ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

INGEGNERIA BIOMEDICA, ELETTRICA E DEI SISTEMI

Ciclo 37

Settore Concorsuale: 09/G1 - AUTOMATICA

Settore Scientifico Disciplinare: ING-INF/04 - AUTOMATICA

LEARNING AND OPTIMAL CONTROL OF UNCERTAIN SYSTEMS VIA DATA-DRIVEN METHODS WITH STABILITY GUARANTEES

Presentata da: Marco Borghesi

Coordinatore Dottorato

Michele Monaci

Supervisore

Giuseppe Notarstefano

Co-supervisore

Ivano Notarnicola

Esame finale anno 2025

Abstract

With the advent of largely available computational power and datasets, a broad number of new algorithms for learning and control of dynamical systems has been proposed; however, the complexity of the considered frameworks has outpaced the theoretical analysis of these new techniques. To reduce this gap, this thesis focuses on the development of both theoretical tools and control algorithms for the learning and optimal control of uncertain systems. The main challenges which arise in these frameworks regard the problems of i) guaranteeing informative systems trajectories, ii) extracting the information from measurable quantities and iii) designing robustly stabilizing and optimal control laws from gathered data. We start by addressing the problem of giving Persistency of Excitation (PE) guarantees in the context of Linear Time-Invariant (LTI) systems, finding necessary and sufficient conditions to obtain this property via an input signal. Our results are developed within a notation which underlines the perfect analogies between the continuous- and discrete-time frameworks. Next, we study data-driven approaches for the stabilization of LTI systems. For the continuous-time framework, we start by addressing the design of model-free observers that extract full-state information from output data, and then we proceed with the design of stabilizing controllers via LMIs when state measurements and derivatives are unavailable. Next, we consider the Linear Quadratic Regulator (LQR) problem, and we propose a nonlinear on-policy controller which globally converges to the optimal feedback preserving the stability of the interconnection during the transient. Finally, leaving the linear framework, we design a model-free optimal control algorithm which, differently from other techniques, takes into account the problem of safety whilst performing the necessary exploration. A distinctive feature of this thesis is its retrospective look at classical techniques and concepts from the adaptive and linear multivariable control field, which we repropose in combination with more recent approaches and which we believe hold the answers to many current questions.

Keywords: Persistent Excitation, Sufficient Richness, Adaptive Control, Data-Driven Control, Reinforcement Learning, Optimal Control, Linear Quadratic Regulator, Numerical Optimal Control.

Acknowledgement

This thesis work was supported by the project PRIN 2022 ECODREAM Energy COmmunity management: DistRibutEd AlgorithMs and toolboxes for efficient and sustainable operations ,code [202228$CTKY$_002] - CUP [$J53D$23000560006], funded by the European Union - NextGenerationEU under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research.

Contents

Introduction

Motivation and challenges

From the first half the XX century, control theory has dealt with the problem of understanding intelligence and the flow of information, aiming to solve the problem of automating certain tasks [181, 194, 208, 220, 221, 225]. These topics have somehow diverged and specialized during the second half of the last century, giving birth to different literature branches: the first, using more heuristic approaches (expert systems, fuzzy logic, neural networks, etc.) and taking inspiration from natural intelligence to deal with mathematically untractable systems [81, 115]; the second, considering idealized dynamical systems and analyzing them in a more systematic and mathematical manner [99, 228].

Today, with the advent of a largely available computational power and datasets, artificial intelligence is becoming more and more pervasive into our lives. Whilst these new instruments - which have proved their potential in several applications such as economics and finance [51], racing [199], games [148, 196] and biology [167] - are undoubtly promising, they often preceed formal understanding and mathematical rigour, and they still have to demonstrate their thrustworthiness in real-world applications. In recent years, several attempts have been made to merge the capabilities of the machine learning to deal with complex and unknown dynamics with the robust guarantees and understanding which are proper of control theory. In this thesis, we take another step in this direction by considering the challenge of combining *learning* and *optimal control* - fundamentals both in machine learning and in control theory - in *uncertain systems*. By following the flow of the *information* needed to control a system, we give guarantees on i) the presence of this information into the system trajectories, ii) the possibility of extracting it entirely from the measurable outputs iii) how to use it to robustly stabilize a system and to achieve closed-loop optimality.

Through the years, several estimation techniques relying on data have been developed depending on the purpose of the estimation, and the major field developing these kind of techniques is system identification [15, 121, 123, 197, 211]. In general, the idea underlying these algorithms is to find a (tipically algebraic) relation between known quantities, and then to reconstruct this relation using a parametrized

function estimator and the collected data. A key aspect which determines the success of these algorithm is the quality of the collected data, which must carry the information about the quantities we are interested in. The problem of achieving successful estimation has been addressed in several seminal works resulting in the definition of the PE concept [33, 72, 132, 154, 155, 188, 206]; however, to the author's knowledge, there is no unified theory regarding how to impose PE on systems trajectories through the input signal (though several sufficient results are available for specific frameworks [42, 142, 226]). The specific framework proposed in this thesis investigates sufficient richness in LTI systems. However, embedding enough information into the collected data is essential in all data-driven frameworks (see, e.g., [84, 159, 200] for similar desirable properties in the context of neural network training).
In general, not all signals in which we are interested (for learning or control purposes) can be directly measured. The devices which reconstruct the internal state of a system are called observers [26, 103, 140], and a typical problem of their design is their reliance on the dynamics knowledge. This problem is addressed by adaptive observers [98, 112, 209], which reconstruct, together with the state information, a parametrization of the plant dynamics (however, introducing sometimes stability issues even in presence of PE trajectories).
Another challenge is understanding how to robustly stabilize an equilibrium or a trajectory in presence of uncertainties in the system model. Addressing this issue is fundamental in particular for safety-critical systems such as collaborative robotics [217] or aircraft control [201]. Several control techniques approach this problem in different ways. In this field, we find robust controllers [85, 234], which tipically rely on the knowledge of an upper bound for system noise or uncertainties (however, which do not exploit the gathered data to improve their performances), adaptive controllers [98, 157, 162, 191], whose founding idea is to adapt (leveraging on the collected data) a parametrized control law to achieve a zero error in the interested outputs, and other more recent data-driven algorithms, which use numerical techniques to estimate control laws directly from the data [64].
At last, we consider the problem of optimality, namely, the minimization of a given performance index in presence of uncertainties. This objective is meaningful since the definition of a performance index is a very intuitive (and broadly used) way to encode control objectives which go beyond the simple stabilization of a system [27, 46]. Several techniques have been developed to solve optimal control problems when the controlled system is unknown, such as black-box optimization, derivative-free optimization, extremum seeking, and others [3, 4, 53, 178, 203]. Though the specific approaches may be different, the underlying idea is to probe the cost function to find a descent direction. Reinforcement learning [177, 202] somehow preserves this idea, but shifting the optimization variable from system trajectories to control policies.
We believe that the clear understanding of all the fundamental mechanisms involved in learning, control and optimization is a necessary step for a clever and meaningful design of complex algorithms, and it is natural to study these mechanisms in a mathematically tractable context before applying them to real-world scenarios.

Summary of the main contributions

From an high-level perspective, the techniques developed in this thesis contribute to the scientific community in the following three aspects.
First, we approach the problem of guaranteeing data informativity (in terms of a persistence of excitation notion) in linear systems via the design of a proper input signal. Whilst the existing literature already contains several results, we address this problem from a more systematic way both for continuous- and discrete-time systems, providing new necessary conditions and highlighting how the multivariable structure of multi-input systems influences the ability to impose PE on system trajectories. Furthermore, we show how to exploit the spectral properties of the input signal together with the linear systems structure to obtain tighter sufficient conditions. We believe that a fundamental analysis of linear systems is the corner stone for nonlinear extensions of these results (as it is for other very important system theoretical properties such as controllability [92, 120]). A full paper with the presented result has been submitted for publication, and it is available on Arxiv [35].
Second, we tackle the problem of extracting the necessary information from the output data. The issue of full state unavailability was approached years ago by the adaptive control literature, and we draw inspiration from continuous-time techniques [77, 185, 209, 210, 231] which emphasize the importance of signal pre-processing in the implementation of identification and adaptation laws. Inspired by classical adaptive observers [162, Ch. 4], [5], we propose a framework for the development of a new type of observer - the gazer - able to obtain a representation of the unavailable state of a linear system without relying on adaptive laws or prior model knowledge. The underlying idea is to trade this necessity with the nonminimality of the resulting representation. We then investigate the feasibility of this approach, studying how the observability of the plant and the controllability of the gazer interact to ensure the existence of solutions. We stress that some of these results are still preliminary results, and further investigations are needed.
Third, we consider the problem of using the extracted information to achieve robust stabilization with optimality guarantees. In an initial study, we consider a simple offline scenario, where optimality is not required and where the data have been previously collected. In this case, the problem is the design of a state-feedback gain. Recent literature often assumes access to state derivatives; however, this is not always possible, and trying to estimate them may introduce additional, unwanted noise. For solving this problem, we leverage the above presented gazer. A conference article containing the results of this work has been submitted for publication and is currently available on Arxiv [38]. Next, we consider an online, on-policy LQR setting which takes into account all the problems of learning, robust stabilization and, optimality above mentioned altogether. In this context, we propose a common framework for the design of online, on-policy regulators aimed at achieving optimality; furthermore, we provide a new algorithm inspired to the model-reference adaptive controllers [205], able to perform learning, stabilization, and optimization altogether. To the authors' knowledge, this is one of the first works possessing all these properties, and a preliminary version of it has been presented in a conference paper CDC 2023 [36]. A full version of this work is currently under peer review and it is available on Arxiv [37].

At last, we extend our investigation to the nonlinear framework. In this case, we consider a finite-horizon optimal control problem where the system is unknown, and we design an optimization algorithm which, differently from reinforcement learning techniques, takes into account the problem of avoiding numerical instabilities whilst the dynamics and the cost function are being probed. In particular, we achieve this desirable feature by drawing inspiration to PRONTO, a solver for nonlinear optimal control which leverages on a projection operator, namely, a stabilizing controller to be designed separately, which ensures robustness to the algorithms evolution and to the dynamics exploration. An article with the presented result is currently under preparation, to be soon submitted.

Organization and chapter contribution

Chapter I

In Chapter I, we formally present the uncertain frameworks central to this thesis. We start with the uncertain LQR problem, and we introduce the reader to the problems of signal reconstruction, uncertainties estimation, system stabilization and optimization of a performance index which are faced in Chapters II, III, IV. Then, we consider the problems of finite-horizon nonlinear optimization and its challenges, which are explored in Chapter V.

Chapter II

This chapter addresses the problem of guaranteeing PE of systems trajectories in the context of LTI systems, and contributes in three ways to the scientific community. First, we find necessary and sufficient conditions on the input for obtaining persistently excited state (state-input) signals.

i) Necessary conditions for sufficient richness are completely new results for the case of multi-input systems, and -to the author's knowledge- the only necessary result existing in the literature is given for a finite-time notion of PE ("Willems' PE", see Section II.1) and only for discrete-time single-input systems in [143, Thm. 3]. In order to find these conditions, we leverage on the concept of Partial Persistence of Excitation -namely, signals which persistently span only subspaces of the whole space they live in-, which firstly appeared in [161, Def. 3] but, to the author's knowledge, was never used up to now. We also remark that the arguments used in the proofs can be used for the design of input sequences that optimize certain "exploration" performance indices (see Remark II.10).

ii) Sufficient conditions to obtain PE in LTI systems are not completely new results (see, e.g., [17, 86, 152]). However, the cited works only provide sufficient conditions for discrete-time systems. For the continuous-time framework, there exists different results, which differ from our result in the following way. In [98, Thm. 5.2.3], a sufficient condition is found on the spectral measure of each scalar input channel only for stationary signals. In [176, Prop. 1], the result is similar to ours; however, the notion of PE used in the article is different from Definition

II.14 since it is a finite-time notion ("Willems' PE", see Section II.1) which does not require the existence of a uniform lower bound in time for the energy of the span in each direction. Next, in [142, Thm. 4.2] the author finds a characterization for sufficient richness which, however, does not consider pure signal derivatives, but some filtered version of them.

iii) At last, we combine the obtained necessary and sufficient results to derive and discuss the shape of the sets of Sufficiently Rich signals for stable, reachable, LTI systems. It turns out that, for the case of single input, SR signals are an open cone in the space of input signals (independent on the particular system and in accordance with [161, Thm. 1]). In the case of multi-input systems, this set is not anymore independent on the considered system. We corroborate the theoretical results by providing two numerical examples which show the tightness of the new results.

The second contribution of this chapter is the proposal of clear, separate concepts of persistence of excitation and sufficient richness, and the development of all the above results within a notation unifying discrete-time and continuous-time framework. This contribution is concrete for the following reasons:

i) Separating the concepts of PE and SR, we clarify how we could use them in a non-overlapping way, and we avoid the need to talk about "PE of order" or "SR of order". The distinction between these concepts reduces the proliferation of equivalent (but apparently different) definitions and terminology that permeates the field. Furthermore, all specific results (which lead to specific definitions) are all seen as particular characterizations of the same property.

ii) By unifying the results and the notation for the discrete-time and the continuous-time framework, we better highlight the structural properties which are intrinsic of linear systems; more specifically, in their ability of obtaining PE signals. We believe this is a fundamental step for proceeding with this research and extending these results to nonlinear systems.

