by* $|d|^2 \leq \omega$ *as in* $\mathscr{D}_{\text{ins}}$ *in* (4.15), *one can infer that the whole unknown disturbance sequence* $\begin{bmatrix} d^0 & \dots & d^{T-1} \end{bmatrix}$ *of the experiment belongs to the set*

$$\mathscr{D}_{\text{e}} := \{D \in \mathbb{R}^{n \times T} : DD^\top \preceq T\omega I\}, \tag{4.24}$$

which we call an energy *bound on the disturbance. By collecting data points in* (4.8) *in matrices*

$$X_1 := \begin{bmatrix} \dot{x}^0 & \dots & \dot{x}^{T-1} \end{bmatrix}, \tag{4.25a}$$

$$Z_0 := \begin{bmatrix} z^0 & \dots & z^{T-1} \end{bmatrix}, V_0 := \begin{bmatrix} v^0 & \dots & v^{T-1} \end{bmatrix}, \tag{4.25b}$$

the set of matrices consistent with data is

$$\mathscr{C}_{\text{e}} := \Big\{[A\ B] : X_1 = AZ_0 + BV_0 + D,\ D \in \mathbb{R}^{n \times T},\ DD^\top \preceq T\omega I\Big\}, \tag{4.26}$$

or, after standard manipulations as in [82, Section 2.3],

$$\begin{aligned}
&\mathscr{C}_{\text{e}} = \left\{[A\ B] = \zeta^\top :\ \begin{bmatrix} I & \zeta^\top \end{bmatrix} \begin{bmatrix} C_{\text{e}} & B_{\text{e}}^\top \\ B_{\text{e}} & A_{\text{e}} \end{bmatrix} \begin{bmatrix} I \\ \zeta \end{bmatrix} \preceq 0 \right\} \\
&A_{\text{e}} := \begin{bmatrix} Z_0 \\ V_0 \end{bmatrix} \begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}^\top,\ B_{\text{e}} := -\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix} X_1^\top, \\
&C_{\text{e}} := -T\omega I + X_1 X_1^\top.
\end{aligned} \tag{4.27}$$

This form is mathematically analogous to that of $\overline{\mathscr{C}}_{\text{i}}$ *in* (4.19) *and one can indeed adapt the results we will find for* $\overline{\mathscr{C}}_{\text{i}}$ *to the set* $\mathscr{C}_{\text{e}}$*, as we will show in Remark 4. However, [86] illustrates that, unless* T *is very large and thus impacts the computational cost, it is advantageous to work with* $\overline{\mathscr{C}}_{\text{i}}$ *instead of* $\mathscr{C}_{\text{e}}$*.*

The results that follow rely on the optimization program in (4.23) being feasible. We would like to show that feasibility of (4.23) is guaranteed under a relatively mild assumption. This assumption is that the data set $\{\dot{x}^j, z^j, v^j\}_{j=0}^{T-1}$ yields a matrix $\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix} := \begin{bmatrix} z^0 & \dots & z^{T-1} \\ v^0 & \dots & v^{T-1} \end{bmatrix}$ with full row rank. This rank condition can be easily checked from data and when not verified, one can always collect additional data points, thereby adding columns to $\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}$, to try and meet the rank condition (in a linear setting, this rank condition is related to classical persistence of excitation, see [82, Section 4.1]). We have then the next result for feasibility of (4.23).

Lemma 2 *If* $\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}$ *has full row rank, then the optimization program* (4.23) *is feasible.*

Proof. With the matrices A_e and B_e defined in (4.27) from data, set in (4.23)

$$\tau_0 = \cdots = \tau_{T-1} = \tau, A_\mathrm{i} = \tau A_\mathrm{e}, B_\mathrm{i} = \tau B_\mathrm{e} \tag{4.28}$$

for $\tau > 0$ to be determined in the proof. Since $\left[\begin{smallmatrix} Z_0 \\ V_0 \end{smallmatrix}\right]$ has full row rank, A_e in (4.27) satisfies $A_\mathrm{e} \succ 0$. Consider the constraints in (4.23); by (4.28), $A_\mathrm{e} \succ 0$ and the selection $\tau > 0$, the statement of the lemma is proven if we choose $\tau > 0$ suitably to verify

$$\begin{bmatrix} -I-\sum_{j=0}^{T-1}\tau c_j & \tau B_\mathrm{e}^\top - \sum_{j=0}^{T-1}\tau b_j^\top & \tau B_\mathrm{e}^\top \\ \tau B_\mathrm{e} - \sum_{j=0}^{T-1}\tau b_j & \tau A_\mathrm{e} - \sum_{j=0}^{T-1}\tau a_j & 0 \\ \tau B_\mathrm{e} & 0 & -\tau A_\mathrm{e} \end{bmatrix} \preceq 0.$$

By recalling (4.22) and (4.27), we have the relations

$$C_\mathrm{e} = \sum_{j=0}^{T-1} c_j = -T\omega I + X_1 X_1^\top, \quad B_\mathrm{e} = \sum_{j=0}^{T-1} b_j = -\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix} X_1^\top, \quad A_\mathrm{e} = \sum_{j=0}^{T-1} a_j = \begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}^\top.$$

By substituting these relations in the previous matrix inequality, we want to choose $\tau > 0$ suitably to verify

$$\begin{bmatrix} -I-\tau C_\mathrm{e} & 0 & \tau B_\mathrm{e}^\top \\ 0 & 0 & 0 \\ \tau B_\mathrm{e} & 0 & -\tau A_\mathrm{e} \end{bmatrix} \preceq 0 \iff \begin{bmatrix} -I-\tau C_\mathrm{e} + \tau B_\mathrm{e}^\top A_\mathrm{e}^{-1} B_\mathrm{e} & 0 \\ 0 & 0 \end{bmatrix} \preceq 0$$

by Schur complement and $\tau A_\mathrm{e} \succ 0$. The last condition is indeed true and the statement is thus proven because $\overline{Q}_\mathrm{e} := B_\mathrm{e}^\top A_\mathrm{e}^{-1} B_\mathrm{e} - C_\mathrm{e} \succeq 0$ by [82, Lemma 1] and a sufficiently small $\tau > 0$ ensures $-I + \tau \overline{Q}_\mathrm{e} \preceq 0$. □

With the set $\mathscr{C}_\mathrm{i}$ of matrices consistent with data in (4.17) and its over-approximation $\overline{\mathscr{C}}_\mathrm{i}$ obtained via the optimization program (4.23), we can solve the considered problem of enforcing invariance for ground truth matrices (A, B). This is achieved by enforcing invariance in (4.14) for all matrices (A, B) in $\overline{\mathscr{C}}_\mathrm{i}$ as in the next main result, so that an invariant set is determined directly from data.

Theorem 7 ***(Data-driven invariance condition)***
For a design parameter $\epsilon > 0$, consider the data generation mechanism (4.9)*, measured data $\{\dot{x}^j, z^j, v^j\}_{j=0}^{T-1}$ as in* (4.8) *and disturbance d satisfying the instantaneous bound $\mathscr{D}_\mathrm{ins}$ in* (4.15) *(i.e., $d^0 \in \mathscr{D}_\mathrm{ins}, \ldots, d^{T-1} \in \mathscr{D}_\mathrm{ins}$).*
Assume that the optimization program in (4.23) *is feasible so that parameters $\overline{\zeta}_\mathrm{i}$, $\overline{P}_\mathrm{i}$ and $\overline{Q}_\mathrm{i}$ in* (4.20) *exist. Assume that there exist decision variables $\ell \in \Pi$, $\eta \in \Pi$, $h \in \Pi$ and $K \in \Pi_{m,1}$ such that for all $x \in \mathbb{R}^n$, $\eta(x) > 0$ and $H(x) \preceq 0$, with H defined as*

$$H(x) := \begin{bmatrix} \left\{ \begin{matrix} \ell(x)h(x) + \epsilon \\ + \frac{\partial h}{\partial x}(x)\overline{\zeta}_\mathrm{i}^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} \end{matrix} \right\} & \star & \star \\ \eta(x)\overline{P}_\mathrm{i}\begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} & -2\eta(x)I & \star \\ \overline{Q}_\mathrm{i}^{1/2}\frac{\partial h}{\partial x}(x)^\top & 0 & -2\eta(x)I \end{bmatrix}. \tag{4.29}$$

Then, the set $\mathscr{I}$ in (4.10) *is invariant for the system in* (4.6)*.*

Proof. Under the assumption that (4.23) is feasible, $A_{\rm i} \succ 0$ by construction, so $\overline{P}_{\rm i}$ and $\overline{Q}_{\rm i}$ in (4.20) satisfy $\overline{P}_{\rm i} \succ 0$ and $\overline{Q}_{\rm i} \succ 0$. These two relations allow us to rewrite (4.21) as

$$\overline{\mathscr{C}}_{\rm i} = \{\zeta^\top \in \mathbb{R}^{n\times(n+m)} \colon \overline{Q}_{\rm i}^{-1/2}(\zeta - \overline{\zeta}_{\rm i})^\top \overline{P}_{\rm i}^{-1} \cdot [\star]^\top \preceq I\} \tag{4.30}$$

$$= \{(\overline{\zeta}_{\rm i} + \overline{P}_{\rm i} Y \overline{Q}_{\rm i}^{1/2})^\top \colon Y \in \mathbb{R}^{(n+m)\times n}, Y^\top Y \preceq I\} \tag{4.31}$$

where (4.30) is obtained from (4.21) since $\overline{Q}_{\rm i}^{-1/2} \succ 0$ and (4.31) is obtained by setting $Y = \overline{P}_{\rm i}^{-1}(\zeta - \overline{\zeta}_{\rm i})\overline{Q}_{\rm i}^{-1/2}$ in (4.30). Since the disturbance d satisfies the instantaneous bound $\mathscr{D}_{\rm ins}$ in (4.15), $(\mathsf{A}, \mathsf{B}) \in \bigcap_{j=0}^{T-1} \mathscr{C}_{\rm i}^j = \mathscr{C}_{\rm i}$ and, thus, $(\mathsf{A}, \mathsf{B}) \in \overline{\mathscr{C}}_{\rm i}$ in (4.31) since (4.23) enforces $\mathscr{C}_{\rm i} \subseteq \overline{\mathscr{C}}_{\rm i}$ by construction. The invariance condition in (4.14) imposed for all matrices in $\overline{\mathscr{C}}_{\rm i}$ reads

$$\forall [A\ B] \in \overline{\mathscr{C}}_{\rm i}, \forall x \in \mathbb{R}^n, \begin{cases} \ell(x)h(x) + \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon \le 0 \\ \dot{x} = AZ(x) + BW(x)K(x). \end{cases} \tag{4.32}$$

Since $[\mathsf{A}\ \mathsf{B}] \in \overline{\mathscr{C}}_{\rm i}$, if condition (4.32) holds, then the set $\mathscr{I}$ in (4.10) is invariant for the ground truth system in (4.6). Therefore, the proof is complete if we verify (4.32), i.e., if we verify that

$$\begin{aligned} &\forall [A\ B] \in \overline{\mathscr{C}}_{\rm i}, \forall x \in \mathbb{R}^n, \\ &\ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)\,[\,A\ B\,]\begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} \\ &= \ell(x)h(x) + \epsilon + \tfrac{1}{\sqrt{2}}\begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top \begin{bmatrix} A^\top \\ B^\top \end{bmatrix} \tfrac{1}{\sqrt{2}}\tfrac{\partial h}{\partial x}(x)^\top \\ &\quad + \tfrac{1}{\sqrt{2}}\tfrac{\partial h}{\partial x}(x)\,[\,A\ B\,]\,\tfrac{1}{\sqrt{2}}\begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} \le 0. \end{aligned}$$

Rewrite this invariance condition in the compact form

$$\begin{aligned} &\forall \zeta^\top \in \overline{\mathscr{C}}_{\rm i}, \forall x \in \mathbb{R}^n, \\ &W(x) + S(x)\zeta R(x) + R(x)^\top \zeta^\top S(x)^\top \le 0 \end{aligned} \tag{4.33}$$

after defining

$$\begin{aligned} &W(x) := \ell(x)h(x) + \epsilon \\ &S(x) := \tfrac{1}{\sqrt{2}}\begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top, \ R(x) := \tfrac{1}{\sqrt{2}}\tfrac{\partial h}{\partial x}(x)^\top. \end{aligned}$$

(4.33), and thus (4.32), is equivalent, by (4.31), to

$$\begin{aligned}
&\forall Y \text{ with } Y^\top Y \preceq I,\ \forall x \in \mathbb{R}^n,\\
&W(x) + S(x)(\overline{\zeta}_{\rm i} + \overline{P}_{\rm i} Y \overline{Q}_{\rm i}^{1/2}) R(x)\\
&\ + R(x)^\top (\overline{\zeta}_{\rm i} + \overline{P}_{\rm i} Y \overline{Q}_{\rm i}^{1/2})^\top S(x)^\top\\
&= W(x) + S(x)\overline{\zeta}_{\rm i} R(x) + R(x)^\top \overline{\zeta}_{\rm i}^\top S(x)^\top\\
&\ + S(x)\overline{P}_{\rm i} Y \overline{Q}_{\rm i}^{1/2} R(x) + R(x)^\top \overline{Q}_{\rm i}^{1/2} Y^\top \overline{P}_{\rm i} S(x)^\top \preceq 0.
\end{aligned} \tag{4.34}$$

If there exist $\eta \in \Pi$ with $\eta(x) > 0$ for all $x \in \mathbb{R}^n$, we have

$$\begin{aligned}
&\forall Y \text{ with } Y^\top Y \preceq I, \forall x \in \mathbb{R}^n,\\
&S(x)\overline{P}_{\rm i} Y \overline{Q}_{\rm i}^{1/2} R(x) + R(x)^\top \overline{Q}_{\rm i}^{1/2} Y^\top \overline{P}_{\rm i} S(x)^\top\\
&\preceq \eta(x) S(x)\overline{P}_{\rm i}\,\overline{P}_{\rm i} S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_{\rm i}^{1/2} Y^\top Y \overline{Q}_{\rm i}^{1/2} R(x)\\
&\preceq \eta(x) S(x)\overline{P}_{\rm i}^2 S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_{\rm i} R(x)
\end{aligned}$$

where we use Young's inequality in the first upperbound and $Y^\top Y \preceq I$ in the second upperbound. Using this last upperbound in (4.34), we obtain that (4.34) is implied by the existence of η, with $\eta(x) > 0$ for all $x \in \mathbb{R}^n$, such that

$$\begin{aligned}
\forall x \in \mathbb{R}^n,\ &W(x) + S(x)\overline{\zeta}_{\rm i} R(x) + R(x)^\top \overline{\zeta}_{\rm i}^\top S(x)^\top\\
&+ \eta(x) S(x)\overline{P}_{\rm i}^2 S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_{\rm i} R(x) \preceq 0.
\end{aligned}$$