At last, we show how to exploit the geometric structure of linear systems together with their ability of preserving the spectral content of an input signal to obtain tighter sufficient conditions for PE in multi-input systems.

i) By leveraging on an insightful decomposition for linear systems [228] and on PE characterizations in the frequency domain [41], we improve the sufficient results given first in [232, Thm. 3, Thm. 4] and then in [98, Thm. 5.2.3]. Specifically, with respect to the first work we give a more generic result involving the spectral content of the input signal, halving the required number of sinusoids and not necessarily asking for almost periodic input signals. Furthermore, our result is not "almost certain" (see [232, Thm. 3]). With respect to the second work, we better highlight the structure of linear systems showing how to reduce the required number of spectral lines in the input signal.

Chapter III

This chapter contribution is twofold. At first, we introduce a new type of observer, the "gazer", which, differently from classical observers, is designed for model-free signal reconstruction and is not focused in obtaining the true state of the linear system.

i) Inspired by adaptive observer design [162, Ch. 4], [5], we propose a framework for the development of filters able to obtain a representation of the unavailable state of a linear system. Requirements and sufficient conditions for their fullfillment are thus given for this type of observer.

ii) Structural properties (such as observability of the plant and controllability of the gazer) are studied in their role of guaranteeing the existence of a solution to the gazing problem via certain matrix equations, whose solution are analyzed. Leveraging all of these results, we show how classical filters can be seen as gazers in filter form, and we derive a new gazer for MIMO systems.

Second, we develop a data-driven stabilization technique for continuous-time systems that eliminates the need for signal derivatives. The approach is first demonstrated for state-feedback stabilization and then extended with minor modifications to the single-input single-output (SISO) output-feedback scenario.

i) Instead of using derivative approximations or methods based on integrals and temporal differences, we leverage the above presented gazers to define a non-minimal realization of the plant.Specifically, we process the input and state/output signals with a low-pass filter that is shown to converge exponentially to an augmented system representation. Since both the state and the derivative of the filter are accessible, we employ LMIs similar to those in [62] to compute the gains of a dynamic, filter-based, stabilizing controller. Feasibility of the LMIs is ensured under suitable excitation conditions, and closed-loop stability is guaranteed regardless of the initial filter transient.

Chapter IV

This chapter's contribution is twofold. At first, we introduce a novel formulation of the on-policy data-driven LQR problem where centrality is given to the stability of the whole closed-loop learning and control system.

i) Concerning our first contribution, we formulate the on-policy data-driven LQR problem in terms of convergence of the controller, the plant, and an exosystem (modeling the dither signal) to an asymptotically stable set . The fundamental property defining this set is that the learned policy is optimal. Additionally, the set becomes smaller as the dither amplitude is reduced. Thanks to this formulation, we ensure that asymptotic stability in the nominal scenario is preserved, practically and semiglobally, also for a broad class of perturbations, see [79, Ch. 7]. With the generality of the proposed framework, we aim to provide a solid foundation for future work in the field.

The second contribution of this chapter to the scientific community is a specific solution to the above presented problem, namely the *Model Reference Adaptive Reinforcement Learning* (MR-ARL), a control algorithm integrating concepts from system identification, adaptive control, and reinforcement learning paradigms. The architecture is structured as a modular actor-critic system with a time-varying reference model bridging the two modules. The actor, inspired to an Model Reference Adaptive Controller (MRAC) architecture, guides the plant to a desired behavior set by the reference model, even in the presence of parametric uncertainties. The reference model is updated online by the critic, which leverages system identification techniques to estimate the dynamics.

i) To impose the desired learning properties, the reference model is driven by a dither signal for which we require suitable richness properties. By relying on analysis tools related to adaptive control, differential inclusions, and singular perturbations, we prove formally that our architecture achieves the following properties for the whole closed-loop system: (i) convergence of the policy to the optimal one; (ii) asymptotic estimation of the true system parameters; (iii) uniform asymptotic stability of an attractor (tunable with the dither amplitude); (iv) robustness in the sense of semiglobal practical asymptotic stability with respect to unmodeled nonlinearities and disturbances. To the best of authors' knowledge, in the context of on-policy data-driven LQR, this algorithm is the first one possessing all these properties.

ii) Differently from other works, we require no assumptions about the initial policy. Further, persistency of excitation, needed to ensure convergence, is not assumed a priori, but rather guaranteed by design by resorting to concepts from nonlinear adaptive systems [171].

iii) Given the inherent robustness of the proposed design framework, we ensure that the algorithm is effective in the presence of slowly varying parameters [79, Cor. 7.27] automatically adapting to changes and recovering optimality without having to stop the system. To validate this property, our numerical simulations cover both the constant parameters case and the one with drifts.

iv) As a byproduct of our approach, we study for the first time in MRAC literature (to the author's knowledge) the stability of an MRAC where the reference model is not fixed but its state and input matrices are time varying.

Chapter V

The main contribution given in this chapter is `DATA-DRIVEN PRONTO`, an algorithm for nonlinear optimal control extending the applicability of `PRONTO` [90] to the model-free scenario. In particular, `DATA-DRIVEN PRONTO` leverages the possibility to probe the system dynamics in a safe way via the application of an independently-designed control law (which is required in [90] under the name of "projection operator"). At each algorithm iteration, we perform L experiments (simulations) in which we use the given controller plus an exploration dither to gather data in the neighbourhood of the current trajectory. This data is then used to identify the dynamics linearizations about the current trajectory, substituting the need of a model in the optimization process. The proposed algorithm achieves the following desirable features:

i) Avoid parametrization errors: `DATA-DRIVEN PRONTO` uses data only to identify linearizations of the dynamics, avoiding any error that would occur when identifying or estimating a parametrization of the nonlinear dynamics or optimal policy.

ii) Fully data-driven: `DATA-DRIVEN PRONTO` itself does not require any knowledge of the system; however, the importance of partial system knowledge is needed to improve the numerical stability of the algorithm via the independent design of a good control law.

iii) Data-efficient: thanks to the efficient algorithm structure, the number of required experiments (simulations) needed is significantly smaller than the one needed by RL algorithms, since at each iteration it is comparable with the system dimension.

We provide theoretical guarantees on the algorithm convergence properties, together with insight on how to choose hyperparameters. The proposed algorithm is proved to be Locally Uniformly Ultimately Bounded to a ball about the optimal solution whose radius is related to the quality of the collected data (which can be augmented by enlarging the number of collected explorations and by reducing the amplitude of the injected exploration dither).

Appendix

We leave in the Appendix most of the proofs of the several lemmas, propositions and theorems. The general criteria is that when a proof does not interrupt the discussion (being it very brief or informative), it is kept in the main body of the thesis; otherwise, it is left in the Appendix. Throughout this thesis, we present some known results for readability reasons. However, all the proofs contained in this thesis (main body or Appendix) refer to new results.

Notation

Abbreviation	**Meaning**
♙	End of lemma
♖	End of proposition
♕	End of theorem
♔	End of definition
ARE	Algebraic Riccati Equation
CT	Continuous-Time
DRE	Differential Riccati Equation
DT	Discrete-Time
LMI	Linear Matrix Inequality
LTI	Linear Time-Invariant
LTV	Linear Time-Varying
MI	Multi-Input
MIMO	Multi-Input Multi-Output
PE	Persistency of Excitation / Persistently Excited
RL	Reinforcement Learning
RMS	Root-Mean-Squared
SI	Single-Input
SISO	Single-Input Single-Output
SR	Sufficient Richness / Sufficiently Rich

Symbol	**Meaning**
$\mathbb{N}$	Natural numbers
$\mathbb{N}_1$	Natural numbers
$\mathbb{R}$	Real numbers
$\mathbb{R}_{\geq 0}$	Positive real numbers
$\mathbb{S}$	Unit circle
C^n	Functions $n-$times continuously differentiable
$\mathrm{x}, \mathrm{u}, \mathrm{w}$	When used with the Roman font, indicates the infinite-dimensional continuous- or discrete-time signal $x(t)$ or x_t
$C_b^\infty(\mathbb{R}^d)$	Linear space of all infinitely differentiable functions $f : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ with bounded derivatives equipped with norm $\|\cdot\|_\infty$
$\ell_\infty(\mathbb{R}^d)$	Linear space of all bounded functions $f : \mathbb{N} \to \mathbb{R}^d$ equipped with norm $\|\cdot\|_\infty$
$\mathbb{B}_r$	Open ball in $\mathbb{R}^n$ space (of appropriate dimension), centered in zero and of radius $r > 0$
$\mathbb{B}_r(x)$	Open ball in $\mathbb{R}^n$ centered in $x \in \mathbb{R}^n$ and of radius $r > 0$
$o_x(y)$	Function of $x \in \mathbb{R}^n, y \in \mathbb{R}^m$ such that $\lim_{x\to 0} \|o_x(y)\|\|y\|^{-1} = 0$.
Class $\mathcal{K}$ function	Function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that it is strictly increasing and $\alpha(0) = 0$
Class $\mathcal{KL}$ function	Function $\phi : [0, a) \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\phi(r, s)$ is a class $\mathcal{K}$ function for all fixed $s \in \mathbb{R}_{\geq 0}$, and $\phi(r, s)$, for all fixed r, is decreasing and such that $\phi(r, s) \to 0$ for $s \to \infty$
$\nabla_1 f(\bar{x}, \bar{u})$	Given $f(x, u)$ differentiable in the first argument at $\bar{x}, \bar{u}$, $\nabla_1 f(\bar{x}, \bar{u}) := \frac{\partial f(x,u)}{\partial x}\Big\|_{\bar{x},\bar{u}}^\top$
$\boldsymbol{S}_0^n, \boldsymbol{S}_+^n$	Cone of $n \times n$ symmetric, positive semidefinite (resp. positive definite) real matrices
$\dot{\xi} = f(t, \xi), \xi \in C \subset \mathbb{R}^r$	Differential equation with flow set and initial state constrained to C [79]

Symbol	**Meaning**
$\lvert \cdot \rvert$	Euclidian norm for vectors and spectral norm for matrices
$\lvert \cdot \rvert_F$	Frobenius norm
$\lVert \cdot \rVert_\infty$	$\mathcal{L}_\infty$ norm
$\boldsymbol{x} = (x_1, x_2, \ldots)$	Column vector stacking a finite number of elements $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}, \ldots$
$\mathrm{diag}(A_1, A_2, \ldots)$	Block-diagonal matrix obtained stacking diagonally the matrices (vectors) $A_1 \in \mathbb{R}^{n_1 \times m_1}$, $A_2 \in \mathbb{R}^{n_2 \times m_2}, \ldots$ and zero in all other elements
$\otimes$	Kronecker product
$\kappa(M)$	Condition number of the matrix $M \in \mathbb{R}^{n \times m}$
$M^\dagger$	Moore-Penrose pseudoinverse of matrix M
$\mathcal{U}(a, b)$	Uniform distribution between $a, b \in \mathbb{R}$
$\sigma(M)$	Spectrum of the square matrix M
$M^\top$	Transpose of the real matrix M
M^*	Complex conjugate of the complex matrix M
I_n	Identity operator on $\mathbb{R}^n$

Chapter I

Optimal control of uncertain systems

In this chapter, we motivate and introduce mathematically the frameworks and the challenges which are explored in the subsequent chapters of this thesis. Similarly to the way the thesis is structured, we start by considering a linear framework and, more specifically, we begin by introducing the LQR problem and its challenges in the context of uncertain systems. Next, motivated by the need of taking into account more complex dynamical systems, we move to a the nonlinear optimal control framework.

I.1 Linear quadratic regulation in presence of uncertainties

Consider a multi-input, multi-output linear system

$$\begin{aligned} \dot{x} &= A(\mu)x + B(\mu)u \\ y &= C(\mu)x \end{aligned} \tag{I.1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input, and $y \in \mathbb{R}^p$ is the measurable output. We represent the uncertainties in the system via the unknown parameter $\mu \in \mathcal{K}_\mu \subseteq \mathbb{R}^q$ and the maps

$$A : \mathcal{K}_\mu \to \mathbb{R}^{n\times n} \qquad B : \mathcal{K}_\mu \to \mathbb{R}^{n\times m} \qquad C : \mathcal{K}_\mu \to \mathbb{R}^{p\times n}. \tag{I.2}$$

Depending on the specific framework, the uncertainties may be both in μ and in the maps A, B, C; furthermore, one may have to deal both with constant or time-varying uncertainties $\mu(t)$, $A(\mu, t)$, $B(\mu, t)$, $C(\mu, t)$ see, e.g., Section IV.3.3 for an example of such systems. While in this thesis the focus is on constant uncertainties, we deal implicitly with slowly time-varying uncertainties and other types of perturbations (process and output noise) leveraging on robustness results.

Example I.1. A Doubly Fed Induction Motor (DFIM) can be modeled [117] with a linear system in the form of (I.1) with state and input

$$x = (i_{1u}, i_{1v}, i_{2u}, i_{2v}) \in \mathbb{R}^4, \qquad u = (u_{1u}, u_{1v}, u_{2u}, u_{2v}) \in \mathbb{R}^4, \tag{I.3}$$

where $i_{1u}, i_{1v}, i_{2u}, i_{2v}$ are the stator and rotor currents, and $u_{1u}, u_{1v}, u_{2u}, u_{2v}$, are the stator and rotor voltages. System matrices are defined as

$$A = \frac{1}{\bar{L}} \begin{bmatrix} -L_2R_1 & -\alpha+\beta & L_mR_2 & \beta_2 \\ \alpha-\beta & -L_2R_1 & -\beta_2 & -L_mR_2 \\ L_mR_1 & -\beta_1 & -L_1R_2 & -\alpha-\beta_{12} \\ \beta_1 & L_mR_1 & \alpha+\beta_{12} & -L_1R_2 \end{bmatrix}, \quad B = \frac{1}{\bar{L}} \begin{bmatrix} L_2 & 0 & -L_m & 0 \\ 0 & L_2 & 0 & -L_m \\ -L_m & 0 & L_1 & 0 \\ 0 & -L_m & 0 & L_1 \end{bmatrix}, \tag{I.4}$$

with

$$\begin{gathered} \bar{L} := L_1L_2 - L_m^2, \quad \alpha := \bar{L}\omega_0, \quad \beta := L_m^2\omega_r, \\ \beta_{12} := L_1L_2\omega_r, \quad \beta_1 := L_1L_m\omega_r, \quad \beta_2 := L_2L_m\omega_r. \end{gathered} \tag{I.5}$$

Parameters R_1, R_2 are the stator and rotor resistances, L_1, L_2, L_m are the stator and rotor auto-inductances and the mutual inductance, and ω_0, ω_r are the electrical angular speeds of the rotor and the rotating reference frame. In this case, we may define the vector of uncertainties μ as

$$\mu := (R_1, R_2, L_1, L_2, L_m, \omega_r, \omega_0), \tag{I.6}$$

and the maps (I.2) can be found by expressions (I.4), (I.5). ◇

We associate to system (I.1) a known quadratic performance index

$$J(y(\cdot), u(\cdot)) := \int_0^\infty \left(y(\tau)^\top Q y(\tau) + u(\tau)^\top R u(\tau) \right) \mathrm{d}\tau, \tag{I.7}$$

and our goal is to solve the LQR problem

$$\begin{aligned} \min_{y(\cdot),x(\cdot),u(\cdot)} \; & \int_0^\infty \left(y(\tau)^\top Q y(\tau) + u(\tau)^\top R u(\tau) \right) \mathrm{d}\tau \\ \text{subj. to:} \; & \dot{x}(t) = A(\mu)x(t) + B(\mu)u(t) \\ & y(t) = C(\mu)x(t). \end{aligned} \tag{I.8}$$

It is known that, under certain assumptions on the matrices A, B, C, Q, R, the solution to (I.8) has a closed form, and the input $u(\cdot)$ has to be chosen according to the feedback law

$$u(\mu, t) = -R^{-1}B^\top(\mu)P(\mu)x(t), \tag{I.9}$$

where $P(\mu)$ is the solution to the algebraic Riccati equation

$$A(\mu)^\top P + PA(\mu) - PB(\mu)R^{-1}B(\mu)^\top P + C(\mu)^\top QC(\mu) = 0. \tag{I.10}$$

In our framework, however, three major problems arise in the implementation of control law (I.9). First, the knowledge of $x(t)$ is required, and the full state measurement is, in general, not always possible. Next, solving (I.10) requires the knowledge of the unknowns $A(\mu)$, $B(\mu)$, $C(\mu)$, which must be learned from the collected data. This means that the gathered data must be informative. Third, in online, on-policy frameworks one needs to consider also the problem of guaranteeing some form of stability of the closed-loop system. In the following, we discuss more in details these challenges.

Guaranteeing informativity of system trajectories

For several reasons - state reconstruction, system stabilization, optimal control or simply identification - one may be interested in finding the unknown matrices $A(\mu)$, $B(\mu)$, $C(\mu)$ (or directly the parameter μ). Supposing that we are interested in finding μ, and that we have the availability of the output y, the input u and some generic, additional signal ζ, we define the estimate $\hat{\mu}$ and look for update laws of the form

$$\dot{\hat{\mu}} = h(\hat{\mu}, y, u, \zeta) \tag{I.11}$$

such that $\hat{\mu} \to \mu$. The shape of h is typically found by leveraging on a known algebraic relation between the available quantities ζ, y, u and μ.

Example I.2. As an example, suppose that the uncertainties are only in matrix $C(\mu) \coloneqq \mu$ in (I.1) and that x is available. Since it holds $y = \mu x$, we introduce the estimate $\hat{\mu}(t)$ and define the instantaneous cost function

$$J(\hat{\mu}(t)) \coloneqq |y(t) - \hat{\mu}(t)x(t)|^2 \tag{I.12}$$

whose gradient in $\hat{\mu}$ is given by $\nabla J(\hat{\mu}) = 2(y - \hat{\mu}x)x^\top$. We then choose a simple gradient rule to define the update

$$\dot{\hat{\mu}} = 2(y - \hat{\mu}x)x^\top. \tag{I.13}$$

◇

Notice however that, without assuming any additional property of the regressor $x(t)$,

$$\lim_{t\to\infty} |y(t) - \hat{\mu}(t)x(t)|^2 = 0 \quad \Longrightarrow \quad \lim_{t\to\infty} |(\mu - \hat{\mu}(t))x(t)|^2 = 0 \quad \not\Rightarrow \quad \lim_{t\to\infty} |\mu - \hat{\mu}(t)|^2 = 0. \tag{I.14}$$

More in general, finding a "good" update law (I.11) to fit some known algebraic relation is not enough to reach the goal $\hat{\mu} \to \mu$. In order to guarantee this objective, some regularity condition on $x(t)$ has to be fulfilled. The problem of the estimation convergence has been deeply analyzed by the adaptive control literature, and it has been found that it is solved if a persistency of excitation condition on the regressor is met.

Thesis contribution I.1. In Chapter II, we find necessary and sufficient conditions on the input $u(t)$ to achieve persistency of excitation of the signals $x(t)$ and $(x(t), u(t))$ in LTI systems. Furthermore, the presented results do not require the model knowledge. ⋆

Extracting the state information from available signals

A key challege in solving almost any control problem is having the access to the full state $x(t)$ of the controlled system. In fact, reconstructing the state allows for i) the estimation of the unknown state matrix, ii) the design of state-feedback controllers, and iii) the implementation of the optimal policy (I.9). More in general, in the continuous-time framework we are interested not only in reconstructing $x(t)$, but also in substituting the knowledge of its derivative $\dot{x}(t)$. Since the parameter μ is unknown, model-based observers are not enough for these purposes. In the adaptive literature, this problem is typically addressed via adaptive observers. Introducing the state estimate $\hat{\zeta} \in \mathbb{R}^n$, adaptive observers may be in general written as dynamical systems in the form

$$\begin{aligned} \dot{\hat{\zeta}} &= f(\hat{\vartheta}, \hat{\zeta}, y, u), \\ \dot{\hat{\vartheta}} &= g(\hat{\vartheta}, \hat{\zeta}, y, u) \end{aligned} \tag{I.15}$$

where $\hat{\vartheta} \in \Theta$ is an estimate of a certain function of the plant parameters $\vartheta(\mu)$ (e.g., the characteristic polynomial of $A(\mu)$). Notice that this technique introduces in the control law an auxiliary dynamics $\dot{\hat{\zeta}}$ and it requires some estimate of the true system parameters (and thus, its convergence properties need to be studied together with the update laws for $\dot{\hat{\vartheta}}$). On the other hand, another way to asymptotically reconstruct the state is to build a nonminimal realization of the plant introducing only an auxiliary variable $\hat{\zeta} \in \mathbb{R}^z$, with $z > n$, governed by dynamics

$$\dot{\hat{\zeta}} = m(\hat{\zeta}, y, u) \tag{I.16}$$

without requiring an estimate of the parameters $\vartheta(\mu)$. In this case, the subsequent design of the controller should consider the stabilization of the nonminimal plant.

Example I.3. Consider a simple scalar dynamics $\dot{x} = ax + bu$ with output map $y = x$, with non-measurable $\dot{x}$ and with a, b unknown. By implementing, for some $\lambda > 0$, the filter

$$\dot{\hat{\zeta}} = -\lambda\hat{\zeta} + \begin{bmatrix} x \\ u \end{bmatrix}, \tag{I.17}$$

defining $\epsilon := x - \Pi\hat{\zeta}$, $\Pi = [a + \lambda \;\; b]$ and calculating $\dot{\epsilon}$, we can rewrite (I.17) as the cascade system

$$\begin{aligned} \dot{\epsilon} &= -\lambda\epsilon \\ \dot{\hat{\zeta}} &= \begin{bmatrix} a & b \\ 0 & -\lambda \end{bmatrix}\hat{\zeta} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 1 \\ 0 \end{bmatrix}\epsilon, \end{aligned} \tag{I.18}$$

which converges asymptotically to

$$\dot{\hat{\zeta}} = \begin{bmatrix} a & b \\ 0 & -\lambda \end{bmatrix} \hat{\zeta} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \tag{I.19}$$

By taking the output $\xi = \Pi\hat{\zeta}$, it can be checked that we have found a nonminimal realization of the system $\dot{x} = ax + bu$, whose derivatives $\dot{\hat{\zeta}}$ are available from (I.17) and do not require knowledge of a, b to be implemented. ◇

Thesis contribution I.2. In Chapter III, we reframe classical filtering techniques developed for continuous-time system identification and adaptive control. Next, we apply them for the robustification of more recent numerical techniques (LMIs) dealing with data batches for the design of stabilizing control laws. ★

Stabilization and convergence to the optimal policy

Provided that the problems of reconstructing unknown signals and estimating unknown parameters are solved, we still need to solve the optimization problem (I.8), reaching the optimal control law (I.9) keeping into account the closed-loop stability of the controlled system. The estimation of the solution of the ARE (I.10), or directly estimating the optimal feedback $K(\mu) := -R^{-1}B(\mu)P(\mu)$ in presence of uncertainties may be approached in two ways.

First, one may implement techniques from the reinforcement learning field such as policy iteration or value iteration to use the collected data $\hat{\zeta}, u$ to directly find an estimate $\hat{P}$ or $\hat{K}$. These techniques read as (seen as continuous-time updates)

$$\dot{\hat{P}} = \varphi(\hat{P}, \hat{\zeta}, u). \tag{I.20}$$

In an on-policy, online framework where one needs to reach optimality, this estimate $\hat{P}$ needs to enter directly in the control law, and thus the stability properties of this update need to be studied together with the rest of the control law.

A second approach, which however requires an available estimate $\hat{\mu}$ of μ, is to apply the certainty equivalence principe and to find the solution $\hat{P}$ of

$$A(\hat{\mu})^\top \hat{P} + \hat{P}A(\hat{\mu}) - \hat{P}B(\hat{\mu})R^{-1}B^\top(\hat{\mu})\hat{P} + C(\hat{\mu})^\top QC(\hat{\mu}) = 0. \tag{I.21}$$

In this case, provided that $\hat{\mu} \to \mu$, we obtain $\hat{P} \to P(\mu)$.

Example I.4. In an online context, solving (I.21) at each time instant may not be practical. However, this problem can be circumvented with different approaches, e.g. i) calculating the solution only via discrete-time jumps or ii) implementing the differential Riccati equation

$$\dot{\hat{P}} = A(\hat{\mu})^\top \hat{P} + \hat{P}A(\hat{\mu}) - \hat{P}B(\hat{\mu})R^{-1}B^\top(\hat{\mu})\hat{P} + C(\hat{\mu})^\top QC(\hat{\mu}) \qquad \hat{P}(0) \geq 0. \tag{I.22}$$

◇

Thesis contribution I.3. In Chapter IV, to formalize the concepts presented in this introductory sections, and to provide a solution to online uncertain LQR problem developing a derivative-free algorithm which guarantees convergence $\hat{\mu} \to \mu$, $\hat{P} \to P$ while stabilizing the closed-loop system. ★

I.2 Nonlinear optimal control in presence of uncertainties

In the last part of the thesis, we consider an uncertain nonlinear system with dynamics given by

$$\dot{x} = f_c(x, u, \mu) \tag{I.23}$$

where $x \in \mathcal{M} \subseteq \mathbb{R}^n$ is the (measurable) state, $u \in \mathbb{R}^m$ is the input and x_{init} is the initial condition. We represent the uncertainties in the system via the constant unknown parameter $\mu \in \mathcal{K}_\mu \subseteq \mathbb{R}^q$ and the unknown static map

$$f_c : \mathcal{M} \times \mathbb{R}^m \times \mathcal{K}_\mu \to \mathbb{R}^n. \tag{I.24}$$

Example I.5. The pendubot [195] is a robot consising of two links and one actuator on the first joint. Its dynamics reads as

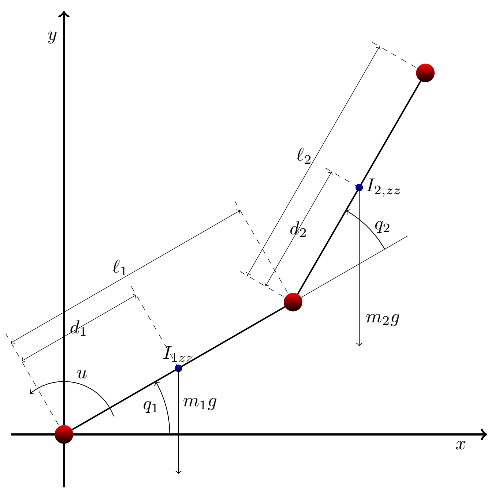

Figure I.1: The pendubot.

$$\begin{bmatrix} a_1 + a_2 + 2a_3\cos(q_2) & a_2 + a_3\cos(q_2) \\ a_2 + a_3\cos(q_2) & a_2 \end{bmatrix}\begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \end{bmatrix} + \begin{bmatrix} a_4\cos(q_1) + a_5\cos(q_1+q_2) \\ a_5\cos(q_1+q_2) \end{bmatrix} + \\ + \begin{bmatrix} -a_3\sin(q_2)\dot{q}_2 + f_1 & -a_3\sin(q_2)(\dot{q}_1+\dot{q}_2) \\ a_3\sin(q_2)\dot{q}_1 & f_2 \end{bmatrix}\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} = \begin{bmatrix} u \\ 0 \end{bmatrix}, \tag{I.25}$$

where $q = (q_1, q_2) \in [0, 2\pi]^2$ stacks the two joint angles $q_1, q_2 \in \mathbb{R}$, $u \in \mathbb{R}$ is the input torque on the first joint, and the coefficients $a_1, \dots, a_5$ are given by

$$\begin{aligned} &a_1 = I_{1,zz} + m_1 d_1^2 + m_2 \ell_1^2, \qquad a_2 = I_{2,zz} + m_2 d_2^2, \qquad a_3 = m_2 \ell_1 d_2, \\ &a_4 = g(m_1 d_1 + m_2 \ell_1), \qquad a_5 = g m_2 d_2. \end{aligned} \tag{I.26}$$