Replacing the explicit expressions of $W(x), S(x), R(x)$ in this inequality, we conclude that the invariance condition (4.32) holds if for all $x \in \mathbb{R}^n$, $\eta(x) > 0$ and

$$\begin{aligned}
0 \ge\ &\ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)\overline{\zeta}_{\rm i}^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}\\
&+ \tfrac{\eta(x)}{2} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top \overline{P}_{\rm i}^2 \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{1}{2\eta(x)} \tfrac{\partial h}{\partial x}(x)\overline{Q}_{\rm i} \tfrac{\partial h}{\partial x}(x)^\top.
\end{aligned}$$

This last condition is equivalent to having for all $x \in \mathbb{R}^n$, $\eta(x) > 0$ and $H(x) \preceq 0$, as one can easily verify applying Schur complement [51, p. 28] on $H(x) \preceq 0$. Hence, (4.32) holds, as we needed to complete the proof. ∎

Remark 4 *In Remark 3, we commented on the possibility of having an energy bound on the disturbance. In that case, the conclusion of Theorem 7 continues to hold if the hypothesis of Theorem 7 is slightly adapted as follows:*

For a design parameter $\epsilon > 0$*, consider the data generation mechanism* (4.9)*, measured data* $\{\dot{x}^j, z^j, v^j\}_{j=0}^{T-1}$ *as in* (4.8) *and disturbance* d *satisfying the* energy bound $\mathscr{D}_\mathrm{e}$ in (4.24) *(i.e.,* $\begin{bmatrix} d^0 & \dots & d^{T-1}\end{bmatrix} \in \mathscr{D}_\mathrm{e}$*).* Assume that $\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}$ has full row rank and parameters $\overline{\zeta}_\mathrm{i}$, $\overline{P}_\mathrm{i}$ and $\overline{Q}_\mathrm{i}$ are equal, *instead of* (4.20), to respectively

$$\overline{\zeta}_\mathrm{e} := -A_\mathrm{e}^{-1} B_\mathrm{e}, \overline{P}_\mathrm{e} := A_\mathrm{e}^{-1/2}, \overline{Q}_\mathrm{e} := B_\mathrm{e}^\top A_\mathrm{e}^{-1} B_\mathrm{e} - C_\mathrm{e} \tag{4.35}$$

for A_e*,* B_e*,* C_e *in* (4.27)*. Assume that there exist decision variables* $\ell \in \Pi$*,* $\eta \in \Pi$*,* $h \in \Pi$ *and* $K \in \Pi_{m,1}$ *such that for all* $x \in \mathbb{R}^n$*,* $\eta(x) > 0$ *and* $H(x) \preceq 0$*, with* H *defined in* (4.29), where $\overline{\zeta}_\mathrm{i}$, $\overline{P}_\mathrm{i}$, $\overline{Q}_\mathrm{i}$ are respectively equal to (4.35).
Then, the set $\mathscr{I}$ *in* (4.10) *is invariant for the system in* (4.6).

This adaptation of Theorem 7 is proven by observing that: (i) $\overline{P}_\mathrm{e}$ *and* $\overline{Q}_\mathrm{e}$*, which are used in place of respectively* $\overline{P}_\mathrm{i}$ *and* $\overline{Q}_\mathrm{i}$*, satisfy* $\overline{P}_\mathrm{e} \succ 0$ *and* $\overline{Q}_\mathrm{e} \succeq 0$ *by the full row rank of* $\begin{bmatrix} Z_0 \\ V_0 \end{bmatrix}$ *and [82, Lemma 1]; (ii) the set* $\mathscr{C}_\mathrm{e}$ *in* (4.26)*, resulting from* $\mathscr{D}_\mathrm{e}$*, can be written in the form* (4.31) *by [82, Prop. 1]; (iii) the rest of the proof of Theorem 7 is the same.*

In this chapter, we have assumed that the first time derivative is measurable. To collect the sequence of $\dot{x}$ it is sufficient to compute the value of $\dot{x}$ at each sample time using the sequence of x as initial values and the same input signal used to collect x. Indeed, it can be recovered from state measurements when the states are sampled densely by finite differences. More sophisticated techniques are devised in the context of continuous-time system identification, see for instance [87] for details. Admittedly, all these techniques do not reconstruct the time derivative exactly, but only with some error. This is something our framework takes care of through the process disturbance d in (4.7) considered in data collection.

Additionally, one can characterize the set of matrices consistent with data starting from the integrals of the state x and of the input u [88]. The main advantage of this integral formulation (over finding the first time derivative from the state measurements) is that the involved integrals are expected to average out noise (instead of amplifying it).For these reasons, we believe that assuming to measure the first time derivative of the state is not as restrictive as it may seem.

Next we discuss how the problem can be solved when we consider measuring the state integrals. Computing these integrals can be done numerically from samples, possibly with some error that can be incorporated into the disturbance term.

With the set $\mathscr{D}_{\text{ins}} := \{d \in \mathbb{R}^n : |d|^2 \le \omega\}$ and times $t_0, \dots, t_{T-1}$, one can obtain that for $j = 0, \dots, T-2$

$$\begin{aligned}
&\Big(\int_{t_j}^{t_{j+1}} d(t)\mathrm{d}t\Big)^\top \Big(\int_{t_j}^{t_{j+1}} d(t)\mathrm{d}t\Big) = \Big|\int_{t_j}^{t_{j+1}} d(t)\mathrm{d}t\Big|^2 = \Big|\int_{t_j}^{t_{j+1}} d_1(t)\mathrm{d}t\Big|^2 + \\
&\qquad + \dots + \Big|\int_{t_j}^{t_{j+1}} d_n(t)\mathrm{d}t\Big|^2 \\
&\le \int_{t_j}^{t_{j+1}} d_1(t)^2\mathrm{d}t \int_{t_j}^{t_{j+1}} \mathrm{d}t + \dots + \int_{t_j}^{t_{j+1}} d_n(t)^2\mathrm{d}t \int_{t_j}^{t_{j+1}} \mathrm{d}t = (t_{j+1} - t_j)\Big(\int_{t_j}^{t_{j+1}} |d(t)|^2\mathrm{d}t\Big) \\
&\qquad \le (t_{j+1} - t_j)^2\omega
\end{aligned}$$

4

where the first bound is obtained by the Cauchy-Schwarz inequality and the second bound uses that $d(t) \in \mathscr{D}_{\text{ins}}$ for all $t \ge 0$. Then, since the data generation mechanism is $\dot{x} = \mathsf{A}Z(x) + \mathsf{B}W(x)u + d$, we have that data satisfy, for $j = 0, \dots, T-2$,

$$\begin{aligned}
&\tilde{x}^j := x(t_{j+1}) - x(t_j) = \int_{t_j}^{t_{j+1}} \dot{x}(t)\mathrm{d}t = \int_{t_j}^{t_{j+1}} \big(\mathsf{A}Z(x(t)) + \mathsf{B}W(x(t))u(t) + d(t)\big)\mathrm{d}t \\
&= \mathsf{A}\int_{t_j}^{t_{j+1}} Z(x(t))\mathrm{d}t + \mathsf{B}\int_{t_j}^{t_{j+1}} W(x(t))u(t)\mathrm{d}t + \int_{t_j}^{t_{j+1}} d(t)\mathrm{d}t =: \mathsf{A}\tilde{z}^j + \mathsf{B}\tilde{v}^j + \int_{t_j}^{t_{j+1}} d(t)\mathrm{d}t.
\end{aligned} \tag{4.36}$$

With this characterization, we can write a set of matrices consistent with the disturbance bound and data points on interval $[t_j, t_{j+1}]$ for $j = 0, \dots, T-2$ as

$$\tilde{\mathscr{C}}_{\mathrm{i}}^j := \Big\{[A\ B] : \tilde{x}^j = A\tilde{z}^j + B\tilde{v}^j + \tilde{d},\ \tilde{d}\tilde{d}^\top \preceq (t_{j+1} - t_j)^2\omega I\Big\}.$$

With this form, the very same developments presented in this section can be carried out due to the analogy with the set $\mathscr{C}_{\mathrm{i}}^j$ (4.16).

DISCRETE SYSTEMS

For discrete-time systems, it does not seem straightforward to obtain a sum-of-squares relaxation following the results presented in this chapter for continuous-time systems. Indeed, consider the discrete-time polynomial dynamics

$$x^+ = AZ(x) + BW(x)K(x) = f_{\text{cl}}(x) \tag{4.37}$$

where the polynomial control law K needs to be determined. Consider also $\mathscr{I} := \{x \in \mathbb{R}^n : h(x) \le 0\}$ as the set to be rendered invariant for a polynomial h, as in (4.10). The set $\mathscr{I}$ is invariant for (4.37) if $f_{\text{cl}}(\mathscr{I}) \subseteq \mathscr{I}$ (Definition 7) or, equivalently,

$f_{\text{cl}}(x) \in \mathscr{I}$ for all $x \in \mathscr{I}$ or, equivalently, $h(f_{\text{cl}}(x)) \le 0$ for all x such that $h(x) \le 0$. This holds if, for some polynomial s, we have that for all x, $s(x) \ge 0$ and

$$0 \ge h(f_{\text{cl}}(x)) - s(x)h(x) = h\big(AZ(x) + BW(x)K(x)\big) - s(x)h(x). \tag{4.38}$$

Now, if A and B were known as in the model-based case, one could envision an alternate procedure where (i) for fixed h, we find K and s such that $s(x) \ge 0$ and (4.38) hold for all x; (ii) for fixed K and s, we find h such that (4.38) holds for all x. However, if we do not know A and B due to data-induced uncertainty and we only know that they belong to the set in (4.21), it seems hard to provide a condition depending only on the data-related parameters of the set in (4.21) based on (4.38), in general.

4.5.1. SAFETY CONSTRAINTS

As discussed in Section 4.1, the invariance condition obtained in Theorem 7 is instrumental to design safe control laws in applications. To effectively link invariance and safe control we introduce the so-called safe set $\mathscr{S}$, by which a user can specify all constraints on the state. Formally, these constraints are expressed by negativity of polynomials $\sigma_1, \dots, \sigma_q$ ($q \in \mathbb{Z}_{\ge 1}$) and the safe set $\mathscr{S}$ is

$$\mathscr{S} := \{x \in \mathbb{R}^n : \sigma_j(x) \le 0, j = 1, \dots, q\}. \tag{4.39}$$

Hence, when designing the invariant set $\mathscr{I}$, one needs to enforce the condition $\mathscr{I} \subseteq \mathscr{S}$ so that when the state belongs to $\mathscr{I}$, it will also comply with all constraints expressed by $\mathscr{S}$. At the same time, it is of interest not only to impose $\mathscr{I} \subseteq \mathscr{S}$, but to ensure that $\mathscr{I}$ is as "large" as possible. Using a classical approach as in, e.g., [72], define the set $\mathscr{L}_\theta$ for a nonnegative polynomial λ (i.e., $\lambda(x) \ge 0$ for all x) and a nonnegative scalar θ as

$$\mathscr{L}_\theta := \{x \in \mathbb{R}^n : \lambda(x) \le \theta\}. \tag{4.40}$$

With $\mathscr{L}_\theta$, $\mathscr{I}$ can be enlarged by imposing $\mathscr{L}_\theta \subseteq \mathscr{I}$ while maximizing $\theta \ge 0$; hence, $\mathscr{L}_\theta$ acts as a variable-size set that dilates $\mathscr{I}$ from the inside according to the shape given by the design parameter λ, which can be chosen based on the form of the safe set $\mathscr{S}$. This approach will be exemplified in Section 4.7, in figure 4.1 you can see a result from Section 4.7 where three sets $\mathscr{S}$, $\mathscr{I}$ and $\mathscr{L}_\theta$ are shown and it is possible to visualized the relation $\mathscr{L}_\theta \subseteq \mathscr{I} \subseteq \mathscr{S}$.

Moreover, we strengthen the positivity conditions of Theorem 7 into more conservative, but computationally tractable, sum-of-squares conditions. Through the safe set $\mathscr{S}$, the variable-size set $\mathscr{L}_\theta$, the strengthening of positivity conditions of Theorem 7 and definition

$$r := 1 + m + 2n,$$

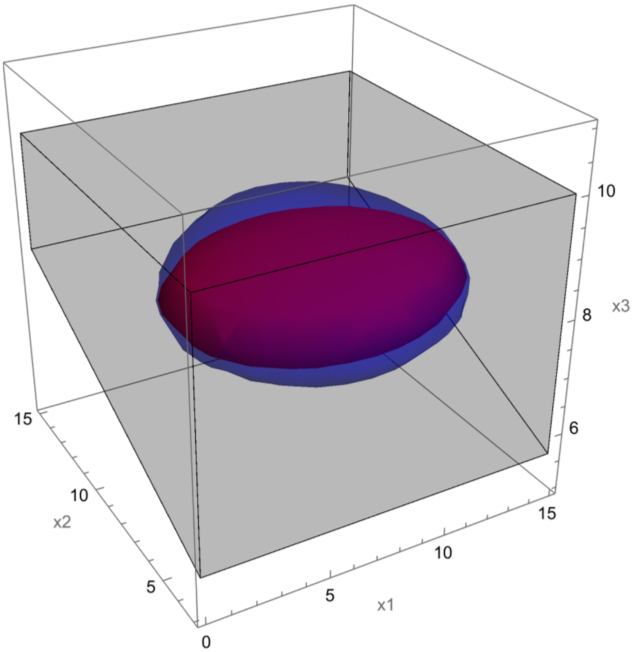

Figure 4.1 | This picture highlight the role of the variable size set $\mathscr{L}_\theta$ (red) in enlarging the invariant set (blue). The invariant set must contains $\mathscr{L}_\theta$ and by increasing θ we can induce the invariant set to grow. The safe set is shown in light gray. Picture obtain with $|d| \leq 10^{-4}$, $T = 500$.

we have the next result for data-driven safe control.