In this case, we may define the vector of uncertainties μ as the vector stacking the unknown masses, moments of inertia, length of links and positions of their centers of mass

$$\mu := (m_1, m_2, d_1, d_2, \ell_1, \ell_2, I_{1,zz}, I_{2,zz}), \tag{I.27}$$

and the map (I.24) can be found by expressions (I.25), (I.26). ◇

We associate to the dynamics (I.23) the finite-horizon nonlinear optimal control problem

$$\begin{aligned} \min_{x(\cdot),u(\cdot)} \quad & \int_0^T \ell(x(\tau), u(\tau)) \mathrm{d}\tau \\ \text{subj. to} \quad & \dot{x}(t) = f_c(x(t), u(t), \mu) \\ & x(0) = x_{\text{init}}, \end{aligned} \tag{I.28}$$

where x_{init} is assumed to be known. With respect to the previous setup, we consider a finite-horizon framework which is better suited for certain applications such as the offline computation of, e.g., industrial robot movement, or minimum-time trajectories for vehicles or aircrafts. Furthermore, since we are interested in solving numerical instances of problem (I.28), we discretize the dynamics, obtaining the problem

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{u}} \quad & \sum_{\tau=0}^{T} \ell(x_\tau, u_\tau) \\ \text{subj. to} \quad & x_{t+1} = f_d(x_t, u_t, \mu) \\ & x_0 = x_{\text{init}}. \end{aligned} \tag{I.29}$$

Notice that, whilst we have added some assumptions with respect to the previous linear framework, these are more than compensated by the complexity of the nonlinearity (and possible nonconvexity of the problem (I.29)). Furthermore, the uncertainties in μ and $f_d(x, u, \mu)$ make the solution of problem (I.29) hard to be found. In the following, we sum up the two challenges that we face in Chapter V.

Optimality guarantees

Being the problem (I.29) uncertain, several approaches may be considered. First, one may suppose to know the set $\mathcal{K}_\mu$ in which the uncertainties are confined, and add as additional constraints the system dynamics for all possible parameters $\mu \in \mathcal{K}_\mu$. In this way, at the price of suboptimality (and provided feasibility is preserved), we may still obtain the best solution with the available information. A problem of this approach is that the elements of $\mathcal{K}_\mu$ are typically infinite, and so a semi-infinite program has to

be solved.
A second way to deal with this problem (the reinforcement learning approach) is having the possibility of repeating experiments multiple times. In particular, one may parametrize a feedback control law in θ, say $\pi(\theta, x)$, and then, at the end or the k-th experiment, update θ^{k+1} according to some gradient information (instead of directly updating the resulting trajectory $\boldsymbol{x}^k, \boldsymbol{u}^k$):

$$\theta^{k+1} = g(\theta^k, \boldsymbol{x}^k, \boldsymbol{u}^k). \tag{I.30}$$

Data efficiency and safety of the dynamics explorations

A problem of the above mentioned approach is that one needs to probe the dynamics function with additional dither to guarantee a successful learning -exactly as for linear systems -. In other words, it is necessary to repeat experiments of the type

$$\begin{aligned} u_t^k &= \pi(\theta^k, x_t^k) + d_t^k \\ x_{t+1}^k &= f(x_t^k, u_t^k, \mu), \end{aligned} \tag{I.31}$$

where k denotes the experiment, in order to obtain improvements in the experiment cost $\sum_{\tau=0}^{T} \ell(x_\tau^k, u_\tau^k)$. The number of required explorations, in order to find good approximations of the solution, depends strongly on the dimension of the parameter θ, which should be large enough to find good approximations of the optimal solution. Furthermore, both the update of θ^k and the addition of the dither d_t^k introduce the risk of possible system instabilities, even when starting from safe initializations.

Thesis contribution I.4. In Chapter V, we tackle these problems by studying how to leverage on a known, well-behaved, fixed control policy $\pi(x)$ which does not necessarily enter in the optimization process, and we exploit its tracking properties to reduce the effects of the possible instabilities of the necessary exploration. Furthermore, we avoid the problem of performing a high number of experiments avoiding policy parametrizations and keeping $(\boldsymbol{x}, \boldsymbol{u})$ as the only optimization variable. ⋆

Chapter II

Persistent excitation and sufficient richness

Whilst the notion of "Persistence of Excitation" arose from the beginning of my P.h.D. (in particular, with the celebrated Willems Lemma [226]), I started investigating it only in the last year. By far, I think this concept one of the most interesting I have encountered, since it appears (for the first time in [16]) as an omnipresent condition that guarantees parameter convergence in data-driven techniques, and it is strictly related to the idea of *information* carried by a signal. Before going into formal details, it is useful to present the key ideas behind the concept of persistency of excitation that will be used in this thesis.

i) *Goal*: a PE signal guarantees exponential convergence of adaptive, identification and learning algorithms.

ii) *Movement*: a PE signal spans in time all the directions y of the space it lives in.

iii) *Energy*: the *movement* in each direction is done with a guaranteed minimum energy α.

iv) *Finite time*: both *movement* and *energy* are guaranteed to be completed in a time span of fixed length T.

v) *Persistency*: *movement, energy, finite time* hold uniformly in time, i.e., for all time intervals $[t, t + T]$.

To recap, there exists a window of maximum length T after which, starting from any time instant t, the signal has spanned all space directions y with a minimum energy $\sqrt{\alpha}$ (in an RMS sense). We propose also a preliminary notion for "Sufficient Richness", inspired by the reasonings in [41, 98, 162]. In real setups, it is often not possible to directly impose PE on the signals in which we are interested. In other words, we can arbitrarily design only *input* signals. This motivates the following informal characterization of sufficient richness.

i) *Goal*: given a certain dynamical system, a SR input signal $\mathtt{u}$ guarantees PE of a specific output signal $\mathtt{y}$.

ii) *Sufficiency*: we look for sufficient properties that may depend on the particular dynamical system and initial condition. We are not looking for equivalent formulations of PE.

iii) *Richness*: the “finite time", “energic" “movements" features required by a PE signal should be embedded into the input signal in some form.

Whilst it can be said there is agreement on the definition of PE (but for some minor details), SR seems a confused notion in the literature: whilst the underlying idea seems the one proposed here, it has never been formalized, and all the proposed definitions for SR are very specific to particular setups.
This Chapter is organized as follows. In Section II.1 we perform a historical review of the concepts of persistence of excitation and sufficient richness, providing also references for the existing results. Next, in Section II.2, we present the problem we want to solve and the considered framework. In Section II.3, we present the mathematical results which have been found, corroborating them with a numerical example showing their tightness and confuting a recently proposed conjecture. At last, in Section II.4, we consider a more restricted framework and show how to exploit the geometry of linear systems together with the properties of sinusoids to obtain tighter sufficient conditions for PE in linear systems. The proofs of the simplest results are given contextually, while the others can be found in Appendix V.4. An article containing the results of this chapter has been submitted as a full paper, [35].

II.1 An introduction to PE and SR

II.1.1 A historical review of the concept of PE

The first notion of PE

From the first work introducing persistency of excitation [16], definitions for "persistency of excitation", "persistence of excitation", "persistent excitation", "PE of order", "sufficient richness", "SR of order", "sufficient excitation" proliferated, being used somehow interchangeably in a first moment (between '70s and '80s) and then specializing for the continuous-time adaptive literature (PE, SR, SR of order) or discrete-time behavioral literature (PE, PE of order). A variety of definitions for both concepts have been proposed, moving away the involved concepts and the related literatures. Scope of this section is (i) to try, starting from the first introduction of these concepts, to provide the reader with a (philological) review of the main definitions that arose until today, (ii) to motivate the ideas behind the "Persistent Excitation" and "Sufficient Richness" definitions that are used in this thesis and (iii) to highlight the importance of the words in making the usage of mathematical concepts more or less clear and understandable. We start from the first definition (1965) for "persistency of excitation of order", given in [16] and reported here for the reader's convenience.

Definition II.1. *[PE of order, [16]] A bounded scalar signal* $\mathrm{u} : \mathbb{N}_1 \to \mathbb{R}$ *is said to be persistently exciting of order* m *if the limits*

$$\bar{u} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_t \qquad\qquad r_u(T) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_t u_{t+T} \tag{II.1}$$

exist and the following matrix is positive definite

$$R_u \coloneqq \begin{bmatrix} r_u(0) & r_u(-1) & \dots & r_u(-m) \\ r_u(1) & r_u(0) & \dots & \dots \\ \dots & \dots & \ddots & \dots \\ r_u(m) & \dots & \dots & r_u(0) \end{bmatrix}. \tag{II.2}$$

♔

Notice that $r_u(T)$ is an even function, namely, considering $u_t = 0$ for all $t < 1$, we get

$$r_u(-T) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_t u_{t-T} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1-T}^{N} u_{t+T} u_t = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_{t+T} u_t = r_u(T), \tag{II.3}$$

so the matrix R_u is symmetric. Later, we will see how $r_u(T)$ is closely related to the power spectrum of the signal u.

The first thing we need to specify is the framework considered in this article: in [16] the authors study an unknown SISO discrete-time linear system, with input signal u and output signal y, on which they design a maximum likelihood method for estimation of the input-output relation. The "persistency of excitation of order" is required on the *input* signal, and its goal is the *identification* of the input-output relation parameters. Notice furthermore that the authors quantify the "amount of excitation" that it is required by the algorithm. This first idea of persistency of excitation embeds somehow both the two concepts of PE and SR that have been introduced at the beginning of this chapter: the *goal* is the convergence of some identification algorithm (as the goal of our PE); however, the requirement is imposed on the input signal, which is available for the design, and it is not imposed on some internal signal (as it is for our PE).

Remark II.1. Considering the five informal "components of PE" that were given at the beginning of this chapter (*goal*, *movement*, *energy*, *finite time*, *persistency*), we see that this original definition includes all of them, though some are hidden. In fact, as made clearer later, *energy*, *finite time*, *persistency* are all implied by condition (II.2), and also the *goal* is the same of our definition. Being this notion for scalar signals, however, *movement* is the hardest to be seen. ◊

PE for UGAS of certain nonautonomous differential equations

We continue our review by passing to the year 1977, where in the two seminal articles [154, 155], the authors studied the uniform asymptotic stability of nonautonomous differential equations of the type

$$\dot{x} = -P(t)x, \qquad \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A(t) & -B(t)^\top \\ B(t) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{II.4}$$

where in the first equation it holds $P(t) = u(t)u(t)^\top \geq 0$ and in the second equation $A(t)+A(t)^\top < 0$, for all $t \geq 0$. These results are fundamental since nonautonomous differential equations of the type (II.4) arise in a huge number of adaptive control and recursive identification schemes such as adaptive observers [97, 98, 111], model reference adaptive controllers [6, 41, 42], gradient-based algorithms [32, 114], reinforcement learning [37, 173, 192, 193] and several others [9, 15, 16, 73, 83, 124, 190, 198, 223] (both discrete and continuous-time). From [154, 155] we now recall - for the reader convenience - the main results (in a shortened version).

Theorem II.1. *[[155], Thm. 1] Let* $\mathrm{u} : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n\times k}$ *be a piecewise continuous and bounded function. Then, the following are equivalent*

i) $\dot{x} = -u(t)u(t)^\top x$ *is uniformly asymptotically stable*

ii) there are real numbers $a > 0$, b *such that, if* $y \in \mathbb{R}^n$ *is a fixed unit vector, then*

$$\int_{t_0}^{t} y^\top u(\tau)u(\tau)^\top y d\tau \geq a(t - t_0) + b, \quad \forall t \geq t_0 \geq 0, \tag{II.5}$$

♕

Theorem II.2. *[[154], Thm. 1] Let* $A : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n\times n}$ *be a piecewise continuous and bounded function such that* $A(t) + A(t)^\top$ *is negative definite. Let* $B : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n\times m}$ *be a piecewise continuous and bounded* *function. Then, the system*

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A(t) & -B(t)^\top \\ B(t) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{II.6}$$

is uniformly asymptotically stable if and only if there are positive numbers T, ϵ and δ such that, given $t \geq 0$ and a unit vector $w \in \mathbb{R}^n$, there is a $t' \in [t, t+T]$ such that

$$\left| \int_{t'}^{t'+\delta} B(\tau)^\top w d\tau \right| \geq \epsilon. \tag{II.7}$$

♕

We refer the reader to the above mentioned articles and to [161, 162] for more insight into the slight differences between these PE expressions (which can be collapsed into the same condition under sufficient smoothness of the involved signals). These theorems - together with other important results from those years which aligned the discrete-time framework and found other important adaptive control or recursive identification schemes which resulted in systems in the form (II.4) - lead to a new notion of PE, which we could state as (being more restrictive to include many proposed definitions, see e.g. the books [8, 98, 162, 191])

Definition II.2. *The piecewise differentiable bounded signal* $\mathrm{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ *is persistently exciting if there exist $T, \alpha > 0$ such that, for all $t \in \mathbb{R}_{\geq 0}$, it holds*

$$\int_t^{t+T} w(\tau) w(\tau)^\top d\tau \geq \alpha I. \tag{II.8}$$

♔

For the discrete-time, the same type of analysis on the stability of equations of the form $x_{t+1} = F_t x_t$ arose, treated e.g. in [33], and from [32] we can infer a definition equivalent to Def. II.2 for the discrete-time domain

Definition II.3. *The bounded signal* $\mathrm{w} : \mathbb{N} \to \mathbb{R}^d$ *is persistently exciting if there exist $T, \alpha > 0$ such that, for all $t \in \mathbb{N}$, it holds*

$$\sum_t^{t+T} w_\tau w_\tau^\top \geq \alpha I. \tag{II.9}$$

♔

A few important observations are in order. First, with this definition there is no more a quantification of "how much" a signal is persistently exciting, namely a signal is PE and it is not "PE of order". Second, whilst the final goal of this new definition is still to guarantee the convergence of adaptive control (recursive identification) schemes, we notice that the PE condition has passed from the input signal u to a generic signal (from now on called "regressor") which plays the role of $u(t)$ or $B(t)$ in equations (II.4).

The shape of the regressor depends strongly on the considered closed-loop scheme, and to achieve this PE property via the design of an appropriate input signal becomes a completely different problem. Of course, one could also study when solutions of autonomous systems are PE depending on the initial conditions and the properties of the system, and this was done very recently in [168].
In 2001, the stability of nonautonomous differential equations more complex than those in (II.4) was considered. In fact, stability of equations in (II.4) is not only uniformly asymptotic, but it is also global and exponential thanks to the linearity of the equation. However, other frequently arising differential equations such as

$$\dot{x} = \begin{bmatrix} A & B\phi(t,x)^\top \\ -\phi(t,x)C^\top & 0 \end{bmatrix} x \tag{II.10}$$

or even more generic equations of the type

$$\dot{x} = F(x,t) \tag{II.11}$$

need more careful analysis. We refer the reader to the following works for other variations of this definition in more complicated frameworks [13, 14, 101, 130–134, 165, 171, 188].

Remark II.2. The definitions for PE presented here have the same *goal, movement, energy, finite-time, persistency* connotations presented at the beginning of this chapter. ◊

Remark II.3. [Excited or exciting?] In this thesis we prefer to say that a signal is "persistently excited" instead of the "persistently exciting" used in the literature. In its original definition, an input signal was exciting when it moved the measurements strongly enough to guarantees a successful identification. With Definition II.2, the property of PE has now been moved to a generic regressor to guarantee stability of certain equations. It is hard to say that the regressor is "exciting" for systems (II.4), since if it is exciting, the origin of these systems is UAS and the state $x(t)$ goes to zero. We prefer then to say that the regressor is just excited since it possesses the *movement, energy, persistence* characteristics introduced at the beginning of the chapter. ◊

Before addressing this problem, we make a short digression on frequency-oriented characterizations for PE that arose from the first years after 1965.

PE in the frequency domain

In 1971, we find in [122] an interesting characterization for PE in the frequency domain. Notice that the considered framework was the same of [16] (namely, a single-input single-output system), and also the goal of this "persistency excitation of order" was the same as in [16]: an input signal has to achieve convergence of recursive identification schemes. In fact, the authors consider a discrete-time scalar signal $\mathrm{u} : \mathbb{N} \to \mathbb{R}$ which should satisfy, similarly to II.1, the existence of the limits

$$\bar{u} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u_t \qquad r_u(T) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} (u_t - \bar{u})(u_{t+T} - \bar{u}). \tag{II.12}$$

Then, the authors implicitely define the function $s_u : [-\pi, \pi] \to \mathbb{R}$ such that

$$r_u(t) = \int_{-\pi}^{\pi} e^{it\omega} \mathrm{d}s_u(\omega) \qquad \forall t \in \mathbb{N}, \tag{II.13}$$

which is unique, symmetric and enjoys some smoothness properties (see [122] and references therein). We are now ready to present the characterization for "PE of order n" in the frequency domain.

Theorem II.3. *[[122], Thm. 1] A necessary and sufficient condition for* u *to be persistently exciting of order n (as per Def. II.1) is that* $\mathrm{supp}(s_u)$ *contains at least n points.* ♛

The most important takeaway from this theorem is that there is a one to one correspondence between the number of points $\omega \in [-\pi, \pi]$ in which $s_u(\omega)$ is not zero and the level of excitation of u. We will later see that s_u is closely related to the spectrum of u, so this means that frequency-domain characterizations may be very helpful in design purposes. Still, we are in a scalar framework, and we expect things to become more complicated as soon as the quantities $r_u(t)$ and $s_u(t)$ become matrix-valued.
As soon as new, multi-valued definitions for PE arose and lost their "order of excitation" quantification, new proposals for a frequency domain characterizations were made. In 1986, in the work [42] the authors considered a definition of PE similar to the one given in Definition II.2, and found an equivalent frequency-domain characterization. Before addressing it, we better formalize and complete the theoretical digression on the relation between the quantities $r_u(t)$ and $s_u(t)$ seen before. Consider a multi-valued continuous-time signal $\mathrm{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$. We define the quantity $R_w : \mathbb{R}_{\geq 0} \to \mathbb{R}^{d \times d}$

$$R_w(t) := \frac{1}{T} \lim_{T \to \infty} \int_0^T w(\tau) w(t+\tau)^\top \mathrm{d}\tau \tag{II.14}$$

as the "autocovariance" of w. Signals for which R_w exists are also called Wide-Sense-Stationary (WSS) [71, p. 388]. Notice that the autocovariance evaluated in zero, namely,

$$R_w(0) = \frac{1}{T} \lim_{T \to \infty} \int_0^T w(\tau) w(\tau)^\top \mathrm{d}\tau \tag{II.15}$$

closely recalls the PE expression in Definition II.2. The main difference is that the condition in II.2 is explicitely required to hold uniformly in time along moving time windows, while in (II.15) all the knowledge of the signal is somehow collapsed into one point of the autocovariance. Notice furthermore that the division by $\frac{1}{T}$ in (II.15) prevents vanishing signals to satisfy $R_w(0) > 0$. We are now ready to state the important result in [42, 191], which holds for a definition of PE similar (less smoothness is required) to II.2.

Lemma II.1. *[[42], Lemma 3.2] Suppose* $\mathrm{w} : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ *has autocovariance* $R_w(t)$*. Then* w *is PE iff* $R_w(0) > 0$. ♙

It was proved in the Wiener-Khinkin theorem [224, Ch. 2] that for WSS signals whose Fourier transform exists (which, in general, are not all WSS signals), their autocovariance R_w and their Power Spectral

Density (PSD) S_w form a Fourier-transform pair, namely,

$$R_w(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_w(\omega)e^{jt\omega}\mathrm{d}\omega, \tag{II.16}$$

from which

$$R_w(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_w(\omega)\mathrm{d}\omega. \tag{II.17}$$

For having PE signals, equation (II.17) together with the above Lemma II.1 clearly requires, as noted in [42], that $S_w(\omega)$ must be non-zero at least in d different points, namely, $\mathrm{supp}(S_w)$ should contain at least d points. Notice that for WSS scalar signals whose Fourier transform exists, the quantity s_u defined in (II.13) is exactly the PSD of the input signal.
In other words, we have now seen that quantity of points in the support of the PSD of a generic signal w has been used for different purposes, namely i) quantifying excitation levels of a scalar signal [122] and (ii) guaranteeing the “simple" PE of a multi-valued signal [42]. Notice furthermore that, by Lemma (II.1), it is not enough to contain enough points in the PSD to guarantee PE of the involved signal, since we need to require

$$R_w(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_w(\omega)\mathrm{d}\omega > 0. \tag{II.18}$$

In other words, if $S_w(\omega)$ is not scalar, we have to guarantee also some form of *independence* between the points $S_w(\omega_1), S_w(\omega_2), \ldots$, to guarantee that the integral is positive definite. An interesting work in which this is addressed is [41]; however, we refer the reader to Section II.4 for an overview of the results regarding spectral lines, since these will be used later.
We refer the reader to [71, 147, 224] for a more detailed discussion on autocovariance, power spectrum and WSS signals, and to [41, 42, 191] for their relation to with PE.

Willem’s PE

In 2005, 40 years after the first notion of PE was introduced, a new notion making use of the so called “Hankel matrix" was proposed for discrete-time systems [226]. Given a discrete-time signal $w_t \in \mathbb{R}^d$ in the interval $t = 1, \ldots, T$, we may build the associated Hankel matrix of depth L as

$$H_L^w := \begin{bmatrix} w_1 & w_2 & \ldots & w_{T-L+1} \\ w_2 & w_3 & \ldots & w_{T-L+2} \\ \ldots & \ldots & \ldots & \ldots \\ w_L & w_{L+1} & \ldots & w_T \end{bmatrix} \in \mathbb{R}^{dL\times(T-L+1)}. \tag{II.19}$$

Definition II.4. *[[226], Pag. 3] The sequence $w_1, \ldots, w_T$, with $w_t \in \mathbb{R}^d$ is said to be persistently exciting of order L if $\mathrm{rank}(H_L^w) = dL$.* ♔

The condition $\mathrm{rank}(H_L^w) = dL$ can be shown to be equivalent to the positive definiteness of following quantity:

$$\sum_{t=1}^{T-L+1} \begin{bmatrix} w_t \\ \dots \\ w_{t+L-1} \end{bmatrix} \begin{bmatrix} w_t \\ \dots \\ w_{t+L-1} \end{bmatrix}^\top > 0. \tag{II.20}$$

A few important observations are in order. First, we notice that the proposed definition takes into account the notion of excitation level in the signal: this is different from more recent continous-time or discrete-time characterizations and it is closer to the original notion given in [16]. The order of excitation is embedded in the number of time-shifts of the signals which one considers in building the matrix (II.20), and in [226][Cor. 1] this notion is used to derive how to obtain excited state-input trajectories from a sufficiently excited input (however, we deal with this in the next section). Second, notice that in this case the PE property is not a property which holds uniformly in time, since it is a property which a finite amount of data satisifies or not. An interesting question is then the following: why should we call this *persistency* if uniformity in time does not enter the definition? Third, by choosing $L = 1$, considering a signal on the whole time axis $\mathrm{w} : \mathbb{N} \to \mathbb{R}^d$ and asking for condition (II.20) to hold for each time window $\tau = t, \dots, t + T$ for all $t \in \mathbb{N}$, we obtain the condition

$$\sum_{\tau=t}^{T-L+1} w_\tau w_\tau^\top > 0 \qquad\qquad \forall t \in \mathbb{N}, \tag{II.21}$$

which is not equivalent to discrete-time definitions of PE II.3 for two reasons: (i) no uniform lower bound is implied by the full rankness of the Hankel matrix, and (ii) also the boundedness of the signal is not implied by this definition.

These features make this definition not well-suited for the analysis of recursive or asymptotic algorithms; however, this notion of persistency of excitation is more adapt to the cases where the number of data is finite (batch methods). Furthermore, the interpretation given by the authors in [226] of this kind of PE condition is also very different from the original one, since it is seen as the ability of gathered data to represent all system trajectories (of a LTI system). Still, also this notion of PE has been proved to be fundamental for the effectiveness an enormous amount of data-based techniques for learning and control (see [23, 62, 64, 127, 213] to cite a few), and very recently ($2022 - 2023$) "robustified" versions of this notion have been proposed [22, 55].

The Hankel matrix notation used for this definition prevented for long time the proposal of equivalent definitions for the continuous-time framework. Only recently some attempts have been made in the direction of fixing this problem. Among the others, in [128] (2022) the authors proposed to solve the problem introducing an Hankel matrix constructed from sampling the continuous-time signal. A different attempt was instead done in [176], where the following definition was proposed

Definition II.5. *[[176], Def. 1] Let $\mathbb{I} = (t_0, t_1) \subseteq \mathbb{R}$. $\mathrm{w} : \mathbb{I} \to \mathbb{R}^d$ is persistently exciting of order k if*

(a) w *is $(k-1)$-times continuously differentiable in $\mathbb{I}$,*

(b) *for every* $v \in \mathbb{R}^{kd}$*, it holds that*

$$v^\top \begin{bmatrix} w(t) \\ \vdots \\ \dot{w}^{(k-1)}(t) \end{bmatrix} = 0 \quad \forall t \in \mathbb{I} \implies v = 0. \tag{II.22}$$

♔

Similarly to Definition II.4, this notion of PE is not a "persistent" property, and thus it is less suited for recursive or asymptotic algorithms; however, it introduces a nice analogy between time shifts and derivatives. In particular, in this definition time derivatives are presented as the continuous-time counterpart of time shifts in quantifying the order of excitation of the signal.

Remark II.4. To recap, the biggest difference between these definitions and the one seen before stands in the concepts of *energy* (no minimum amount of energy is guaranteed in each direction) and *persistence*, since this property holds only on a finite time window and not persistently. Furthermore, the *goal* of these definitions is different, (and justifies the lack of *persistency*, term which should be more carefully used), since it regards the ability of a seguence of measurements of x_t, u_t to represent systems trajectories. ◇

Remark II.5. Since in this thesis we deal with both the two notions, we will use the term "persistence of excitation" for the Definitions II.14, II.13 presented at the beginning of this chapter, and we will use only the term "excitation" for the finite-time Definition II.4 and similar continuous-time counterparts. ◇

II.1.2 A historical review of the concept of SR

To the author's knowledge, one of the first works (1977) mentioning "sufficient richness" for some signal is [154]. In that work, the authors consider the stability of the differential equation (II.6), and require "sufficient richness" on the matrix signal $B(t)$. Although "sufficient richness" is not formally defined, it is interesting to notice that the required condition, namely the one given in Theorem II.2 is the same that will evolve (under sufficient smoothness of signals) into the "persistence of excitation" as per Definition II.2.

In the same year, we can find "sufficient richness" mentioned in [6, Pag. 5]. In the paper, the author is in the context of multivariable adaptive identification, in which the problem is the identification of rational transfer functions. The term "sufficiently rich" is required, generically, on the input signal $v^p(t)$ in order to achieve convergence of the identification scheme to the true transfer function parameters.