Theorem 8 ***(Data-driven safe control)***
For given polynomials $\sigma_1, \ldots, \sigma_q$ and nonnegative polynomial λ, consider the set $\mathscr{S}$ in (4.39) *and the set $\mathscr{L}_\theta$ in* (4.40) *parametrized by θ, along with H defined in* (4.29) *and a given $\bar{\eta} > 0$.*
Under the same hypothesis of Theorem 7, assume the program

$$
\begin{aligned}
&\text{maximize} && \theta && \text{(4.41a)}\\
&\text{s.t.} && \{\ell, \eta, h\} \subseteq \Pi, K \in \Pi_{m,1}, \{s_1, \ldots, s_q, \varsigma\} \subseteq \Sigma, \theta \geq 0 && \text{(4.41b)}\\
& && \eta - \bar{\eta} \in \Sigma, -H \in \Sigma_r && \text{(4.41c)}\\
& && \varsigma(\lambda - \theta) - h \in \Sigma, && \text{(4.41d)}\\
& && s_j h - \sigma_j \in \Sigma, \; j = 1, \ldots, q. && \text{(4.41e)}
\end{aligned}
$$

4

has a solution. Then, the set $\mathscr{I}$ in (4.10) *is invariant for the system in* (4.6) *and satisfies $\mathscr{L}_\theta \subseteq \mathscr{I} \subseteq \mathscr{S}$.*

Proof. (4.41c) implies that for all $x \in \mathbb{R}^n$, $\eta(x) \geq \bar{\eta} > 0$ and $H(x) \preceq 0$ (by Definition 3) and, under the same hypothesis of Theorem 7, these two conditions were shown in Theorem 7 to imply that $\mathscr{I}$ is invariant for (4.6). The statement is then proven if we show that with (4.41b), (4.41d) implies $\mathscr{L}_\theta \subseteq \mathscr{I}$ and (4.41e) implies $\mathscr{I} \subseteq \mathscr{S}$. Since the reasoning is the same, we show only the latter. $\mathscr{I} \subseteq \mathscr{S}$ is equivalent to the set inclusion

$$\{x \in \mathbb{R}^n : \; h(x) \leq 0\} \subseteq \{x \in \mathbb{R}^n : \; \sigma_j(x) \leq 0\}$$

holding for all $j = 1, \ldots, q$ or, equivalently, to the empty-set condition

$$\{x \in \mathbb{R}^n : \; -h(x) \geq 0, \sigma_j(x) \geq 0, \sigma_j(x) \neq 0\} = \emptyset$$

holding for all $j = 1, \ldots, q$ or, equivalently by Fact 1, to the existence, for all $j = 1, \ldots, q$, of polynomials $s_{j,0}$, $s_{j,1}$, $s_{j,2}$, $s_{j,3}$ in Σ and $k_j \in \mathbb{Z}_{\geq 0}$, $j = 1, \ldots, q$ such that

$$s_{j,0} - s_{j,1} h + s_{j,2} \sigma_j - s_{j,3} h \sigma_j + \sigma_j^{2k_j} = 0.$$

This is implied by the existence, for $j = 1, \ldots, q$, of $k_j = 1$, $s_{j,0} = 0$, $s_{j,1} = 0$ and $s_{j,2}$, $s_{j,3}$ in Σ such that $\sigma_j(s_{j,2} - s_{j,3} h + \sigma_j) = 0$, which is implied by (4.41e). ■

To conclude the section, we show in the next remark how input constraints can be readily incorporated in the proposed design to account for actuator limitations.

Remark 5 *Suppose that input u needs to be bounded in norm, i.e., $|u| \leq u_{\mathrm{M}}$ for some $u_{\mathrm{M}} > 0$. This constraint is enforced by imposing that for each x, $|K(x)| \leq u_{\mathrm{M}}$, which is equivalent to $K(x)^\top K(x) \leq u_{\mathrm{M}}$ and $\begin{bmatrix} u_{\mathrm{M}} & K(x)^\top \\ K(x) & I \end{bmatrix} \succeq 0$. This condition can be relaxed as $\begin{bmatrix} u_{\mathrm{M}} & K^\top \\ K & I \end{bmatrix} \in \Sigma_{m+1}$ and added to the conditions in Theorem 8.*

4

4.6. NUMERICAL EXAMPLE: VAN DER POL OSCILLATOR

In this section, we show the ability of Theorem 7 to find an invariant set and a suitable state feedback controller.

To find a solution to the conditions of Theorem 7, we turn the positivity conditions into more conservative, but tractable, sum-of-squares conditions. Specifically, we replace the positive semidefiniteness condition (4.29), i.e., $-H(x) \succeq 0$ for all $x \in \mathbb{R}^n$, with the requirement that $-H$ be an SOS matrix polynomial, i.e., $-H \in \Sigma_r$ with $r := 1+m+2n$; we also replace $\eta(x) > 0$ for all $x \in \mathbb{R}^n$ with $\eta \in \Sigma$. We note that $\eta \in \Sigma$ implies only nonnegativity, and not positivity, of η; however, as pointed out in [89, p. 41], interior-point algorithms find automatically $\eta \in \Sigma$ that is strictly positive, if it exists. Hence, we impose $\eta \in \Sigma$ and verify a-posteriori its positivity, rather than imposing $\eta - \bar{\eta} \in \Sigma$ for a small $\bar{\eta} > 0$. Besides these SOS relaxations, due to the products between decision variables in H, the optimization is solved in two alternate steps.

The discussed strategy is summarized in Algorithm 1.

Algorithm 1 SOS relaxation of Theorem 7

Initialize: $c = 1$, $K(x) = 0$, $\ell(x) = 1$.

Find $h \in \mathscr{P}, \eta \in \mathscr{P}$
subject to $\eta \in \Sigma$, $-H \in \Sigma_r$.
Update h, η.

Find $\ell \in \mathscr{P}$, F
subject to $-H \in \Sigma_r$.

The SOS programs are solved in Matlab with YALMIP [90] [75] and MOSEK. The degrees of the polynomials ℓ, η, h are respectively $2,2,4$. The proposed solution was tested with two kinds of van der Pol oscillator.

CASE 1

We consider a van der Pol oscillator with unstable origin and *stable* limit cycle for $u = 0$:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + (1 - x_1^2)x_2 + u. \end{aligned} \tag{4.42}$$

The SOS program is solved with a collection of $T = 20$ samples. The obtained solution consist of a controller $K(x)$ and an invariant set $\mathscr{I} = \{x \in \mathbb{R}^2 : h(x) - 1 \le 0\}$

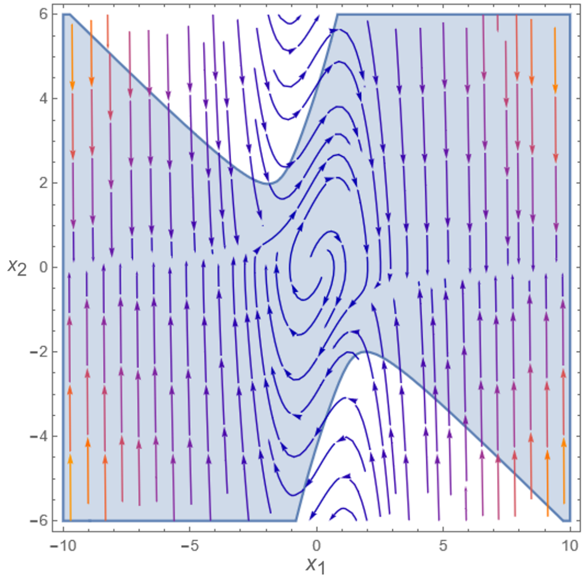

Figure 4.2 | Closed-loop phase plot of Case 1 of Section 4.6 with invariance set (4.44) shown in light blue, state-feedback controller (4.43) and bounded disturbance $|d|^2 \leq 0.1$.

with

$$K(x) = 10^{-3}(-3.2x_1 - 0.37x_2 - 0.34x_1^2 + 0.32x_1x_2) \tag{4.43}$$

$$\begin{aligned} h(x) = &-2.879x_1^2 + 0.695x_1x_2 - 5.773x_2^2 - 0.001x_1^3 + \\ &-0.002x_1^2x_2 - 0.017x_1x_2^2 + 0.007x_2^3 - 1.522x_1^4 + \\ &-1.470x_1^3x_2 + 0.845x_1^2x_2^2 - 1.303x1x_2^3 + 0.327x_2^4. \end{aligned} \tag{4.44}$$

In this and all results, we remove the smaller coefficients ($< 10^{-5}$). In Figure 4.2, we can see the invariant set plotted over the trajectories of system (4.42) with $u = K(x)$.

CASE 2

We consider a van der Pol oscillator with stable origin and *unstable* limit cycle for $u = 0$:

$$\begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= x_1 - (1 - x_1^2)x_2 + u. \end{aligned} \tag{4.45}$$

The SOS program is solved with a collection of $T = 20$ samples with a sample time of 0.5 seconds. The obtained solution consist of a controller $K(x)$ and an invariant

4

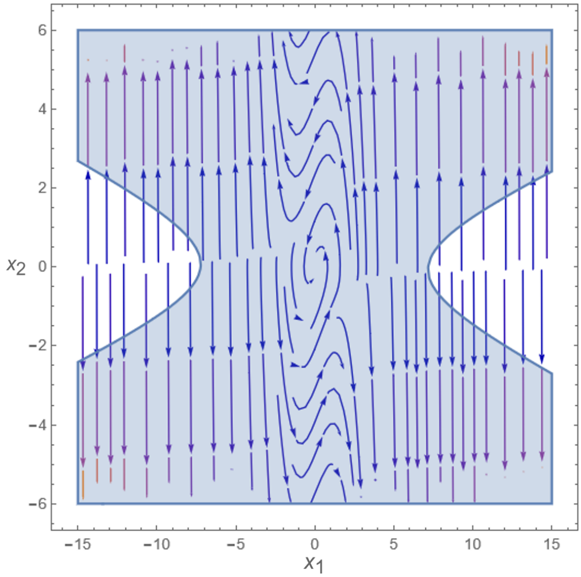

Figure 4.3 | Closed-loop phase plot of Case 2 of Section 4.6 with invariance set (4.47) shown in light blue, state-feedback controller (4.46) and bounded disturbance $|d|^2 \leq 0.01$.

set $\mathscr{I} = \{x \in \mathbb{R}^2 \colon h(x) - 1 \leq 0\}$ with

$$K(x) = 10^{-3}(6.6x_1 + 4.9x_2), \tag{4.46}$$

$$\begin{aligned} h(x) = &-112.45x_1^2 + 24.66x_1x_2 - 38.56x_2^2 + 0.11x_1^3 + \\ &- 0.27x_1^2x_2 + 0.04x_1x_2^2 + 0.0007x_2^3 + 2.16x_1^4 + \\ &- 1.58x_1^3x_2 - 56.28x_1^2x_2^2 + 13.18x_1x_2^3 - 30.41x_2^4. \end{aligned} \tag{4.47}$$

In figure 4.3, we can see the invariant set plotted over the trajectories of system (4.45) with $u = K(x)$.

4.7. NUMERICAL EXAMPLE: CAR PLATOONING

As a safety-critical system, we consider two cars moving in a platoon formation. The system can be modeled as

$$\dot{x}_1 = u_1 - \gamma_1 - \beta_1 x_1 - \alpha_1 x_1^2 \tag{4.48a}$$

$$\dot{x}_2 = u_2 - \gamma_2 - \beta_2 x_2 - \alpha_2 x_2^2 \tag{4.48b}$$

$$\dot{x}_3 = x_1 - x_2 \tag{4.48c}$$

where: the components x_1, x_2, x_3 of state x represent respectively the velocity of the front vehicle, the velocity of the rear vehicle and the relative distance between the two; the components u_1, u_2 of input u represent the forces normalized by vehicle mass; γ_k, β_k, α_k for $k = 1,2$ are the static, rolling and aerodynamic-drag friction coefficients normalized by vehicle mass. We impose safety constraints by the set

$$\begin{aligned}\mathscr{S} := \{x \in \mathbb{R}^3 \colon\ & d_0 + \tau_{\rm h} x_2 \le x_3,\ x_3 \le d_1,\\ & 0 \le x_1,\ x_1 \le v_{\rm M},\ 0 \le x_2,\ x_2 \le v_{\rm M}\}\\ =: \{x \in \mathbb{R}^3 \colon\ & \sigma_1(x) \le 0, \dots, \sigma_6(x) \le 0\} \end{aligned} \tag{4.49}$$

where $d_0 + \tau_{\rm h} x_2$ is a relative distance required to avoid collisions with the front vehicle ($d_0 > 0$ is a standstill distance and $\tau_{\rm h} > 0$ is a time headway), d_1 is a distance required to keep the benefits of platooning (especially, aerodynamic drag reduction), and $v_{\rm M}$ is a maximum velocity allowed on the road. Finally, we consider as point of interest $\bar{x} := (\bar{v}, \bar{v}, \bar{d})$ where $\bar{v}$ is a predefined cruise velocity and $\bar{d}$ is a reference safety distance. Numerical values of parameters are in Table 4.1.

γ_1	$0.005\,\frac{\rm N}{\rm kg}$	β_1	$0.1\,\frac{\rm Ns}{\rm kgm}$	α_1	$0.02\,\frac{\rm Ns^2}{\rm kgm^2}$
γ_2	$0.005\,\frac{\rm N}{\rm kg}$	β_2	$0.2\,\frac{\rm Ns}{\rm kgm}$	α_2	$0.04\,\frac{\rm Ns^2}{\rm kgm^2}$
$\tau_{\rm h}$	0.2s	d_0	5m	d_1	10m
$v_{\rm M}$	$22.2\,\frac{\rm m}{\rm s}$	$\bar{v}$	$8.5\,\frac{\rm m}{\rm s}$	$\bar{d}$	8m

Table 4.1 | Values of the parameters for the two cars platoon simulations in Section 4.7.

The numerical program to find invariant set and controller is in Algorithm 2 and is implemented in Matlab with YALMIP [75, 90] and MOSEK. We now comment Algorithm 2, which consists of an initialization (lines 1-2) and a main part (lines 3-15) made of two steps.

As for the initialization of Algorithm 2, we set the degrees of polynomials h, η, s_j ($j = 1, \dots, 6$) and ς to 4, 2, 2, 2. Moreover, the shape of the safe set $\mathscr{S}$ in (5.27) (light blue set in Figure 4.4) is wider in the coordinates x_1 and x_2 and narrow in the coordinate x_3. Hence, we dilate the invariant set $\mathscr{I}$ through an ellipsoid shaped similarly to $\mathscr{S}$ and with center $\bar{x}$ since we would like $\mathscr{I}$ to contain $\bar{x}$; i.e., we dilate $\mathscr{I}$ through

$$\mathscr{L}_\theta := \{x \in \mathbb{R}^3 \colon (x - \bar{x})^\top \begin{bmatrix} 0.02 & 0 & 0 \\ 0 & 0.05 & 0 \\ 0 & 0 & 1 \end{bmatrix} (x - \bar{x}) =: \lambda(x) \le \theta\}$$

as in (4.40). Finally, a guess of h is needed to initialize the iterations in the procedure. In a data-based fashion, we can use as a guess of h the Lyapunov function obtained *from data* by using [82, Th. 2] and performing a preliminary experiment with "small" input and state signals around the equilibrium $\bar{x}$.

Algorithm 2 Car Platooning

1: Excite the system around an equilibrium point of interest, collect input/state data and obtain a Lyapunov function for the linearized system from data by [82, Th. 2].
2: **initialize:** $\iota = 0$ (a counter), $\theta_\iota = \theta_0 = 0.01$, $\epsilon = 0.01$, $\eta(x) = 1$, h equal to the Lyapunov function found above.
3: **repeat**
4: Find $\ell \in \Pi$, $K \in \Pi_{m,1}$ and $s_1, \dots, s_6, \varsigma \in \Sigma$
5: subject to $-H \in \Sigma_r$
6: $\varsigma(\lambda - \theta_\iota) - h \in \Sigma$
7: $s_j h - \sigma_j \in \Sigma, \quad j = 1, \dots, 6.$
8: Update ℓ, K, $s_1, \dots, s_6$, ς.
9:
10: Maximize θ over $\eta, h \in \Pi$ and $\theta \geq 0$,
11: subject to $\eta \in \Sigma$, $-H \in \Sigma_r$
12: $\varsigma(\lambda - \theta) - h \in \Sigma$
13: $s_j h - \sigma_j \in \Sigma, \quad j = 1, \dots, 6.$
14: Update η, h, $\theta_\iota = \theta$, $\iota \leftarrow \iota + 1$.
15: **until** $\theta_\iota - \theta_{\iota-1}$ is less than some tolerance (10^{-3}).

As for the main part of Algorithm 2, it corresponds to the SOS program (4.41) in Theorem 8. However, since (4.41) presents products between decision variables, we first fix h, η and θ in (4.41) and solve for the other decision variables (lines 4-8), and then fix ℓ, K and s_1, ..., s_6, ς in (4.41) and solve for the other decision variables (lines 10-14). Moreover, we asked $\eta - \bar{\eta} \in \Sigma$ in Theorem 8 for (small) $\bar{\eta} > 0$; here, we ask the weaker condition $\eta \in \Sigma$ because, if the constraint $\eta \in \Sigma$ is feasible, interior-point algorithms automatically find [89, p. 41] a strictly positive η, hence satisfying $\eta - \bar{\eta} \in \Sigma$, which we verified a-posteriori. Finally, since $-H \in \Sigma_r$ in (4.41) is homogeneous with respect to (h, η), we prune solutions by fixing the 0-degree coefficient of h to a *given* constant.

We also remark that in H the terms Z and W appear. The choice of the monomials considered in (4.5) for Z and W is important to obtain the best result from our solution. The simplest choice is to consider all monomials in x_1, x_2, x_3 up to a maximum degree; with noisy data, however, the coefficient of each monomial becomes uncertain and coping with it results in more conservative solutions. A smarter choice is to include high-level prior knowledge [91]. For platooning, we use $Z(x) = [x_1\ x_2\ x_3\ x_1^2\ x_2^2]^\top$ and $W(x) = I$, since we know from physical considerations that the time derivative of the relative distance x_3 depends only on the velocities x_1 and x_2, and $\dot{x}_1$, $\dot{x}_2$ depend only on x_1, x_1^2, x_2, x_2^2 and u.

$$h(x) = 90.30 + 202.66x_1^2 + 364.18x_2^2 - 458.07x_1x_2 + 28.19x_1x_3 - 283.29x_2x_3 + 158.34x_3^2 \\ - 0.95x_1^3 - 0.63x_1^2x_2 - 47.29x_1^2x_3 + 3.70x_1x_2^2 - 3.89x_2^3 - 83.13x_2^2x_3 + 108.33x_1x_2x_3 - \\ 7.36x_1x_3^2 + 65.74x_2x_3^2 - 38.24x_3^3 + 0.10x_1^4 - 0.10x_1^3x_2 + 0.36x_1^2x_2^2 - 0.25x_1x_2^3 + 0.29x_2^4 \\ - 0.21x_1^3x_3 - 0.36x_1^2x_2x_3 - 0.43x_1x_2^2x_3 - 0.46x_2^3x_3 + 3.47x_1^2x_3^2 - 5.93x_1x_2x_3^2 \\ + 6.15x_2^2x_3^2 - 0.18x_1x_3^3 - 4.95x_2x_3^3 + 2.75x_3^4. \tag{4.50}$$

By using Algorithm 2, we obtain an invariant set $\mathscr{I} = \{x \in \mathbb{R}^3 : h(x) \le 0\}$ with h as in (4.50), displayed over two columns, and a controller $u = K(x) = \begin{bmatrix} K_1(x) \\ K_2(x) \end{bmatrix}$ with

4

$$K_1(x) = 14.39x_3 - 10.45x_1 + 44.05x_2 - 0.85x_1^2 - 4.24x_2^2,$$
$$K_2(x) = 71.28x_3 - 2.47x_1 - 23.32x_2 + 0.10x_1^2 - 4.09x_2^2.$$

We removed monomials coefficients smaller than 10^{-5} in h and K.

We compared our data-driven solution in Algorithm 2 against a model-based solution in which perfect model knowledge is assumed. Thereby providing a baseline for what we can achieve with the data-based scheme. The model-based implementation, which is the counterpart of (4.41), corresponds to

$$\begin{array}{ll} \text{maximize} & \theta \\ \text{s.t.} & \{\ell, h\} \subseteq \Pi, K \in \Pi_{m,1}, \{s_1, \dots, s_6, \varsigma\} \subseteq \Sigma, \theta \ge 0 \\ & -(\ell h + \epsilon + \frac{\partial h}{\partial x} [\mathsf{A} \ \mathsf{B}] \begin{bmatrix} Z \\ WK \end{bmatrix}) \in \Sigma \\ & \varsigma(\lambda - \theta) - h \in \Sigma,\ s_j h - \sigma_j \in \Sigma, j = 1, \dots, 6. \end{array}$$

In Figure 4.4, we can see that the model-based and the data-driven solutions are comparable as for the sizes of the resulting invariant sets for data points affected by a disturbance satisfying $|d| \le 10^{-3}$. In both cases safety constraints are not violated since both invariant sets are within $\mathscr{S}$. In Figure 4.5, the invariant set $\mathscr{I}$ of the data-driven solution is plotted together with trajectories corresponding to the vector field (4.48) in closed-loop with controller $u = K(x)$. Trajectories are initialized close to the boundary of $\mathscr{I}$ to show that once in the set $\mathscr{I}$, they never leave it, thereby confirming that $\mathscr{I}$ is invariant.

4.8. CONCLUSIONS

In this chapter, we addressed the problem of enforcing invariance for a polynomial system based on data. We assumed that the open-loop data are corrupted by

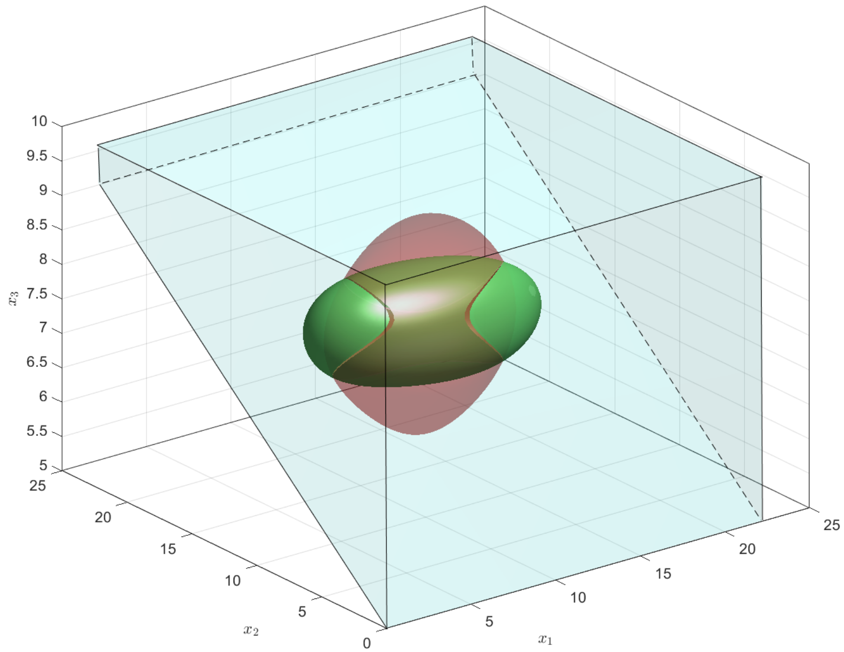

Figure 4.4 | Invariant set for the model-based case in light red and for the data-driven case in green ($|d| \leq 10^{-3}$, $T = 1000$). The safe set in light blue highlights that the two invariant sets comply with safety constraints.

noise, whose nature and characteristic are unknown except for an (instantaneous) bound. The presence of noise resulted then in a set of dynamics consistent with data, which we over-approximated via matrix ellipsoids and took into account in the design. Our solution provided a data-dependent SOS optimization program to obtain a state feedback controller and an invariant set for the closed-loop system, and we optimized the size of the invariant set under the constraint that it remains contained in a user-defined safety set. Finally, we tested our data-driven algorithm on a platooning example where we showed that, for a reasonable noise level, our solution compares well with the case of perfect model knowledge.

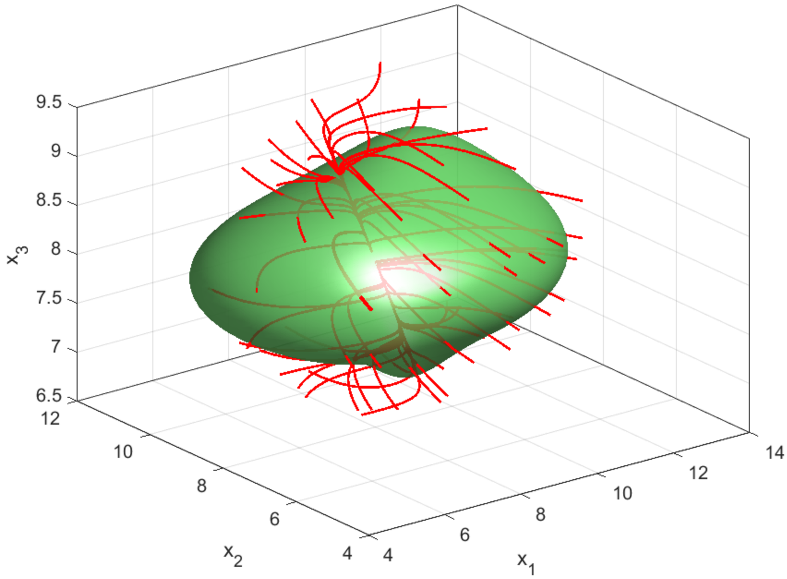

Figure 4.5 | To show that the invariant set $\mathscr{I}$ (green) given by (4.50) is invariant, we simulate the closed-loop system with initial conditions slightly outside $\mathscr{I}$, close to the boundary. All simulated trajectories (red) never leave the set after entering it.

5

Data-driven design of robust safe control for polynomial systems

---

Abstract

In this chapter we continue the discussion on the design of safe controllers for polynomial systems, but with one key difference. Instead of considering the condition of nominal invariance for the controlled closed-loop system we explicitly take into account the disturbance during closed-loop operation in the design by using a robust invariance condition. In this way the resulting controller guarantees positive invariance of a certain set even in case of disturbances acting on the system during operation. The designed robust controller is then compared with the controller presented in Chapter 4 for the control of a platoon of cars with safety constraints.

5.1. Preliminaries

We consider perturbed polynomial systems of the form

$$\dot{x} = f(x) + g(x)u + d \tag{5.1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $d \in \mathbb{R}^n$ is the disturbance, and f and g are polynomial vector fields. The specific expressions of f and g are unknown, and d is also unknown but satisfies

$$\mathscr{D}_{\text{ins}} := \{d \in \mathbb{R}^n : |d|^2 \leq \omega\}. \tag{5.2}$$

Following the same procedure used in the previous chapter we can write the polynomial system (5.1) into a linear-like form

$$\dot{x} = \mathsf{A}Z(x) + \mathsf{B}W(x)u + d \tag{5.3}$$

where $\mathsf{A} \in \mathbb{R}^{n\times N_\mathsf{A}}$ and $\mathsf{B} \in \mathbb{R}^{n\times N_\mathsf{B}}$ are *unknown* constant matrices, the *known* $N_\mathsf{A} \times 1$ vector $Z(x)$ contains as entries the distinct monomials in x that may appear in f, and the *known* $N_\mathsf{B} \times m$ matrix $W(x)$ contains as entries the monomials that may appear in g.

The closed-loop dynamics results in

$$\dot{x} = \mathsf{A}Z(x) + \mathsf{B}W(x)K(x) + d = \begin{bmatrix}\mathsf{A} & \mathsf{B}\end{bmatrix}\begin{bmatrix}Z(x)\\ W(x)K(x)\end{bmatrix} + d \tag{5.4}$$

5

and the data generation mechanism is described by

$$\dot{x}^j = \mathsf{A}z^j + \mathsf{B}v^j + d^j, j = 0,1,\ldots,T-1. \tag{5.5}$$

with

$$\begin{aligned} \dot{x}^j &:= \dot{x}\left(j\tau_\text{s}\right), \quad z^j := Z\left(x\left(j\tau_\text{s}\right)\right),\\ v^j &:= W\left(x\left(j\tau_\text{s}\right)\right)u\left(j\tau_\text{s}\right),\\ d^j &:= d\left(j\tau_\text{s}\right). \end{aligned} \tag{5.6}$$

5.2. ROBUST INVARIANCE

First we need to clarify what kind of disturbance signal we consider in this chapter. So, we define the set

$$\mathscr{N} := \{d : \mathbb{R}_{\geq 0} \to \mathbb{R}^n : d \text{ is continuous}, d(t) \in \mathscr{D}_i \text{ for all } t \geq 0\} \tag{5.7}$$

which is the set of functions for the disturbance signals. In Chapter 4 we considered this kind disturbances only in the measured data used for the design. In this chapter we consider the presence of a disturbance also during operation. All disturbances belong to the set $\mathscr{N}$.