An interesting comment on the "richness" property was made in [232]:

The intuition behind the concept of persistent excitation is that the input should be (recurrently) "rich" enough to excite all the "modes" of the system which is being probed. Thus, from the system theoretic point of view, we can say that the input $u(\cdot)$ *is persistently exciting for the system* (1) *if the combined vector function* $[x'(\cdot), u'(\cdot)]$ *is persistently spanning.*

A few comments are in order. First, the authors use "rich" for an input signal that should impose excitation on some output signal $(x(t), u(t))$. Second, the authors recognize the need of saying *for which system* a certain input is persistently exciting. An interesting takeaway is thus the following: it makes sense to separate the spanning condition (in other works, called "mixing" [198, 223], which we called *movement*), from the means to achieve it. Furthermore, the spanning condition may be required on regressors which depend on the particular application: this means that a certain input may be "rich" for a system but not for another. We conclude from this that either we should quantify "how much" a signal is persistently exciting or we should use a different term ("sufficient richness") for what concerns input signals. Five years later, in [9] we read

We shall show that when the output of the unknown plant is required to follow a sufficiently rich reference trajectory, in the sense that conditions like the output-only conditions of Section 2*– see especially Theorem* 2.4 *are satisfied by the reference trajectory, then the estimate of the plant parameters will converge exponentially fast to the true value, as will the plant output to the reference trajectory.*

In this case, we are in the context of discrete-time systems and the authors are investigating the stability of identification and adaptive control algorithms. In particolar, in Section 2 the authors show how to impose PE (as per Def. II.2) on the regressor $x_t = (y_t, \ldots, y_{t-n+1}, u_t, \ldots, u_{t-m+1})$, with y_t and u_t input and output of a SISO system. The condition to which this quote refer is asked on the output trajectory u_t (or y_t) is the existence of some $N \in \mathbb{N}, \rho_1, \rho_2 > 0$ such that

$$\rho_1 I \geq \sum_{\tau=t}^{t+N} \begin{bmatrix} u_{\tau+n} \\ \cdots \\ u_{\tau-m+1} \end{bmatrix} \begin{bmatrix} u_{\tau+n} \\ \cdots \\ u_{\tau-m+1} \end{bmatrix}^\top \geq \rho_2 I \qquad \forall t \in \mathbb{N}. \tag{II.23}$$

An important observation is the following: this condition is very close to the original "persistency of exitation of order" (Def. II.1) and to Willems' one (Def. II.4). However, in this case, its role is to directly impose PE on x_t, and it is only indirectly looking for the convergence of the control/identification scheme. Furthermore, we can notice how the "order of excitation" required depends on the dimension of the considered system. Moving on, whilst in [41, Pag. 3] the term "sufficient richness" is used with the same meaning as PE (as per Def. II.2), three years later in 1986 the same authors [42] propose a different idea

The terms sufficiently rich (SR) and persistently exciting (PE) have been used somewhat interchangeably in the literature. It is proposed that PE refers to property (2.7) *for a vector of signals, and that sufficient richness be a property of the reference signal (scalar valued). A vector of signals is thus PE or not, but whether or not a reference signal is SR depends on the MRAC being studied. More specifically it depends only on the number of unknown parameters in the system, so it is proposed that a reference signal which results in a PE in an N-parameter MRAC be referred to as sufficiently rich of order N.*

With this notion of SR in mind, we are close to the definition proposed in this thesis. Whilst in this article the authors are focused on the MRAC field, they clearly distinguish between the idea of PE and the idea of guaranteeeing the PE via an input signal, and they recognize that different conditions may be required depending on the number of parameters to be identified. Another inportant step in this direction was made in [152]

Since the plant states x_k cannot be manipulated directly, but only via the plant inputs u_k,it is important to translate "persistency of excitation" conditions on the states of the plant, to"persistence of excitations" or "sufficiently rich" conditions on the plant inputs and noise.

where the idea of imposing PE via suitable properties is very clear. The condition found in [152] for achieving PE in a DT-LTI system leverages, once again, on a PE-like condition on the time-shifts of the input signals. In 1985 [17], this notion is formalized and a first proposal for a mathematical definition of this concept is given for discrete-time framework:

Definition II.6. *[[17] Def. 2] A sequence $x_t \in \mathbb{R}^n$ is said to be sufficiently rich (SR) of order m (in N steps), if there exists $N \in \mathbb{N}, \alpha > 0$ such that*

$$\sum_{\tau=t+1}^{t+N} \begin{bmatrix} x_{\tau+1} \\ \dots \\ x_{\tau+m} \end{bmatrix} \begin{bmatrix} x_{\tau+1} \\ \dots \\ x_{\tau+m} \end{bmatrix}^\top \geq \alpha I \qquad \forall t \in \mathbb{N}. \tag{II.24}$$

♔

This definition was also characterized in the frequency domain, at first for scalar signals

Lemma II.2. *[[17] Lemma 2, [98] Def. 5.2.3] If a scalar sequence u_t has an autocovariance, then u_t is SR of order n if and only if the spectral measure of u_t is not concentrated on $k < n$ points.* ♙

and later for multivariable ones [165].
Notice however, that whilst a frequency domain characterization for SR is available for both the continuous-time and discrete-time frameworks, an equivalent of Definition II.6 for the continuous time is still missing. An interesting attempt in this direction was done by [142] (1988).

Definition II.7. *[[142], Def 4.1] A signal $\mathrm{u} : \mathbb{R} \to \mathbb{R}^m$ is called sufficiently rich of order $(n_1, \dots, n_m/n)$ where $n \geq \max(n_i)$ if, for any $\gamma > 0$, there exist constants $t_1, \alpha > 0$, and $T > 0$ such that*

$$\frac{1}{T} \int_t^{t+T} \psi(\tau)\psi(\tau)^\top d\tau \geq \alpha I \qquad \forall t \geq t_1, \tag{II.25}$$

with

$$\psi(t) := \frac{1}{(s+\gamma)^{n-1}} \left[\begin{array}{cccc} 1 & 0 & \dots & 0 \\ s & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ s^{n_1-1} & 0 & \dots & \dots \\ 0 & 1 & \dots & \dots \\ \dots & s & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & s^{n_2-1} & \dots & 0 \\ --- & --- & --- & --- \\ \dots & \dots & \dots & 1 \\ \dots & \dots & \dots & s \\ \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & s^{n_m-1} \end{array}\right] u(t). \tag{II.26}$$

♔

Notice that the proposed definition involves again a PE-like condition on a certain function of the input signal; in this case however the authors do not take pure derivatives of the considered signal, but a filtered version of them. This was not the case in [97], where, studying the convergence guarantees of an adaptive observer with observability indices n_i (see [140]), the authors proposed the following definition for sufficient richness.

Definition II.8. *[[97], Pag. 5] Let n_0 denote $n_0 = \max_i n_i$, and p denote differential operator $p = d/dt$. Input $u(t)$ is said to be sufficiently rich if, for any t, $r((n+n_0-1))$ functions*

$$u_1, pu_1, \dots, p^{n+n_0-1}, u_2, \dots, p^{n+n_0-1}u_r \tag{II.27}$$

are linearly independent on the interval $[t_0, \infty)$, that is, for any $n+n_0-1$-th order polynomials

$$z_i(p) = z_{i,0} + \dots, z_{i,n+n_0-1}p^{n+n_0-1}, \quad i = 1, \dots, r, \tag{II.28}$$

$$\sum_{i=1}^{r} z_i(p)u_i(t) = 0 \tag{II.29}$$

does not hold. ♔

Although this article does not seem to be well-recognized, we anticipate that a very similar condition on the input derivatives will be found out in the next section to provide sufficient conditions for imposing PE on systems trajectories.

We conclude here our historical digression on the evolution of the concepts of PE and SR, which was made i) for motivating our choices for the Defs. II.14, II.13, II.15 and ii) as a contribution itself to the community working in this field. We refer the reader to, e.g., [19, 22, 78, 129, 160, 171, 176] for other more recent works in which the term "richness" of a signal is somehow involved; however, we stress the fact that these concepts were more used for online schemes, and terms such as "sufficient richness" are not in auge anymore.

II.2 Problem setup

II.2.1 Preliminaries: infinite-dimensional notation and operators

Since, as previously mentioned, we are dealing with properties which should hold uniformly in time for signals, it is useful to introduce some notation to lighten infinite-dimensional operators and results. We start by defining what a dynamical system is, for both the discrete-time and the continuous-time frameworks.

Definition II.9. *[BIBO Dynamical system - DT] A Bounded Input Bounded State Output discrete-time dynamical system is a map*

$$\begin{aligned} \sigma : \ell_\infty(\mathbb{R}^m) \times \mathbb{R}^n &\to \ell_\infty(\mathbb{R}^p) \\ \mathrm{u}, x_0 &\longmapsto \mathrm{y} \end{aligned} \tag{II.30}$$

that can be written in the form

$$\begin{aligned} x_{t+1} &= f(x_t, u_t), \qquad\qquad x_0 \in \mathbb{R}^n, \\ y_t &= h(x_t, u_t), \end{aligned} \tag{II.31}$$

where $u_t \in \mathbb{R}^m, x_t \in \mathbb{R}^n, y_t \in \mathbb{R}^p$ are the input, the state and the output at time t and f, h are the dynamics and the output map. ♔

Definition II.10. *[BIBO Dynamical system - CT] A Bounded Input Bounded Output continuous-time dynamical system is a map*

$$\begin{aligned} \sigma : C_b^\infty(\mathbb{R}^m) \times \mathbb{R}^n &\to C_b^\infty(\mathbb{R}^p) \\ \mathrm{u}, x(0) &\longmapsto \mathrm{y} \end{aligned} \tag{II.32}$$

that can be written in the form

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), \qquad\qquad x(0) \in \mathbb{R}^n, \\ y(t) &= h(x(t), u(t)), \end{aligned} \tag{II.33}$$

where $u(t) \in \mathbb{R}^m, x(t) \in \mathbb{R}^n, y(t) \in \mathbb{R}^p$ are the input, the state and the output at time t and f, h are the dynamics and the output map. ♔

Clearly, we could have defined dynamical system in a more general way, e.g. including unbounded signals, discontinuities, ecc.; we require instead smoothness and boundedness of the involved signals just for simplifying the exposition of the subsequently given result. Next, we introduce a list of useful operators to operate with discrete-time signals (shift operators) and continuous-time signals (derivative operators).

Definition II.11. *[Shift Operators] Given $k, d \in \mathbb{N}$ and the space $\ell_\infty(\mathbb{R}^d)$, we define the shift operator as*

$$\begin{aligned} \mathrm{q}^k : \ell_\infty(\mathbb{R}^d) &\to \ell_\infty(\mathbb{R}^d) \\ \mathrm{w}_t &\longmapsto \begin{cases} \mathrm{w}_{t-k} & t \geq k \\ 0 & t < k. \end{cases} \end{aligned} \tag{II.34}$$

Furthermore, we introduce the following multi-shift operator:

$$\begin{aligned} \mathrm{Q}^k : \ell_\infty(\mathbb{R}^d) &\to \ell_\infty(\mathbb{R}^{dk}) \\ \mathrm{w} &\longmapsto (\mathrm{q}^{k-1}\mathrm{w}, \ldots, \mathrm{q}^0\mathrm{w}). \end{aligned} \tag{II.35}$$

♔

Definition II.12. *[Derivative Operators] Given $k, d \in \mathbb{N}$ and the space $C_b^\infty(\mathbb{R}^d)$, we define the derivative operator as*

$$\begin{aligned} \mathrm{d}^k : C_b^\infty(\mathbb{R}^d) &\to C_b^\infty(\mathbb{R}^d) \\ \mathrm{w}(t) &\longmapsto \frac{d^k \mathrm{w}}{dt^k}(t). \end{aligned} \tag{II.36}$$

Furthermore, we introduce the following multi-derivative operator:

$$\begin{aligned} \mathrm{D}^k : C_b^\infty(\mathbb{R}^d) &\to C_b^\infty(\mathbb{R}^{dk}) \\ \mathrm{w} &\longmapsto (\mathrm{d}^{k-1}\mathrm{w}, \ldots, \mathrm{d}^0\mathrm{w}). \end{aligned} \tag{II.37}$$

♔

II.2.2 Persistent excitation and sufficient richness

We are now ready to state the definitions for Persistence of Excitation which will be used throughout the thesis.

Definition II.13. *[Discrete-Time PE] A signal $\mathrm{w} \in \ell_\infty(\mathbb{R}^d)$ is Persistently Excited (PE) if there exist positive scalars α, and T such that, for all $t \in \mathbb{N}$,*

$$\sum_{\tau=t}^{t+T} w_\tau w_\tau^\top \geq \alpha I. \tag{II.38}$$

♔

Definition II.14. *[Continuous-Time PE [7]] A signal* $\mathrm{w} \in C_b^\infty(\mathbb{R}^d)$ *is Persistently Excited (PE) if there exist positive scalars* α *and* T *such that, for all* $t \in \mathbb{R}_{\geq 0}$

$$\int_t^{t+T} w(\tau)w(\tau)^\top d\tau \geq \alpha I. \tag{II.39}$$

♔

As done before for the considered systems, the given definitions for PE are more conservative than other present in the literature, since in general only boundedness and some degree of smoothness are required, and this is done for simplicity reasons. See however Section II.1 for a more comprehensive list of possible definitions. Similarly as done in [161, Def. 2], we define the subsets

$$\begin{aligned} \Omega_d^{\mathrm{D}} &:= \{\mathrm{w} \in \ell_\infty(\mathbb{R}^d) : \mathrm{w} \text{ satistisfies (II.38)}\} \\ \Omega_d^{\mathrm{C}} &:= \{\mathrm{w} \in C_b^\infty(\mathbb{R}^d) : \mathrm{w} \text{ satistisfies (II.39)}\}, \end{aligned} \tag{II.40}$$

which are the set of all the PE signals in the discrete and continuous-time setting. The following lemma characterizes these sets.