Definition 8 (Robust invariant set) *For $a\colon \mathbb{R}^n \to \mathbb{R}^n$ polynomial, i.e., $a \in \Pi_{n,1}$, for any disturbance $d \in \mathscr{N}$, and for an arbitrary $x_0 \in \mathbb{R}^n$, denote $t \mapsto \alpha(t, x_0, d(t))$ the unique solution to $\dot{x} = a(x) + d$ with initial condition $x_0 = \alpha(0, x_0, d(0))$ and defined on the interval $[0, T(x_0))$ (with $T(x_0)$ possibly $+\infty$). A set $\mathscr{I}$ is* robustly invariant *for $\dot{x} = a(x) + d$ if $x_0 \in \mathscr{I}$ and $d \in \mathscr{N}$ implies that for each $t \in [0, T(x_0))$, $\alpha(t, x_0, d(t)) \in \mathscr{I}$.*

As in Chapter 4 the invariant set is defined as

$$\mathscr{I} := \{x \in \mathbb{R}^n : h(x) \leq 0\} \tag{5.8}$$

with $h \in \Pi$. To impose that $\mathscr{I}$ is a robust invariant set for (5.4) as in Definition 8, we require that for each $d \in \mathscr{D}_i$ as in (5.2)

$$\{x \in \mathbb{R}^n, h(x) = 0\} \subseteq \{x \in \mathbb{R}^n, d \in \frac{\partial h}{\partial x}(x)\dot{x} \leq -\epsilon\} \tag{5.9}$$

where $\dot{x}$, used for brevity, takes the expression in (5.4) and the parameter $\epsilon > 0$ is introduced to guarantee some degree of robustness at the boundary of $\mathscr{I}$.

Lemma 3 (Differentiability, in Section 3.1 [44]) *Consider a nonlinear system $\dot{x} = f(x,t)$. If $f(x,t)$ is continuous in t and x, then the solution $t \mapsto \phi(x_0, t)$ will be continuously differentiable.*

5

Lemma 4 *Let h be a polynomial. Consider the set* (5.8) *and a generic system with additive disturbance $\dot{x} = a(x) + d$, with $d \in \mathscr{N}$. If*

$$\left\{x \in \mathbb{R}^n,\ d \in \mathscr{D}_i : h(x) = 0\right\} \subseteq \left\{x \in \mathbb{R}^n,\ d \in \mathscr{D}_i : \frac{\partial h}{\partial x}(x)\dot{x} \leq -\epsilon\right\} \tag{5.10}$$

then, the set $\mathscr{I}$ is robustly invariant (according to Definition 8).

Proof. Suppose by contradiction that Definition 8 does not hold, i.e., there is some $x_0 \in \mathscr{I}$ and $d \in \mathscr{N}$ for which it does not hold that $\alpha(t, x_0, d(t)) \in \mathscr{I}$, $\forall t \in [0, T(x_0))$, i.e., there is some $x_0 \in \mathscr{I}$, some $d(t) \in \mathscr{D}_i$, some $\bar{t} \in [0, T(x_0))$ and some $\bar{s} \in [0, T(x_0))$ such that

$$\bar{s} > \bar{t}, \alpha\left(\bar{t}, x_0, d(\bar{t})\right) \in \mathscr{I}, \alpha(t, x_0, d(t)) \notin \mathscr{I} \text{ for all } t \in (\bar{t}, \bar{s}]. \tag{5.11}$$

Consider then the function given by $t \mapsto g(t) := h(\alpha(t, x_0, d(t)))$. g is continuously differentiable because it is a composition of two continuously differentiable functions (h is a polynomial in x and α is continuously differentiable in t by Lemma 3). By (5.11), $g(\bar{t}) = h\left(\alpha\left(\bar{t}, x_0, d(\bar{t})\right)\right) = 0$ and $g(t) = h(\alpha(t, x_0, d(t))) > 0$ for all $t \in (\bar{t}, \bar{s}]$.

$$\dot{g}(\bar{t}) := \lim_{\varepsilon \to 0^+} \frac{g(\bar{t}+\varepsilon) - g(\bar{t})}{\varepsilon} = \lim_{\varepsilon \to 0^+} \frac{g(\bar{t}+\varepsilon)}{\varepsilon} \geq 0$$

because of $g(\bar{t}) = h\left(\alpha\left(\bar{t}, x_0, d(\bar{t})\right)\right) = 0$ and $g(t) = h(\alpha(t, x_0, d(t))) > 0$ for all $t \in (\bar{t}, \bar{s}]$. Consider the point $\bar{x} := \alpha\left(\bar{t}, x_0, d(\bar{t})\right)$. $\bar{x}$ satisfies

$$h(\bar{x}) = h\left(\alpha\left(\bar{t}, x_0, d(\bar{t})\right)\right) = g(\bar{t}) = 0 \tag{5.12}$$

and

$$\begin{aligned}\frac{\partial h}{\partial x}(\bar{x})(a(\bar{x})+d(\bar{t})) &= \frac{\partial h}{\partial x}\left(\alpha\left(\bar{t},x_0,d(\bar{t})\right)\right)(a(\bar{x})+d(\bar{t})) \\ &= \frac{\partial h}{\partial x}\left(\alpha\left(\bar{t},x_0,d(\bar{t})\right)\right)\dot{\alpha}\left(\bar{t},x_0,d(\bar{t})\right) = \dot{g}(\bar{t}) \geq 0\end{aligned} \tag{5.13}$$

The existence of $\bar{x}$ satisfying both (5.12) and (5.13) violates (5.10). □

By using Fact 1, the condition (5.9) is implied by

$$\ell(x)h(x)+\epsilon+\frac{\partial h}{\partial x}(x)(\mathsf{A}Z(x)+\mathsf{B}W(x)K(x)+d) \leq 0 \tag{5.14}$$

for all $x \in \mathbb{R}^n$ and $d \in \mathscr{D}_i$.

Now that we have defined and proved when a set is robustly invariant for system (5.4), we can present the main result of this chapter.

5

Theorem 9 ***(Data-driven robust invariance condition)***
For a design parameter $\epsilon > 0$, consider the data generation mechanism (5.5)*, measured data $\{\dot{x}^j, z^j, v^j\}_{j=0}^{T-1}$ as in* (5.6) *and a disturbance $d \in \mathcal{N}$ with $\mathcal{N}$ defined as in* (5.7) *and $\omega > 0$.*
Assume that the optimization program in (4.23) *is feasible so that parameters $\overline{\zeta}_{\mathrm{i}}$, $\overline{P}_{\mathrm{i}}$ and $\overline{Q}_{\mathrm{i}}$ in* (4.20) *exist. Assume that there exist decision variables $\ell \in \Pi$, $\eta \in \Pi$, $\tilde{\varrho} \in \Pi$, $h \in \Pi$ and $K \in \Pi_{m,1}$ such that for all $x \in \mathbb{R}^n$, $\eta(x) > 0$, $\tilde{\varrho}(x) > 0$ and $H(x) \preceq 0$, with H defined as*

$$H(x) := \begin{bmatrix} \left\{\begin{matrix} \ell(x)h(x)+\epsilon+\tilde{\varrho}(x) \\ +\frac{\partial h}{\partial x}(x)\overline{\zeta}_{\mathrm{i}}^{\top}\begin{bmatrix} Z(x) \\ W(x)K(x)\end{bmatrix}\end{matrix}\right\} & \star & \star & \star \\ \eta(x)\overline{P}_{\mathrm{i}}\begin{bmatrix} Z(x) \\ W(x)K(x)\end{bmatrix} & -2\eta(x)I & \star & \star \\ \overline{Q}_{\mathrm{i}}^{1/2}\frac{\partial h}{\partial x}(x)^{\top} & 0 & -2\eta(x)I & \star \\ \frac{1}{2}\sqrt{\omega}\frac{\partial h}{\partial x}(x)^{\top} & 0 & 0 & -\tilde{\varrho}(x)I \end{bmatrix}. \tag{5.15}$$

Then, the set $\mathscr{I}$ in (5.8) *is robustly invariant for the system in* (5.4).

Proof. The first part of the proof follows the same exact steps as the proof of Theorem 7. Under the assumption that (4.23) is feasible, $A_{\mathrm{i}} \succ 0$ by construction, so $\overline{P}_{\mathrm{i}}$ and $\overline{Q}_{\mathrm{i}}$ in (4.20) satisfy $\overline{P}_{\mathrm{i}} \succ 0$ and $\overline{Q}_{\mathrm{i}} \succ 0$. These two relations allow us to rewrite (4.21) as

$$\overline{\mathscr{C}}_{\mathrm{i}} = \{\zeta^{\top} \in \mathbb{R}^{n\times(n+m)} \colon \overline{Q}_{\mathrm{i}}^{-1/2}(\zeta-\overline{\zeta}_{\mathrm{i}})^{\top}\overline{P}_{\mathrm{i}}^{-1}\cdot[\star]^{\top} \preceq I\} \tag{5.16}$$

$$= \{(\overline{\zeta}_{\mathrm{i}}+\overline{P}_{\mathrm{i}}Y\overline{Q}_{\mathrm{i}}^{1/2})^{\top} \colon Y \in \mathbb{R}^{(n+m)\times n}, Y^{\top}Y \preceq I\} \tag{5.17}$$

where (5.16) is obtained from (4.21) since $\overline{Q}_{\mathrm{i}}^{-1/2} \succ 0$ and (5.17) is obtained by setting $Y = \overline{P}_{\mathrm{i}}^{-1}(\zeta - \overline{\zeta}_{\mathrm{i}})\overline{Q}_{\mathrm{i}}^{-1/2}$ in (5.16). Since the disturbance $d \in \mathcal{N}$, $(\mathsf{A}, \mathsf{B}) \in \bigcap_{j=0}^{T-1} \mathscr{C}_{\mathrm{i}}^{j} = \mathscr{C}_{\mathrm{i}}$ and, thus, $(\mathsf{A}, \mathsf{B}) \in \overline{\mathscr{C}}_{\mathrm{i}}$ in (5.17) since (4.23) enforces $\mathscr{C}_{\mathrm{i}} \subseteq \overline{\mathscr{C}}_{\mathrm{i}}$ by construction. The invariance condition in (5.14) imposed for all matrices in $\overline{\mathscr{C}}_{\mathrm{i}}$ reads

$$
\begin{gathered}
\forall [A\ B] \in \overline{\mathscr{C}}_{\mathrm{i}}, \forall x \in \mathbb{R}^n, \forall d \in \mathscr{D}_i, \\
\begin{cases} \ell(x)h(x) + \frac{\partial h}{\partial x}(x)\dot{x} + \epsilon \leq 0 \\ \dot{x} = AZ(x) + BW(x)K(x) + d. \end{cases}
\end{gathered} \tag{5.18}
$$

Since $[\mathsf{A}\ \mathsf{B}] \in \overline{\mathscr{C}}_{\mathrm{i}}$, if condition (5.18) holds, then the set $\mathscr{I}$ in (5.8) is robustly invariant for the ground truth system in (5.4). Therefore, the proof is complete if we verify (5.18), i.e., if we verify that

$$
\begin{aligned}
&\forall [A\ B] \in \overline{\mathscr{C}}_{\mathrm{i}}, \forall x \in \mathbb{R}^n, \forall d \in \mathscr{D}_i \\
&\ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)\,[A\ B] \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{\partial h}{\partial x}(x)d \leq 0 \\
&= \ell(x)h(x) + \epsilon + \tfrac{1}{\sqrt{2}} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^{\top} \begin{bmatrix} A^{\top} \\ B^{\top} \end{bmatrix} \tfrac{1}{\sqrt{2}} \tfrac{\partial h}{\partial x}(x)^{\top} \\
&\quad + \tfrac{1}{\sqrt{2}} \tfrac{\partial h}{\partial x}(x)\,[A\ B]\, \tfrac{1}{\sqrt{2}} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{\partial h}{\partial x}(x)d \leq 0.
\end{aligned}
$$

Rewrite this invariance condition in the compact form

$$
\begin{aligned}
&\forall \zeta^{\top} \in \overline{\mathscr{C}}_{\mathrm{i}}, \forall x \in \mathbb{R}^n, \forall d \in \mathscr{D}_i \\
&W(x,d) + S(x)\zeta R(x) + R(x)^{\top}\zeta^{\top}S(x)^{\top} \leq 0
\end{aligned} \tag{5.19}
$$

after defining

$$
\begin{aligned}
&W(x,d) := \ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)d \\
&S(x) := \tfrac{1}{\sqrt{2}} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^{\top},\ R(x) := \tfrac{1}{\sqrt{2}} \tfrac{\partial h}{\partial x}(x)^{\top}.
\end{aligned}
$$

(5.19), and thus (5.18), is equivalent, by (5.17), to

$$
\begin{aligned}
&\forall Y \text{ with } Y^{\top}Y \preceq I, \forall x \in \mathbb{R}^n,\ \forall d \in \mathscr{D}_i \\
&W(x,d) + S(x)(\overline{\zeta}_{\mathrm{i}} + \overline{P}_{\mathrm{i}}Y\overline{Q}_{\mathrm{i}}^{1/2})R(x) \\
&\ + R(x)^{\top}(\overline{\zeta}_{\mathrm{i}} + \overline{P}_{\mathrm{i}}Y\overline{Q}_{\mathrm{i}}^{1/2})^{\top}S(x)^{\top} \\
&= W(x,d) + S(x)\overline{\zeta}_{\mathrm{i}}R(x) + R(x)^{\top}\overline{\zeta}_{\mathrm{i}}^{\top}S(x)^{\top} \\
&\ + S(x)\overline{P}_{\mathrm{i}}Y\overline{Q}_{\mathrm{i}}^{1/2}R(x) + R(x)^{\top}\overline{Q}_{\mathrm{i}}^{1/2}Y^{\top}\overline{P}_{\mathrm{i}}S(x)^{\top} \leq 0.
\end{aligned} \tag{5.20}
$$

5

If there exist $\eta \in \Pi$ with $\eta(x) > 0$ for all $x \in \mathbb{R}^n$, we have

$$\begin{aligned}
&\forall Y \text{ with } Y^\top Y \preceq I, \forall x \in \mathbb{R}^n \\
&S(x)\overline{P}_i Y \overline{Q}_i^{1/2} R(x) + R(x)^\top \overline{Q}_i^{1/2} Y^\top \overline{P}_i S(x)^\top \\
&\preceq \eta(x) S(x)\overline{P}_i \overline{P}_i S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_i^{1/2} Y^\top Y \overline{Q}_i^{1/2} R(x) \\
&\preceq \eta(x) S(x)\overline{P}_i^2 S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_i R(x)
\end{aligned}$$

where we use Young's inequality in the first upper bound and $Y^\top Y \preceq I$ in the second upper bound. Using this last upperbound in (5.20), we obtain that (5.20) is implied by the existence of η, with $\eta(x) > 0$ for all $x \in \mathbb{R}^n$, such that

$$\begin{aligned}
\forall x \in \mathbb{R}^n,\ \forall d \in \mathscr{D}_i,\ & W(x,d) + S(x)\overline{\zeta}_i R(x) + R(x)^\top \overline{\zeta}_i^\top S(x)^\top \\
&+ \eta(x) S(x)\overline{P}_i^2 S(x)^\top + \tfrac{1}{\eta(x)} R(x)^\top \overline{Q}_i R(x) \preceq 0.
\end{aligned}$$

Replacing the explicit expressions of $W(x,d), S(x), R(x)$ in this inequality, we conclude that the invariance condition (5.18) holds if for all $x \in \mathbb{R}^n$, $\forall d \in \mathscr{D}_i$, $\eta(x) > 0$ and

$$\begin{aligned}
0 \geq\ & \ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)d + \tfrac{\partial h}{\partial x}(x)\overline{\zeta}_i^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} \\
&+ \tfrac{\eta(x)}{2} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top \overline{P}_i^2 \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{1}{2\eta(x)} \tfrac{\partial h}{\partial x}(x)\overline{Q}_i \tfrac{\partial h}{\partial x}(x)^\top.
\end{aligned}$$

As last step we need to remove the unknown disturbance d from

$$\begin{aligned}
0 \geq\ & \ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)\overline{\zeta}_i^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{\eta(x)}{2} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top \overline{P}_i^2 \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \\
&+ \tfrac{1}{2\eta(x)} \tfrac{\partial h}{\partial x}(x)\overline{Q}_i \tfrac{\partial h}{\partial x}(x)^\top + \frac{1}{2}\tfrac{\partial h}{\partial x}(x)d + d^\top \frac{1}{2}\tfrac{\partial h}{\partial x}(x)^\top.
\end{aligned}$$

First, we use Young's inequality on the cross product $\frac{1}{2}\frac{\partial h}{\partial x}(x)d + d^\top \frac{1}{2}\frac{\partial h}{\partial x}(x)^\top$ to obtain

$$\frac{1}{2}\tfrac{\partial h}{\partial x}(x)d + d^\top \frac{1}{2}\tfrac{\partial h}{\partial x}(x)^\top \leq \varrho(x)\frac{1}{4}\tfrac{\partial h}{\partial x}(x)dd^\top \tfrac{\partial h}{\partial x}(x)^\top + \frac{1}{\varrho(x)}$$

with $\varrho(x) > 0$ for all $x \in \mathbb{R}^n$. Lastly recall that we assumed that d is bounded as (5.2). The bound can be equivalently written as $dd^\top \preceq \omega I$ so we have

$$\varrho(x)\frac{1}{4}\tfrac{\partial h}{\partial x}(x)dd^\top \tfrac{\partial h}{\partial x}(x)^\top + \frac{1}{\varrho(x)} \leq \varrho(x)\omega\frac{1}{4}\tfrac{\partial h}{\partial x}(x)\tfrac{\partial h}{\partial x}(x)^\top + \frac{1}{\varrho(x)}$$

Putting this all together leads to

$$\begin{aligned}
0 \geq\ & \ell(x)h(x) + \epsilon + \tfrac{\partial h}{\partial x}(x)\overline{\zeta}_i^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \tfrac{\eta(x)}{2} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix}^\top \overline{P}_i^2 \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} + \\
&+ \tfrac{1}{2\eta(x)} \tfrac{\partial h}{\partial x}(x)\overline{Q}_i \tfrac{\partial h}{\partial x}(x)^\top + \varrho(x)\omega\frac{1}{4}\tfrac{\partial h}{\partial x}(x)\tfrac{\partial h}{\partial x}(x)^\top + \frac{1}{\varrho(x)}.
\end{aligned}$$

The last inequality is equivalent to

$$H(x) := \begin{bmatrix} \left\{ \begin{matrix} \ell(x)h(x)+\epsilon+\frac{1}{\varrho(x)} \\ +\frac{\partial h}{\partial x}(x)\overline{\zeta}_{\rm i}^\top \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} \end{matrix} \right\} & \star & \star & \star \\ \overline{P}_{\rm i} \begin{bmatrix} Z(x) \\ W(x)K(x) \end{bmatrix} & -\frac{2}{\eta(x)}I & \star & \star \\ \overline{Q}_{\rm i}^{1/2} \frac{\partial h}{\partial x}(x)^\top & 0 & -2\eta(x)I & \star \\ \frac{1}{2}\sqrt{\omega}\frac{\partial h}{\partial x}(x)^\top & 0 & 0 & -\frac{1}{\varrho(x)}I \end{bmatrix} \preceq 0. \tag{5.21}$$

by applying Schur complement [51, p. 28].

Finally to obtain (5.15) we define $\tilde{\varrho}(x) := \frac{1}{\varrho(x)}$ and we pre- and post-multiply $H(x)$ as follows

$$\begin{bmatrix} I & & & \\ & \eta I & & \\ & & I & \\ & & & I \end{bmatrix} H(x) \begin{bmatrix} I & & & \\ & \eta I & & \\ & & I & \\ & & & I \end{bmatrix} \preceq 0. \tag{5.22}$$

■

The robust invariance condition obtained in Theorem 9 is a key component to design safe robust control laws in applications. As in Chapter 4 we introduce the so-called safe set $\mathscr{S}$, by which all constraints on the state can be expressed by positivity of polynomials $\sigma_1, \ldots, \sigma_q$ ($q \in \mathbb{Z}_{\geq 1}$) and the safe set $\mathscr{S}$ is

$$\mathscr{S} := \{x \in \mathbb{R}^n : \sigma_j(x) \leq 0, j = 1, \ldots, q\}. \tag{5.23}$$

To design a safe robust controller, we need to enforce besides the invariance condition (5.9) a safety condition $\mathscr{I} \subseteq \mathscr{S}$ so that when the state belongs to $\mathscr{I}$, it will also comply with all constraints expressed by $\mathscr{S}$. To obtain a robust invariant set that is as "large" as possible, we impose also $\mathscr{I} \subseteq \mathscr{S}$ as in Section 4.5. We recall here the definition of the variable-size set $\mathscr{L}_\theta$ used to enlarge $\mathscr{I}$

$$\mathscr{L}_\theta := \{x \in \mathbb{R}^n : \lambda(x) \leq \theta\}. \tag{5.24}$$

where $\theta \geq 0$ is used to enlarge the $\mathscr{L}_\theta$. The safe condition and the variable-size set inclusion are relaxed in sum-of-squares conditions that are computationally tractable.

Theorem 10 ***(Data-driven robust safe control)***
For given polynomials $\sigma_1, \ldots, \sigma_q$ and nonnegative polynomial λ, consider the set $\mathscr{S}$ in (5.23) *and the set $\mathscr{L}_\theta$ in* (5.24) *parametrized by θ, along with H defined*

in (5.15) *and a given* $\bar{\eta} > 0$.
Under the same hypothesis of Theorem 9, assume the program

$$\begin{aligned}
\textit{maximize} \quad & \theta && (5.25a)\\
\textit{s.t.} \quad & \{\ell, \eta, h\} \subseteq \Pi, K \in \Pi_{m,1}, \{s_1, \dots, s_q, \varsigma\} \subseteq \Sigma, \theta \geq 0 && (5.25b)\\
& \eta - \bar{\eta} \in \Sigma, -H \in \Sigma_r, \tilde{\varrho} \in \Sigma && (5.25c)\\
& \varsigma(\lambda - \theta) - h \in \Sigma, && (5.25d)\\
& s_j h - \sigma_j \in \Sigma, j = 1, \dots, q. && (5.25e)
\end{aligned}$$

has a solution. Then, the set $\mathscr{I}$ *in* (5.8) *is robustly invariant for the system in* (5.4) *and satisfies* $\mathscr{L}_\theta \subseteq \mathscr{I} \subseteq \mathscr{S}$.

The proof follows the one of Theorem 8 and is therefore omitted.

5

5.3. NUMERICAL EXAMPLE: CAR PLATOONING

To prove the effectiveness of our design, in this section we show how to obtain a robust safe controller for a safety-critical system using only data. For comparison with the safe controller designed in Chapter 4, we consider the same system with two cars moving in a platoon formation with the addition of a disturbance. The system can be modeled as

$$\begin{aligned}
\dot{x}_1 &= u_1 - \gamma_1 - \beta_1 x_1 - \alpha_1 x_1^2 + \mu_1 && (5.26a)\\
\dot{x}_2 &= u_2 - \gamma_2 - \beta_2 x_2 - \alpha_2 x_2^2 + \mu_2 && (5.26b)\\
\dot{x}_3 &= x_1 - x_2 && (5.26c)
\end{aligned}$$

where: the components x_1, x_2, x_3 of state x represent respectively the velocity of the front vehicle, the velocity of the rear vehicle and the relative distance between the two; the components u_1, u_2 of input u represent the forces normalized by vehicle mass; γ_k, β_k, α_k for $k = 1,2$ are the static, rolling and aerodynamic-drag friction coefficients normalized by vehicle mass; finally to avoid confusion with the notation the disturbance components are not denoted with d but with μ_1, μ_2. We impose the safety constraints represented by the set

$$\begin{aligned}
\mathscr{S} := \{x \in \mathbb{R}^3 :\ & d_0 + \tau_h x_2 \leq x_3,\ x_3 \leq d_1,\\
& 0 \leq x_1,\ x_1 \leq v_M,\ 0 \leq x_2,\ x_2 \leq v_M\}\\
=: \{x \in \mathbb{R}^3 :\ & \sigma_1(x) \leq 0, \dots, \sigma_6(x) \leq 0\} && (5.27)
\end{aligned}$$

where $d_0 + \tau_h x_2$ is a relative distance required to avoid collisions with the front vehicle ($d_0 > 0$ is a standstill distance and $\tau_h > 0$ is a time headway), d_1 is a distance required to keep the benefits of platooning (especially, aerodynamic drag reduction), and v_{M} is a maximum velocity allowed on the road. As in Chapter 4, we consider as point of interest $\bar{x} := (\bar{v}, \bar{v}, \bar{d})$ where $\bar{v}$ is a predefined cruise velocity and $\bar{d}$ is a reference safety distance. We used the same numerical values reported in Chapter 4 in table 4.1 for the parameters.

A robust safe controller for the perturbed system is then obtained using Theorem 10, Algorithm 3 reports all the details of the implementation. Compared with Algorithm 2, the H matrix now is equal to (5.15) and an additional polynomial $\tilde{\varrho}$ is added. All the other parameters and initialization values used are equal to the ones in Algorithm 2: for the polynomials h, η, $\tilde{\varrho}$, s_j $(j = 1,\dots,6)$ and ς the degrees are set, respectively, to $4, 2, 2, 2, 2$ with the shaping function $\mathscr{L}_\theta$,

$$\mathscr{L}_\theta := \{x \in \mathbb{R}^3 : (x - \bar{x})^\top \begin{bmatrix} 0.02 & 0 & 0 \\ 0 & 0.05 & 0 \\ 0 & 0 & 1 \end{bmatrix} (x - \bar{x}) =: \lambda(x) \le \theta\}$$

and with the monomials $Z(x) = [x_1 \ x_2 \ x_3 \ x_1^2 \ x_2^2]^\top$ and $W(x) = I$.

Algorithm 3 Car Platooning (Robust)

initialize: $\iota = 0$ (a counter), $\theta_\iota = \theta_0 = 0.01$, $\epsilon = 0.01$, $\eta(x) = 1$. To initialize the polynomial $h(x)$ is it possible to proceed as in algorithm 2 using the Lyapunov function for the linearized system or $h(x)$ can be initialized using the result of algorithm 2.
repeat
    Find $\ell \in \Pi, \rho \in \Pi$, $K \in \Pi_{m,1}$ and $s_1,\dots,s_6, \varsigma \in \Sigma$
    subject to $\quad -H \in \Sigma_r$
        $\varsigma(\lambda - \theta_\iota) - h \in \Sigma$
        $s_j h - \sigma_j \in \Sigma, \quad j = 1,\dots,6.$
    Update ℓ, K, $s_1,\dots,s_6$, ς.

    Maximize θ $\quad$ over $\eta, h \in \Pi$, $\rho \in \Pi$, and $\theta \ge 0$,
    subject to $\quad \eta \in \Sigma$, $-H \in \Sigma_r$
        $\varsigma(\lambda - \theta) - h \in \Sigma$
        $s_j h - \sigma_j \in \Sigma, \quad j = 1,\dots,6.$
    Update η, h, $\theta_\iota = \theta$, $\iota \leftarrow \iota + 1$.
until $\theta_\iota - \theta_{\iota-1}$ is less than some tolerance (10^{-3}).

While algorithm 2 could work with noisy data, it does not provide any theoretical guarantees that due to the presence of the disturbance the closed-loop system can not violate the safety constraints. Algorithm 3 solves this problem by implementing Theorem 10, that it is based on the definition of robust invariance, and it can be used to improve upon the result of algorithm 2. Instead of initializing

$h(x)$ with a Lyapunov function for the linearized system we can use the result of algorithm 2 to initialize $h(x)$. This solution significantly reduces the number of iterations needed for algorithm 3 to provide a valid result.

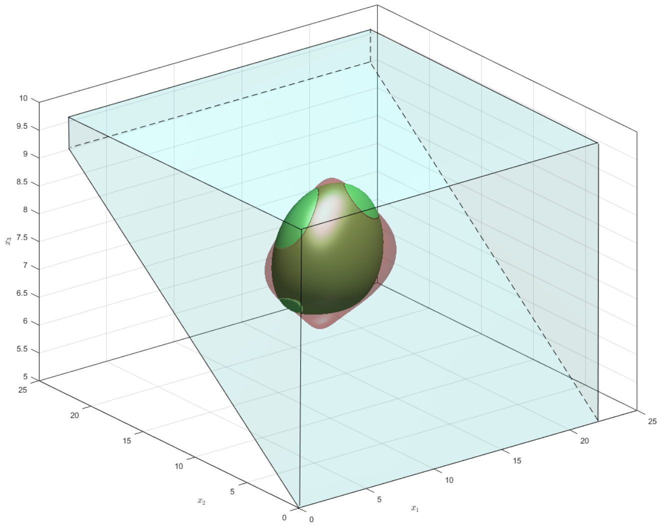

Figure 5.1 | The data-driven invariant set from Chapter 4 in light red and the robust invariance data-driven design in green ($|d| \leq 10^{-3}$, $T = 1000$). The safe set in light blue highlights that the two invariant sets comply with safety constraints.