Lemma II.3. Ω_d^{D} *(resp.,* Ω_d^{C}*) is an open cone in* $\ell_\infty(\mathbb{R}^d)$ *(resp.,* $C_b^\infty(\mathbb{R}^d)$*).* ♙

The proof of Lemma II.3 is in Appendix V.4.1.

Remark II.6. Notice that the property of being open sets is a reformulation of the robustness (in $\mathcal{L}_\infty$ sense) of the PE property given in [191, Lemma 6.1.2]. In other words, being Ω_d^{C} an open cone, for any given signal $\mathrm{w} \in \Omega_d^{\mathrm{C}}$ we can always find a sufficiently close signal w' such that $\mathrm{w}' \in \Omega_d^{\mathrm{C}}$. However, this robustness is not the same for every PE signal. ◇

Before addressing the main questions we want to solve, we formally define how the notion of "Sufficient Richness" is intended in this work. Consistently with the adaptive literature, we say an *input* signal sufficiently rich if, injected into a *dynamical system*, it guarantees persistence of excitation of the *output* signal we are interested in. Notice that imposing the PE to some output signal y via input u is a problem which is not in general guaranteed to be solvable. Furthermore, an input u may be sufficiently rich for a specific system but not for a different one. Specifically, the following two aspects determine the solvability (and the solution) of the problem:

i) the structural properties of the dynamical system σ we are considering (dynamics and output map)

ii) the set of initial conditions $x(0)$ of the systems in which we are interested.

These reasonings motivate the following definition for SR.

Definition II.15. *[Sufficient Richness] Given a discrete (resp. continuous)-time dynamical system* σ *and initial condition* x_0 *(resp.* $x(0)$*) we say that the input signal* $\mathrm{u} \in \ell_\infty(\mathbb{R}^m)$ *(resp.* $C_b^\infty(\mathbb{R}^m)$*) is SR for the* *tuple* (σ, x_0) *if*

$$\sigma(\mathrm{u}, x_0) \in \Omega_p^{\mathrm{D}} (\textit{resp. } \Omega_p^{\mathrm{C}}) \tag{II.41}$$

♔

We can define the set of all SR input signal for system σ and initial condition $x_0 \in \mathbb{R}^n$ as

$$\begin{aligned} \mathcal{D}_{\mathrm{SR}}(\sigma, x_0) &:= \{\mathrm{u} \in \ell_\infty(\mathbb{R}^m) : \sigma(\mathrm{u}, x_0) \in \Omega_p^{\mathrm{D}}\} \\ C_{\mathrm{SR}}(\sigma, x(0)) &:= \{\mathrm{u} \in C_b^\infty(\mathbb{R}^m) : \sigma(\mathrm{u}, x(0)) \in \Omega_p^{\mathrm{C}}\}. \end{aligned} \tag{II.42}$$

Notice that any characterization of the sets $C_{\mathrm{SR}}(\cdot), \mathcal{D}_{\mathrm{SR}}(\cdot)$ for the SR property as per Definition II.15 requires explicitely the considered system (dynamics and output map) and initial conditions for which the characterization of SR is valid.

II.2.3 Problem statement: sufficient richness in LTI systems

Consider a discrete or continuous-time linear time invariant system in the form

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \qquad & \dot{x}(t) &= Ax(t) + Bu(t), \\ y_t &= Cx_t + Du_t & y(t) &= Cx(t) + Du(t) \end{aligned} \tag{II.43}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $y \in \mathbb{R}^p$ is the output, while $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are the system matrices. The aim of this chapter is twofold, namely

i) given meaningful classes of systems in form (II.43) and initial conditions, find an explicit characterization of the sets of SR inputs $C_{\mathrm{SR}}(\cdot), \mathcal{D}_{\mathrm{SR}}(\cdot)$

ii) obtain a characterization of $C_{\mathrm{SR}}(\cdot), \mathcal{D}_{\mathrm{SR}}(\cdot)$ which underlines the analogies between the discrete-time and the continuous-time framework.

Notice that while the specific framework proposed in this thesis investigates these properties only for linear systems, the proposed questions and concept of persistent excitation and sufficient richness are exportable also to nonlinear frameworks. An example of an apparently different setup is the training of neural networks: it was demostrated [84] that richness of the training data ensures well posedness of the estimation problem. Very recently [159, 200], it has been shown that richness conditions on both the input data and the loss functions can be found so that the resulting trained neural network is more robust to adversarial attacks [50, 96].
Concerning the first objective, since we are not interested in solving a control problem (and having defined PE only for bounded signals), we consider only asymptotically stable systems, i.e., systems in which A is Schur (resp., Hurwitz). Intuitively, this means that the answer we are looking for will not depend on the chosen initial condition. Next, given a stable system it is known [161, Lemma 5], [98] that a necessary condition for having PE state trajectories is that (A, B) must be reachable. We formalize

these classes of systems with the following notation

$$\begin{aligned} \mathbb{L}^{\mathrm{D}} &:= \{\sigma \text{ in form (II.43)} : A \text{ Shur}, (A, B) \text{ reachable}, \dim(x) = n, \dim(u) = m\} \\ \mathbb{L}^{\mathrm{C}} &:= \{\sigma \text{ in form (II.43)} : A \text{ Hurwitz}, (A, B) \text{ reachable}, \dim(x) = n, \dim(u) = m\}. \end{aligned} \tag{II.44}$$

Next, since one may be interested in considering different output maps depending on the uncertainties in the matrices A, B, we define system classes

$$\begin{aligned} &\mathbb{L}^{\mathrm{D}}_{\mathrm{x}} := \{\sigma \in \mathbb{L}^{\mathrm{D}} : C = I_n, D = 0_m\}, && \mathbb{L}^{\mathrm{C}}_{\mathrm{x}} := \{\sigma \in \mathbb{L}^{\mathrm{C}} : C = I_n, D = 0_m\}, \\ &\mathbb{L}^{\mathrm{D}}_{\mathrm{xu}} := \left\{\sigma \in \mathbb{L}^{\mathrm{D}} : C = \begin{bmatrix} I_n \\ 0_{m\times n} \end{bmatrix}, D = \begin{bmatrix} 0_{n\times m} \\ I_m \end{bmatrix}\right\}, && \mathbb{L}^{\mathrm{C}}_{\mathrm{xu}} := \left\{\sigma \in \mathbb{L}^{\mathrm{C}} : C = \begin{bmatrix} I_n \\ 0_{m\times n} \end{bmatrix}, D = \begin{bmatrix} 0_{n\times m} \\ I_m \end{bmatrix}\right\}. \end{aligned} \tag{II.45}$$

At last, we are interested in separating the results for single-input and multi-input systems, for which we introduce the notation

$$\begin{aligned} \mathbb{L}^{\mathrm{D}}_{\mathrm{x},1} &:= \{\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}} : m = 1\}, & \mathbb{L}^{\mathrm{C}}_{\mathrm{x},1} &:= \{\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}} : m = 1\}, \\ \mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1} &:= \{\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}} : m > 1\}, & \mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1} &:= \{\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}} : m > 1\}. \end{aligned} \tag{II.46}$$

Concerning the second objective, namely the problem of unifying the continuous and discrete-time, our major concern is to avoid incompatible descriptions. An example of incompatible description is the Hankel matrix notation, since it is not well-suited for the continuous-time framework. The idea is to rely on the correspondence between time shifts in discrete-time and derivatives in continuous-time, exploiting the infinite-dimensional operators defined in II.11, II.12.

II.3 Necessary and sufficient conditions for SR in LTI systems

II.3.1 Necessary conditions

Discrete-time

We start by consider the problem of giving necessary results for sufficient richness in the case of discrete-time systems. As previously mentioned, these results are new, since to the author's knowledge necessary conditions have been recently found only for single-input systems [143] in the framework of finite-time excitation. Before presenting the theorem, we introduce a notion to characterize signals that persistently span only subspaces of the space they live in.

Definition II.16. *[Discrete-Time Partial PE] A signal* $\mathrm{w} \in \ell_\infty(\mathbb{R}^d)$ *is Partially Persistently Exciting (PPE) of degree* $d' < d$ *(and write* $\mathrm{w} \in \Omega^{\mathrm{D}}_{d,d'}$*) if there exists a canonical projection* $P : \mathbb{R}^d \to \mathbb{R}^{d'}$ *such that* $P\mathrm{w} \in \Omega^{\mathrm{D}}_{d'}$. ♔

Remark II.7. An almost equivalent notion of PPE was introduced for the continuous-time in [161, Def.

3], [162, Def. 6.3], from which we also adopt the simplified notation $\Omega^{\mathrm{D}}_{d,d'}$. ◇

Introducing this notion is important since, in the following, we will prove that the only necessary condition for achieving PE trajectories in multi-input systems is a PPE condition on $\mathrm{Q}^n(\mathrm{u})$.
It turns out that if the degree of PPE of signals in the form $\mathrm{Q}^n(\mathrm{u})$ (namely, stacks of time-shifts of the same signal u) is sufficiently small, then it never increases when increasing the number of considered time shifts. The intuition behind this property is that shifts a certain time window of the same signal are not completely independent. Consider e.g., the sequence of scalars $\{w_0, w_1, \dots, w_n\}$ and its time shift $\{w_1, w_2, \dots, w_{n+1}\}$. Notice that the two sequencies share $n-1$ elements, and the same holds true also considering the successive window $\{w_2, w_3, \dots, w_{n+2}\}$. This constraint can be shown to bind the number of directions possibly spanned by the moving window $\{w_t, w_{t+1}, \dots, w_{t+n-1}\}$ depending on the number of directions persistently spanned by the original signal w. This is formalized in the following lemma.

Lemma II.4. *Let* $\mathrm{w} \in \ell_\infty(\mathbb{R}^d)$*. Let* $d' \in \mathbb{N}$ *be the biggest natural for which* $\mathrm{Q}^n(\mathrm{w}) \in \Omega^{\mathrm{D}}_{nd,d'}$*. If* $d' \leq d(n-1)$*, then for all* $k \geq n$*, the biggest natural for which* $\mathrm{Q}^k(\mathrm{w}) \in \Omega^{\mathrm{D}}_{kd,d''}$ *holds is such that* $d'' \leq d'$*.* ♙

The proof of Lemma II.4 is provided in Appendix V.4.2.
We are now ready to state a necessary condition to obtain PE in discrete-time LTI multivariable systems.

Theorem II.4. *[Necessary condition for DT systems] Let* $\mathrm{u} \in \ell_\infty(\mathbb{R}^m)$*. For all initial conditions* $x_0 \in \mathbb{R}^n$*,*

i) If $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}}$ *and* $\sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n$*, then* $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm,n}$*.*

ii) If $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{xu}}$ *and* $\sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_{n+m}$*, then* $\mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(n+1)m,n+m}$*.*

♕

The proof of Theorem II.4 is provided in Appendix V.4.7.

Remark II.8. Notice that, in the case of single-input systems, this result collapses into a "full" persistency of excitation requirement on $\mathrm{Q}^n(\mathrm{u})$ and $\mathrm{Q}^{n+1}(\mathrm{u})$, namely, for all $x_0 \in \mathbb{R}^n$

$$\begin{aligned} \sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}}, \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n &\implies \mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_n, \\ \sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{xu}}, \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_{n+1} &\implies \mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{n+1}. \end{aligned} \tag{II.47}$$

◇

We sketch here the proof of Theorem II.4, since it provides arguments which are perfectly repeatable in the continuous-time framework and gives insight on why PPE of $\mathrm{Q}^n(\mathrm{u})$ is a necessary condition. This result is proven by contraposition, namely, we want show that if the input signal u is not PPE of a sufficiently high degree then the resulting x is not PE. We follow these steps

1. In asymptotically stable systems, x can be approximated arbitrarily well by a linear combination of a large enough number $k \in \mathbb{N}$ of time-shifts of the input, $\mathrm{Q}^k(\mathrm{u})$, namely, $\mathrm{x} \approx K\mathrm{Q}^k(\mathrm{u})$.

2. If $Q^n(u)$ is PPE of degree at most $n-1$, then for any $k \geq n$, the PPE of $Q^k(u)$ does not increase (using Lemma II.4) by picking any $k > n$.

3. If $Q^k(u)$ is PPE of degree at most $n-1$, then x cannot be PE, since it is a linear combination of a signal persistently spanning at most $n-1$ directions.

Remark II.9. For this necessary condition, we considered A Schur. Notice that in this case this assumption is required not only to guarantee boundedness of the state, but also to approximate the state as a linear combination only of a finite number of inputs. ◊

Remark II.10. The arguments used for these proofs give insight into how one may specifically design inputs to achieve PE trajectories (provided that the matrices A, B were known). In particular, one may leverage the approximation $x \approx KQ^k(u)$ (for sufficiently high k) and the knowledge of K (given in (V.108)) to guarantee the energy of the input is used in the "best way" possible, according to some performance index. As an example, one may desire to maximize α (from Definition II.13) for a fixed input energy. Notice furthermore that, in general, solving this task requires a good "planning" of the inputs, namely, one should solve the problem for a sufficiently high k and thus finding the input for several time instants. ◊

Continuous-time

Before presenting the results for the continuous-time framework, we need to consider some problems which arise when applying Definition II.14 (that are not present in the discrete-time framework). With the following lemma, we leverage the smoothness properties of the considered signals to guarantee that a non-PE signal is arbitrarily small for arbitrarily long periods along certain directions.