By using algorithm 3 on the platoon of cars, we obtain a robust invariant set $\mathscr{I} = \{x \in \mathbb{R}^3 \colon h(x) \leq 0\}$ with

$$\begin{aligned} h(x) = {} & 2733.46 + 3664.65x_1^2 + 4290.05x_2^2 + 19653.25x_3^2 - 2333.36x_1x_2 - 5831.70x_1x_3 + \\ & - 13891.54x_2x_3 - 26.55x_1^3 - 71.02x_1^2x_2 - 802.53x_1^2x_3 - 178.46x_1x_2^2 - 15.55x_2^3 + \\ & - 937.89x_2^2x_3 + 1145.50x_1x_2x_3 + 1062.86x_1x_3^2 + 3089.22x_2x_3^2 - 4662.43x_3^3 + \\ & + 12.43x_1^4 - 10.93x_1^3x_2 + 40.22x_1^2x_2^2 - 28.81x_1x_2^3 + 24.13x_2^4 - 35.69x_1^3x_3 \\ & - 40.54x_1^2x_2x_3 + 33.49x_1x_2^2x_3 - 65.20x_2^3x_3 + 129.33x_1^2x_3^2 - 71.65x_1x_2x_3^2 + \\ & + 148.39x_2^2x_3^2 - 114.18x_1x_3^3 - 252.75x_2x_3^3 + 318.01x_3^4 \end{aligned} \tag{5.28}$$

and a controller $u = K(x) = \begin{bmatrix} K_1(x) \\ K_2(x) \end{bmatrix}$ with

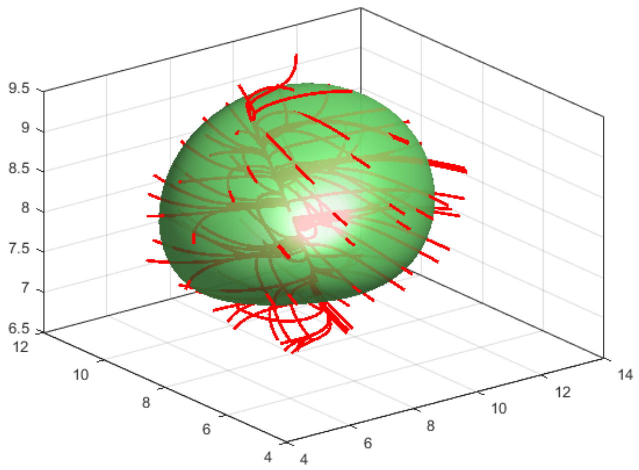

Figure 5.2 | To show that the invariant set $\mathscr{I}$ (green) given by (5.28) is invariant, we simulate the closed-loop system with initial conditions slightly outside $\mathscr{I}$, close to the boundary. All simulated trajectories (red) never leave the set after entering it.

$$\begin{aligned} K_1(x) &= 138.45 + 575.46x_1 - 74.51x_2 - 435.66x_3 + 99.60x_1^2 + 119.35x_2^2 + 233.01x_3^2 \\ &\quad - 162.64x_1x_2 - 201.18x_1x_3 - 82.53x_2x_3 \\ K_2(x) &= 233.81 + 523.61x_1 + 440.53x_2 - 843.22x_3 + 128.01x_1^2 + 122.09x_2^2 + 258.79x_3^2 \\ &\quad - 227.42x_1x_2 - 100.05x_1x_3 - 180.11x_2x_3. \end{aligned} \tag{5.29}$$

We compared the result given by algorithm 3 with the result for nominal invariance reported in Section 4.7. In Figure 5.1, we can see that the robust invariant set found by Algorithm 3 is not worse than the one found by algorithm 2, the data points are affected by a disturbance satisfying $|d| \leq 10^{-3}$. In both cases safety constraints are not violated since both invariant sets are within $\mathscr{S}$. In Figure 5.2, the invariant set $\mathscr{I}$ for (5.28) is plotted together with trajectories of the vector field (5.26) in closed-loop with the controller $K(x)$ in (5.29). Trajectories are initialized close to the boundary of $\mathscr{I}$ to show that once in the set $\mathscr{I}$, they never leave it, thereby confirming that $\mathscr{I}$ is invariant.

5.4. CONCLUSIONS

In this chapter, we investigated how the safe controller design presented in Chapter 4 can be extended to account for noise during operation. Using the concept of robust invariant set, we have formulated a semi-definite program to obtain a stabilizing controller that theoretically guarantees the satisfaction of all safety constraints. The problem was then relaxed in a more tractable form using sum-of-squares programming. The resulting problem has one additional polynomial decision variable $\varrho(x)$ and a larger matrix polynomial decision variable $H(x)$ when compared to the nominal (non-disturbed) case of Chapter 4.

6

CONCLUSION

In this thesis, we proposed new methods to design stabilizing controllers directly from data. The detailed conclusions are as follows.

- In Chapter 2, we solved the absolute stability problem for nonlinear systems that can be represented with a nonlinear component that satisfies quadratic constraints. It was shown that necessary and sufficient conditions for the design of a stabilizing controller can be obtained directly from input/state measurements without the need to know the model. These conditions can be verified directly from data by solving a unique semi-definite programming problem. We have also noted that our result can be viewed as a data-dependent Kalman-Yakubovitch-Popov Lemma. Finally, we have proved the effectiveness of our results with numerical examples.

- In Chapter 3, we have extended the design of a stabilizing controller for the nonlinear systems studied in Chapter 2 by considering the presence of noise in the measurements. Assuming a bounded disturbance, all the results obtained for the nominal case were reformulated with the inclusion of the disturbance demonstrating that our solution is suitable also for real applications.

- In Chapter 4, the problem of designing a safe controller for polynomial systems was discussed. The Positivstellensatz was used to formulate the problem in a tractable form by relaxing polynomial positivity conditions with sum-of-squares conditions. For this chapter, we considered the presence of the noise only on the measurement collected from open-loop experiments to estimate the set of compatible system matrices. The operation of the controller in closed-loop was assumed noiseless. The proposed solution was tested on a safety-critical system, a platoon of cars, and the resulting safe controller was compared with the one obtained with a model-based design with perfect knowledge. Using this numerical example, we have shown that our solution returns a controller with comparable performance to the one derived knowing the system model.

- In Chapter 5, we introduced the definition of robust invariance to provide theoretical guarantees on the robustness of the derived controller. Using the Positivstellensatz to relax the problem to a sum-of-squares condition, we showed how the algorithm derived in Chapter 4 can be modified to guarantee stability and safety despite the presence of noise during operation.

6.1. FUTURE RESEARCH

The results presented in this thesis are a contribution in the direction of replacing all model-based solutions for controller synthesis with direct data-driven algorithms that do not require any knowledge of the system to control.

We have presented in Chapters 2-3 a data-driven solution for the problem of absolute stability and we considered the disturbance only in the data collected from the open-loop system. A natural extension of these works would be to consider the noise during the execution of the control task and study the stability problem with respect to external perturbations. Handling noise in the data is one of the main concerns of any data-driven solution and improving the robustness to disturbance is critical to apply new data-driven algorithms in real applications.

In Chapters 4-5 we have presented an algorithm to derive an invariant set and a safe controller directly from data. The invariance property of the set was enforced only around one predefined state, and it would be interesting to extend our design to the case of tracking a prescribed trajectory without violating any constraint. As for the work on absolute stability one of the main areas where there is room for improvement is in the noise robustness. To improve the robustness new options must be explored since reducing the relaxations we used is not a solution. The SOS and P-Satz relaxations would be present also in a model based solution and $\overline{\mathscr{C}}$ is not a critical simplification. One possibility to increase noise robustness can be to use not raw input/state data but filtered data. Singular spectrum analysis (SSA) [92][93] is a promising solution to discard all the components that are due to the presence of noise. SSA is a technique based on singular value decomposition used to de-noise time series.

The research for completely data-driven control algorithms was also inspired by the success obtained by machine learning techniques in solving complex problems. So, for future research, it is worth to explore more the possibility of using machine learning algorithms to solve control problem from data. One of the main differences between classical machine learning algorithms and the solution we have presented in this thesis is the theoretical guarantees we were able to provided for our results. In the future it would be interesting to investigate how to provide the same level of theoretical guarantees, necessary especially for safety critical systems, with deep neural networks. A good overview of the work done on this subject can be found in [94].

SUMMARY

The recent successes of machine learning solutions have inspired the research of new control algorithms derived directly from the available data without any intermediate step. Being able to design a stabilizing controller directly from data has the main advantage that, since it does not rely on a model of the system to control, the controller design is not influenced by any modeling error.

Most of the time real systems are simplified with linear models to reduce the overall complexity in the controller design discarding all the complex non-linear behaviors. A linear approximation could be an excessive simplification for complex system where the presence of nonlinear dynamics are important to understand those processes and nonlinearities can not be ignored. However, the analysis and control of a nonlinear model is often challenging. Most of the time real systems are simplified with linear models to reduce the overall complexity in the controller design discarding all the complex non-linear behaviors. A linear approximation could be an excessive simplification for complex system where the presence of nonlinear dynamics are important to understand those processes and nonlinearities can not be ignored. However, the analysis and control of a nonlinear model is often challenging.

This thesis investigates data-based control methods for continuous and discrete-time nonlinear systems. In particular we have develop a solution to obtain a stabilizing state feedback controller for the case of nonlinear systems where the nonlinearities can be bounded with a quadratic constraint. This is a notable class of systems that includes: systems with sector bounded or passive nonlinearities and fully recurrent neural networks. As opposed to classic model based solutions our approach can be applied to both continuous and discrete systems without any significant change.

Stabilizing a closed-loop system is critical but sometimes it is not enough. Safety is another important criteria considered in the design of a controller. The inclusion of safety requirements in the design complicates the overall problem that now includes additional constraints. Safety constraints can be formulated as a safe set where the state of the system must never leave. To solve this design problem using only data, we have developed a new approach for the case of nonlinear systems that can be modeled as polynomial functions. Using sum-of-squares programming and the Positivstellensatz theorem we were able to formulate a tractable semi-definite problem to find a stabilizing controller that

can also guarantee that the state of the system never violate the safety constraints.

One of the major problem in the application of data-driven algorithms in real cases is the presence of noise in the measurements. For all the solutions presented we discuss how to handle noise to obtain stabilizing controllers also with real measurements.

SAMENVATTING

De recente successen van machine learning-oplossingen hebben geleid tot het onderzoek naar nieuwe besturingsalgoritmen die rechtstreeks zijn afgeleid van de beschikbare gegevens zonder enige tussenstap. Het kunnen ontwerpen van een stabiliserende controller rechtstreeks vanuit data heeft het grote voordeel dat, aangezien het niet afhankelijk is van een model van het te besturen systeem, het ontwerp van de controller niet wordt beinvloed door enige modelleringsfout.

Meestal worden echte systemen vereenvoudigd met lineaire modellen om de algehele complexiteit in het ontwerp van de controller te verminderen, waarbij alle complexe niet-lineaire gedragingen worden weggegooid. Een lineaire benadering zou een buitensporige vereenvoudiging kunnen zijn voor complexe systemen waar de aanwezigheid van niet-lineaire dynamiek belangrijk is om die processen te begrijpen en niet-lineariteiten niet kunnen worden genegeerd. De analyse en besturing van een niet-lineair model is echter vaak een uitdaging. Meestal worden echte systemen vereenvoudigd met lineaire modellen om de algehele complexiteit in het controllerontwerp te verminderen, waarbij alle complexe niet-lineaire gedragingen worden weggegooid. Een lineaire benadering zou een buitensporige vereenvoudiging kunnen zijn voor complexe systemen waar de aanwezigheid van niet-lineaire dynamiek belangrijk is om die processen te begrijpen en niet-lineariteiten niet kunnen worden genegeerd. De analyse en controle van een niet-lineair model is echter vaak een uitdaging.

Dit proefschrift onderzoekt op data gebaseerde regelmethoden voor continue en discrete tijd niet-lineaire systemen. In het bijzonder hebben we een oplossing ontwikkeld om een stabiliserende toestandsfeedbackcontroller te verkrijgen voor het geval van een niet-lineair systeem waar de niet-lineariteiten kunnen worden begrensd met een kwadratische beperking. Dit is een opmerkelijke klasse van systemen die omvat: systeem met sectorgebonden of passieve niet-lineariteiten en volledig terugkerende neurale netwerken. In tegenstelling tot klassieke model-gebaseerde oplossingen kan onze aanpak zonder noemenswaardige verandering worden toegepast op zowel continue als discrete systemen.

Het stabiliseren van een gesloten-lussysteem is van cruciaal belang, maar soms is het niet genoeg. Veiligheid is een ander belangrijk criterium bij het ontwerp van een controller. Het opnemen van veiligheidseisen in het ontwerp compliceert het algehele probleem dat nu extra beperkingen bevat. Veiligheids-beperkingen kunnen worden geformuleerd als een veilige set waar de toestand van het systeem nooit mag weggaan. Om dit ontwerpprobleem op te lossen met alleen data, hebben we een nieuwe benadering ontwikkeld voor het geval van niet-lineaire systemen die kunnen worden gemodelleerd als polynoomfuncties. Met behulp van kwadratische programmering en de Positivstellensatz-stelling konden we een hanteerbaar semi-definitief probleem formuleren om een stabiliserende controller te vinden die ook kan garanderen dat de toestand van het systeem nooit de veiligheidsbeperkingen schendt.

Een van de grootste problemen bij de toepassing van datagestuurde algoritmen in reële gevallen is de aanwezigheid van ruis in de metingen. Voor alle gepresenteerde oplossingen bespreken we hoe we met ruis kunnen omgaan om stabiliserende controllers te verkrijgen, ook met echte metingen.

Acknowledgements

This thesis is the culmination of a four years work, and I would like to thank all the peoples that help me to arrive at this milestone. First, I would like to thank my supervisors Claudio, Pietro and Bart for their guidance. Claudio thank you for all your help and patience with me, your help was essential to complete my PhD. I would like to express my greatest gratitude to Andrea, I could not complete this thesis with your support, thank you. It was a pleasure to work with you, and I learned a lot from your corrections and detailed suggestions. Many thanks to my reading committee members for taking the time to read and comment this thesis. Finally, I would like to thank all my friend and my family for their support during the years of my PhD.

Bibliography

References

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

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

[3] J. Coulson, J. Lygeros, and F. Dörfler. Data-enabled predictive control: In the shallows of the deepc. In *2019 18th European Control Conference (ECC)*, pages 307–312. IEEE, 2019.

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

[5] T. Holicki, C. W. Scherer, and S. Trimpe. Controller design via experimental exploration with robustness guarantees. *IEEE Control Systems Letters*, 5(2):641–646, 2020.

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

[7] Claudio De Persis and Pietro Tesi. Low-complexity learning of linear quadratic regulators from noisy data. *Automatica*, 128:109548, 2021.

[8] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi. From noisy data to feedback controllers: non-conservative design via a matrix s-lemma. *IEEE Transactions on Automatic Control*, 2020.

[9] J. Berberich, C. W. Scherer, and F. Allgöwer. Combining prior knowledge and data for robust controller design. *arXiv preprint arXiv:2009.05253*, 2020.

[10] A. Romer, J. M. Montenbruck, and F. Allgöwer. Determining dissipation inequalities from input-output samples. *IFAC-PapersOnLine*, 50(1):7789–7794, 2017.

[11] Tábitha E Rosa and Bayu Jayawardhana. On the one-shot data-driven verification of dissipativity of lti systems with general quadratic supply rate function. In *2021 European Control Conference (ECC)*, pages 1291–1296. IEEE, 2021.

[12] A. Bisoffi, C. De Persis, and P. Tesi. Data-based stabilization of unknown bilinear systems with guaranteed basin of attraction. *Sys. & Contr. Lett.*, 145:104788, 2020.

[13] M. Sharf. On the sample complexity of data-driven inference of the l 2-gain. *IEEE Control Systems Letters*, 4(4):904–909, 2020.

[14] C. Novara, S. Formentin, S. M. Savaresi, and M. Milanese. Data-driven design of two degree-of-freedom nonlinear controllers: the d2-ibc approach. *Automatica*, 72:19–27, 2016.

[15] Eurika Kaiser, J Nathan Kutz, and Steven L Brunton. Data-driven discovery of Koopman eigenfunctions for control. *Machine Learning: Science and Technology*, 2(3):035023, 2021.

[16] B. Lusch, J. N. Kutz, and S. L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature communications*, 9(1):1–10, 2018.

[17] Bahadir Saltik, Bayu Jayawardhana, and Ashish Cherukuri. Koopman operator based iterative learning and model predictive control for repetitive nonlinear systems. In *61st IEEE Conference on Decision and Control*. IEEE, 2022.

[18] Z. Yuan and J. Cortés. Data-driven optimal control of bilinear systems. *arXiv preprint arXiv:2112.15510*, 2021.

[19] M. Guo, C. De Persis, and P. Tesi. Data-driven stabilization of nonlinear polynomial systems with noisy data. *IEEE Transactions on Automatic Control*, pages 1–1, 2021.

[20] T. Dai and M. Sznaier. A semi-algebraic optimization approach to data-driven control of continuous-time nonlinear systems. *IEEE Control Systems Letters*, 5(2):487–492, 2020.

[21] J. G. Rueda-Escobedo and J. Schiffer. Data-driven internal model control of second-order discrete Volterra systems. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 4572–4579, 2020.

[22] Y. Guo, Z. Hou, S. Liu, and S. Jin. Data-driven model-free adaptive predictive control for a class of mimo nonlinear discrete-time systems with stability analysis. *IEEE Access*, 7:102852–102866, 2019.

[23] H. J. Van Waarde and M. K. Camlibel. A matrix finsler's lemma with applications to data-driven control. *arXiv preprint arXiv:2103.13461*, 2021.

[24] C. De Persis, M. Rotulo, and P. Tesi. Learning controllers from data via approximate nonlinearity cancellation. *arXiv:2201.10232*, 2020.

[25] E. Garone, S. Di Cairano, and I. Kolmanovsky. Reference and command governors for systems with constraints: A survey on theory and applications. *Automatica*, 75:306–328, 2017.

[26] J. Wolff and M. Buss. Invariance control design for constrained nonlinear systems. In *IFAC Proc. Vol.*, 2005.

[27] K. P. Wabersich and M. N. Zeilinger. Scalable synthesis of safety certificates from data with application to learning-based control. In *Proc. Eur. Contr. Conf.*, 2018.

[28] W. Shaw Cortez and D. V. Dimarogonas. Safe-by-design control for Euler-Lagrange systems. arXiv preprint arXiv:2009.03767, 2021.

[29] M. Nagumo. Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen. *Proc. Physico-Mathem. Soc. of Japan, 3rd Ser.*, 1942.

[30] J. P. Aubin, A. M. Bayen, and P. Saint-Pierre. *Viability theory: new directions*. Springer, 2011.

[31] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[32] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames. Robustness of control barrier functions for safety critical control. In *IFAC-PapersOnLine*, 2015.

[33] X. Xu, J. W. Grizzle, P. Tabuada, and A. D. Ames. Correctness guarantees for the composition of lane keeping and adaptive cruise control. *IEEE Trans. Autom. Science Eng.*, 15(3):1216–1229, 2017.

[34] P. Wieland and F. Allgöwer. Constructive safety using control barrier functions. *IFAC Proc. Vol.*, 2007.

[35] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *Proc. Int. Worksh. Hybr. Sys. Comp. Contr.*, 2004.

[36] M. Jankovic. Robust control barrier functions for constrained stabilization of nonlinear systems. *Automatica*, 96:359–367, 2018.

[37] K. Garg and D. Panagou. Robust control barrier and control Lyapunov functions with fixed-time convergence guarantees. In *Proc. Amer. Contr. Conf.*, 2021.

[38] B. T. Lopez, J.J.E. Slotine, and J. P. How. Robust adaptive control barrier functions: An adaptive and data-driven approach to safety. *IEEE Contr. Sys. Lett.*, 5(3):1031–1036, 2020.

[39] A. Alan, A.J. Taylor, C. R. He, G. Orosz, and A. D. Ames. Safe controller synthesis with tunable input-to-state safe control barrier functions. *IEEE Contr. Sys. Lett.*, 2021.

[40] A.I. Lurie. Some nonlinear problems in the theory of automatic control. *London: HM Stationary Office*, 1957.

[41] V.M. Popov. Relaxing the sufficiency conditions for absolute stability. *Automation and Remote Control*, 19(1):1–7, 1958.

[42] V. A. Yakubovich. Solution of certain matrix inequalities in the stability theory of nonlinear control systems. *Soviet. Math. Dokl.*, 3:620–623, 1962.

[43] A.L. Fradkov, I.V. Miroshnik, and V.O. Nikiforov. *Nonlinear and adaptive control of complex systems*, volume 491. Springer Science & Business Media, 2013.

[44] H.K. Khalil. *Nonlinear systems*. Prentice Hall Upper Saddle River, NJ, 2002.

[45] D.D. Šiljak and D. M. Stipanovic. Robust stabilization of nonlinear systems: The LMI approach. *Mathematical problems in Engineering*, 6(5):461–493, 2000.

[46] W.M. Haddad and D.S. Bernstein. Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability. part ii: Discrete-time theory. *International Journal of Robust and Nonlinear Control*, 4(2):249–265, 1994.

[47] A. Zemouche and M. Boutayeb. On LMI conditions to design observers for Lipschitz nonlinear systems. *Automatica*, 49(2):585–591, 2013.

[48] M. Jin and J. Lavaei. Control-theoretic analysis of smoothness for stability-certified reinforcement learning. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 6840–6847, 2018.

[49] L. Lessard, B. Recht, and A. Packard. Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM Journal on Optimization*, 26(1):57–95, 2016.

[50] N. E. Barabanov and D. V. Prokhorov. Stability analysis of discrete-time recurrent neural networks. *IEEE Transactions on Neural Networks*, 13(2):292–303, 2002.

[51] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear matrix inequalities in system and control theory*. Siam, 1994.

[52] C. De Persis and P. Tesi. Designing experiments for data-driven control of nonlinear systems. *24th International Symposium on Mathematical Theory of Networks and Systems (Cambridge, UK), invited paper on "Data-driven control", Markowsky, Mishra, Zeilinger (Organizers)*, 2020.

[53] V. A. Yakubovich, G. A. Leonov, and A.Kh. Gelig. *Stability of stationary sets in control systems with discontinuous nonlinearities*, volume 14. World Scientific, 2004.

[54] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. cvxr.com/cvx, 2014.

[55] H. J. Van Waarde, J. Eising, H. Trentelman, and M. K. Camlibel. Data informativity: a new perspective on data-driven analysis and control. *IEEE Transactions on Automatic Control*, 65(11):4753–4768, 2020.

[56] A. Rantzer. On the Kalman-Yakubovich-Popov lemma for positive systems. *IEEE Transactions on Automatic Control*, 61(5):1346–1349, 2015.

[57] M. Arcak, M. Larsen, and P. Kokotović. Circle and Popov criteria as tools for nonlinear feedback design. *IFAC Proceedings Volumes*, 35(1):85–90, 2002.

[58] M. Schlegel and B. R. Noack. On long-term boundedness of Galerkin models. *Journal of Fluid Mechanics*, 765:325–352, 2015.

[59] A. A. Kaptanoglu, J. L. Callaham, A. Aravkin, C. J. Hansen, and S. L. Brunton. Promoting global stability in data-driven models of quadratic nonlinear dynamics. *Physical Review Fluids*, 6(9):094401, 2021.

[60] S.S. Sastry and A. Isidori. Adaptive control of linearizable systems. *IEEE Transactions on Automatic Control*, 34(11):1123–1131, 1989.

[61] S.L. Brunton, J.L. Proctor, J.N. Kutz, and W. Bialek. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 113(15):3932–3937, 2016.

[62] E. D. Sontag. Further facts about input to state stabilization. *IEEE Transactions on Automatic Control*, 35(4):473–476, 1990.

[63] Fred Schweppe. Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28, 1968.

[64] Dimetri Bertsekas and Ian Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Transactions on Automatic Control*, 16(2):117–128, 1971.

[65] A. Bisoffi, C. De Persis, and P. Tesi. Controller design for robust invariance from noisy data. arXiv preprint available at `https://arxiv.org/abs/2007.13181`, 2020.

[66] S. I. Chernyshenko, P. Goulart, D. Huang, and A. Papachristodoulou. Polynomial sum of squares in fluid dynamics: a review with a look ahead. *Phil. Trans. Royal Soc. A: Mathem., Phys. Eng. Sciences*, 372(2020), 2014.

[67] M. V. Lakshmi, G. Fantuzzi, J. D. Fernández-Caballero, Y. Hwang, and S. I. Chernyshenko. Finding extremal periodic orbits with polynomial optimization, with application to a nine-mode model of shear flow. *SIAM J. Appl. Dyn. Sys.*, 19(2):763–787, 2020.

[68] A. Majumdar, A. A. Ahmadi, and R. Tedrake. Control design along trajectories with sums of squares programming. In *Proc. IEEE Int. Conf. Rob. Autom.*, 2013.

[69] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, Cal. Inst. Tech., 2000.

[70] G. Chesi. LMI techniques for optimization over polynomials in control: a survey. *IEEE Trans. Autom. Contr.*, 55(11):2500–2510, 2010.

[71] A. Papachristodoulou and S. Prajna. A tutorial on sum of squares techniques for systems analysis. In *Proc. Amer. Contr. Conf.*, 2005.

[72] Z. Jarvis-Wloszek, R. Feeley, W. Tan, K. Sun, and A. Packard. Control applications of sum of squares programming. In *Positive Polynomials in Control*, pages 3–22. Springer, 2005.

[73] W. Tan and A. Packard. Stability region analysis using polynomial and composite polynomial Lyapunov functions and sum-of-squares programming. *IEEE Trans. Autom. Contr.*, 53(2):565–571, 2008.

[74] Graziano Chesi, Andrea Garulli, Alberto Tesi, and Antonio Vicino. Lmi-based computation of optimal quadratic lyapunov functions for odd polynomial systems. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 15(1):35–49, 2005.

[75] J. Löfberg. Pre- and post-processing sum-of-squares programs in practice. *IEEE Trans. Autom. Contr.*, 54(5):1007–1011, 2009.

[76] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou. Chordal and factor-width decompositions for scalable semidefinite and polynomial optimization. *Ann. Rev. Contr.*, 2021.

[77] I. Markovsky and P. Rapisarda. Data-driven simulation and control. *Int. J. Contr.*, 81(12):1946–1959, 2008.

[78] J. Coulson, J. Lygeros, and F. Dörfler. Data-enabled predictive control: In the shallows of the DeePC. In *Proc. Eur. Contr. Conf.*, 2019.

[79] C. De Persis and P. Tesi. Formulas for data-driven control: Stabilization, optimality and robustness. *IEEE Trans. Autom. Contr.*, 65(3):909–924, 2020.

[80] T. Dai and M. Sznaier. A semi-algebraic optimization approach to data-driven control of continuous-time nonlinear systems. *IEEE Contr. Sys. Lett.*, 5(2):487–492, 2021.

[81] M. Guo, C. De Persis, and P. Tesi. Data-driven stabilization of nonlinear polynomial systems with noisy data. *IEEE Trans. Autom. Contr.*, 2021. Early access.

[82] A. Bisoffi, C. De Persis, and P. Tesi. Data-driven control via Petersen's lemma. arXiv preprint available at `https://arxiv.org/abs/2109.12175`, 2021.

[83] H. Hjalmarsson and L. Ljung. A discussion of "unknown-but-bounded" disturbances in system identification. In *Proc. IEEE Conf. Dec. Contr.*, 1993.

[84] M. Milanese and C. Novara. Set membership identification of nonlinear systems. *Automatica*, 40(6):957–975, 2004.

[85] T. Dai and M. Sznaier. A moments based approach to designing MIMO data driven controllers for switched systems. In *IEEE Conf. Dec. Contr.*, 2018.

[86] A. Bisoffi, C. De Persis, and P. Tesi. Trade-offs in learning controllers from noisy data. *Sys. & Contr. Lett.*, 154:104985, 2021.

[87] H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data: implementation issues and performance evaluation. *International journal of Control*, 76(13):1337–1357, 2003.

[88] C. De Persis, R. Postoyan, and P. Tesi. Koopman operator based iterative learning and model predictive control for repetitive nonlinear systems. In *preprint*, 2022.

[89] A. A. Ahmadi. Non-monotonic Lyapunov functions for stability of nonlinear and switched systems: theory and computation. Master's thesis, Mass. Inst. Tech., 2008.

[90] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proc. IEEE Int. Symp. Comp. Aid. Contr. Sys. Des.*, 2004.

[91] A. A. Ahmadi and B. El Khadir. Learning dynamical systems with side information. In *Proc. 2nd Conf. Learn. Dynam. Contr.*, 2020.

[92] Zhigang Bao, Xiucai Ding, and Ke Wang. Singular vector and singular subspace distribution for the matrix denoising model. *The Annals of Statistics*, 49(1):370–392, 2021.

[93] A. M. Tomé, D. Malafaia, A. R. Teixeira, and E. W. Lang. On the use of singular spectrum analysis. *arXiv*, 2018.

[94] Steven L Brunton and J Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control.* Cambridge University Press, 2022.