Lemma II.5. *Let* $w \in C_b^\infty(\mathbb{R}^d)$*. If* $w \notin \Omega_d^C$*, then for any* $T, \epsilon > 0$ *there exist* $t > 0$, $z \in \mathbb{R}^d$ *such that, for all* $\tau \in [t, t+T]$*, it holds that*

$$|z^\top w(\tau)| \leq \epsilon. \tag{II.48}$$

♙

The proof of Lemma II.5 is provided in Appendix V.4.3.
Notice that whilst the above result is immediate in the discrete-time framework, in the continuous time it is not necessarily true - without assuming some degree of smoothness of the signal.
Next, in order to state necessary results for PE in continuous-time systems, we introduce the notion of PPE also for this framework.

Definition II.17. *[Continuous-Time Partial PE] A signal* $w \in C^d$ *is Partially Persistently Exciting (PPE) of degree* $d' \leq d$ *(and write* $w \in \Omega_{d,d'}^C$*) if there exists a canonical projection* $P : \mathbb{R}^d \to \mathbb{R}^{d'}$ *such that* $Pw \in \Omega_{d'}^C$*.* ♔

To get the necessary conditions of PE in the continuous-time framework by following the same steps of the discrete-time one, we need an analogous of Lemma II.4.

Lemma II.6. *Let* $\mathrm{w} \in C_b^\infty(\mathbb{R}^d)$. *Let* $d' \in \mathbb{N}$ *be the biggest natural for which* $\mathrm{W} := \mathrm{D}^n(\mathrm{w}) \in \Omega^{\mathrm{C}}_{nd,d'}$. *If* $d' \leq d(n-1)$, *then for any* $0 < \bar{T} \leq T$, $\epsilon > 0$ *there exist* $t > 0, \bar{N} \in \mathbb{N}$, $d'' \leq d'$, *and* $G \in \mathbb{R}^{nd(N+1)\times d''}$ *such that, for all* $N \geq \bar{N}$,

$$\begin{bmatrix} W(\tau) \\ W\left(\tau - \frac{\bar{T}}{N}\right) \\ \dots \\ W(\tau - \bar{T}) \end{bmatrix} = G\lambda(\tau) + \tilde{W}(\tau), \quad \forall \tau \in [t, t+T] \tag{II.49}$$

for some $\lambda(\tau) \in \mathbb{R}^{d''}$ *and* $\tilde{W}(\tau) \in \mathbb{R}^{nd(N+1)}$ *such that* $|\tilde{W}(\tau)| \leq \epsilon$. ♙

The proof of Lemma II.6 is provided in Appendix V.4.4.
Unfortunately, the above lemma is not as elegant as its discrete-time counterpart (Lemma II.4). However, the intuition is the same: under a certain threshold of PPE, and sampling with a sufficiently small sample time, the number of directions spanned persistently by a certain stack of time shifts of $\mathrm{D}^n(\mathrm{w})$ does not increase if the time shifts are increased. Given Lemma II.6, and following the same steps of the discrete-time case, we obtain the following necessary conditions (the proof is in the Appendix).

Theorem II.5. *[Necessary condition for CT systems] Let* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$. *For all* $x(0) \in \mathbb{R}^n$,

i) If $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}$ *and* $\sigma(\mathrm{u}, x(0)) \in \Omega^{\mathrm{C}}_n$, *then* $\mathrm{D}^n(\mathrm{u}) \in \Omega^{\mathrm{C}}_{nm,n}$.

ii) If $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{xu}}$ *and* $\sigma(\mathrm{u}, x(0)) \in \Omega^{\mathrm{C}}_{n+m}$, *then* $\mathrm{D}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{C}}_{(n+1)m,n+m}$.

♕

The proof of Theorem II.5 is provided in Appendix V.4.9.

Remark II.11. As in the case of discrete-time systems, notice that when $m = 1$ this result collapses into a "full" persistency of excitation requirement on $\mathrm{D}^n(\mathrm{u})$ and $\mathrm{D}^{n+1}(\mathrm{u})$, namely, for all $x(0) \in \mathbb{R}^n$, it holds that

$$\begin{aligned} \sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}, \sigma(\mathrm{u}, x(0)) \in \Omega^{\mathrm{C}}_n &\implies \mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{C}}_n, \\ \sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{xu}}, \sigma(\mathrm{u}, x(0)) \in \Omega^{\mathrm{C}}_{n+1} &\implies \mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{C}}_{n+1}. \end{aligned} \tag{II.50}$$

◊

II.3.2 Sufficient conditions

Discrete-time

We proceed by finding sufficient conditions for the case of discrete-time systems. Whilst the following result is known in the literature [17, 86, 152], it is worth noting that the proof given here is new, provides insight on the obtained result, and its principles are completely repeatable in the continuous-time framework.

Theorem II.6. *[Sufficient condition for DT systems] Let* $\mathrm{u} \in \ell_\infty(\mathbb{R}^m)$*. For all* $x_0 \in \mathbb{R}^n$*,*

i) If $\sigma \in \mathbb{L}_{\mathrm{x}}^{\mathrm{D}}$ *and* $\mathrm{Q}^n(\mathrm{u}) \in \Omega_{nm}^{\mathrm{D}}$*, then* $\sigma(\mathrm{u}, x_0) \in \Omega_n^{\mathrm{D}}$*.*

ii) If $\sigma \in \mathbb{L}_{\mathrm{xu}}^{\mathrm{D}}$ *and* $\mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega_{(n+1)m}^{\mathrm{D}}$*, then* $\sigma(\mathrm{u}, x_0) \in \Omega_{n+m}^{\mathrm{D}}$*.*

♛

The proof of Theorem II.6 is provided in Appendix V.4.6.
Though the proof is given in the Appendix directly for the multi-input case, in order to have an easier understanding of the underlying arguments, consider the single-input case. We prove the above result by contraposition following these three arguments:

1. Any solution $\mathrm{x} = \sigma(\mathrm{u}, x(0)) \in \ell_\infty(\mathbb{R}^n)$ of system σ which is not PE constrains the (scalar) system input to be arbitrarily close to a feedback gain for arbitrarily long intervals, namely, $u_t \approx Kx_t$.
2. Such an input renders the system an autonomous system for arbitrarily long intervals. This means that time shifts $x_t, x_{t-1}, \ldots, x_{t-n+1}$ of the state can be written as a linear function of the state x_{t-n+1}.
3. Time shifts of inputs are given by state feedbacks of time shifts of the state $((u_t, \ldots, u_{t-n+1}) \approx (Kx_t, \ldots, Kx_{t-n+1}))$, which are a linear function of the state x_{t-n+1}. This means that if x is not PE, $\mathrm{Q}^n(\mathrm{u})$ is not PE.

The principles behind the proof for multi-input systems are the same as those for single-input systems, with the differences explained in the following steps.

1. Any solution $\mathrm{x} = \sigma(\mathrm{u}, x(0)) \in \ell_\infty(\mathbb{R}^n)$ of system σ which is not PE does not fully constrain the input signal to be arbitrarily close to a feedback gain of the state. In particular, we obtain an input which can be written as $u_t \approx Kx_t + v_t$.
2. v_t is constrained, however, to span in a space which is at most $m-1$ dimensional.
3. It can be checked that in the end this means $\mathrm{Q}^n(\mathrm{u})$ is not PE, since it is a function of x and $\mathrm{Q}^n(\mathrm{v})$ which, altogether, span persistently a subspace of $\mathbb{R}^{nm}$ which is at most $nm-1$ dimensional.

Remark II.12. For this sufficient condition we considered A Schur. However, we stress that we require it only to fulfill an upper bound on the state evolution. In order to guarantee only a spanning condition with a lower bound (see, e.g., [82, Def. 3]), it is not required A to be Schur. Another application of this result in absence of a stable system is when the gathered data is finite (Willems' PE): in those cases, it is possible to recover the same results as those in [226]. ◊

Continuous-time

In order to repeat the proof given for the sufficient results in the discrete-time domain, we need now to find a relation between the PE of a signal and the PE of its time derivative.

Lemma II.7. *Let* $\mathrm{w} \in C_b^\infty(\mathbb{R}^d)$*. If* $\mathrm{w} \notin \Omega_d^{\mathrm{C}}$*, then for any* $T, \epsilon > 0$ *there exist* $t > 0$, $z \in \mathbb{R}^d$ *such that*

$$|z^\top w(\tau)| \le \epsilon, \qquad\qquad |z^\top \dot{w}(\tau)| \le \epsilon, \tag{II.51}$$

for all $\tau \in [t, t+T]$. ♙

The proof of Lemma II.7 is provided in Appendix V.4.5.
Once Lemmas II.5 and II.7 are established, we can repeat the proof of Theorem II.6 as it is to derive an equivalent result for continuous-time systems (the proof is in the Appendix).

Theorem II.7. *[Sufficient condition for CT systems] Let* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$*. For all* $x(0) \in \mathbb{R}^n$*,*

i) If $\sigma \in \mathbb{L}_{\mathrm{x}}^{\mathrm{C}}$ *and* $\mathrm{D}^n(\mathrm{u}) \in \Omega_{nm}^{\mathrm{C}}$*, then* $\sigma(u, x(0)) \in \Omega_n^{\mathrm{C}}$*.*

ii) If $\sigma \in \mathbb{L}_{\mathrm{xu}}^{\mathrm{C}}$ *and* $\mathrm{D}^{n+1}(\mathrm{u}) \in \Omega_{(n+1)m}^{\mathrm{C}}$*, then* $\sigma(u, x(0)) \in \Omega_{n+m}^{\mathrm{C}}$*.*

♕

The proof of Theorem II.7 is provided in Appendix V.4.8.

II.3.3 The cone of SR signals

Having collected all these necessary and sufficient conditions in the above theorems, we are now interested in understanding the shape of the set of the signals which are SR *for all* stable systems sharing certain structural properties, namely, the input and state dimension (since, in general, we are interested in PE when the system matrices A, B are not completely known). At first, consider the case of single input systems of fixed dimension $\mathbb{L}_{\mathrm{x},1}^{\mathrm{D}}, \mathbb{L}_{\mathrm{x},1}^{\mathrm{C}}$ defined in (II.46). We are interested in the sets

$$\mathcal{D}_{\mathrm{SR}}(\mathbb{L}_{\mathrm{x},1}^{\mathrm{D}}) \coloneqq \bigcap_{\substack{x_0 \in \mathbb{R}^n \\ \sigma \in \mathbb{L}_{\mathrm{x},1}^{\mathrm{D}}}} \mathcal{D}_{\mathrm{SR}}(\sigma, x_0), \qquad \mathcal{C}_{\mathrm{SR}}(\mathbb{L}_{\mathrm{x},1}^{\mathrm{C}}) \coloneqq \bigcap_{\substack{x(0) \in \mathbb{R}^n \\ \sigma \in \mathbb{L}_{\mathrm{x},1}^{\mathrm{C}}}} \mathcal{C}_{\mathrm{SR}}(\sigma, x_0), \tag{II.52}$$

where $\mathcal{D}_{\mathrm{SR}}(\sigma, x_0), \mathcal{C}_{\mathrm{SR}}(\sigma, x_0)$ are defined in (II.42).

Lemma II.8. *Given system classes* $\mathbb{L}_{\mathrm{x},1}^{\mathrm{D}}, \mathbb{L}_{\mathrm{x},1}^{\mathrm{C}}$ *in* (II.46)*, the sets of sufficiently rich inputs* (II.52) *are given by*

$$\begin{aligned} \mathcal{D}_{\mathrm{SR}}(\mathbb{L}_{\mathrm{x},1}^{\mathrm{D}}) &= \{\mathrm{u} \in \ell_\infty(\mathbb{R}) : \mathrm{Q}^n(\mathrm{u}) \in \Omega_n^{\mathrm{D}}\}, \\ \mathcal{C}_{\mathrm{SR}}(\mathbb{L}_{\mathrm{x},1}^{\mathrm{C}}) &= \{\mathrm{u} \in C_b^\infty(\mathbb{R}) : \mathrm{D}^n(\mathrm{u}) \in \Omega_n^{\mathrm{C}}\}. \end{aligned} \tag{II.53}$$

♙

Proof. It is sufficient to apply Theorems II.6, II.4, II.7, II.5 to notice that for all $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x},1}$ and $x_0 \in \mathbb{R}^n$, it holds

$$Q^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_n \iff \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n. \tag{II.54}$$

※

Remark II.13. Notice sets $\mathcal{D}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{D}}_{\mathrm{x},1}), \mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},1})$ are open cones in $\ell_\infty(\mathbb{R}), C^\infty_b(\mathbb{R})$ (the proof is the same as the one for Lemma II.3). As per per Remark II.6, this means that sufficiently small perturbations of SR signals are still SR signals. ◊

Notice that this characterization is complete, namely, we have found *all* the input signals which are SR for systems in $\mathbb{L}^{\mathrm{D}}_{\mathrm{x}}, \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}$. Indeed, the single-input case is simplified by the fact that, as stated in [161, Thm. 1], any signal which is SR for a certain $\sigma_1 \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}}$ must be SR also for $\sigma_2 \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}}$ (which means that in (II.52) we intersect always the same set). The same is not true for multi-input systems, and in this case a complete characterization of the inputs which are SR for *all* the stable systems sharing input and state dimension seems not easy to obtain. Considering the classes $\mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}, \mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}$ defined in (II.46), we are interested in the sets

$$\mathcal{D}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}) := \bigcap_{\substack{x_0 \in \mathbb{R}^n \\ \sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}}} \mathcal{D}_{\mathrm{SR}}(\sigma, x_0), \qquad \mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}) := \bigcap_{\substack{x(0) \in \mathbb{R}^n \\ \sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}}} \mathcal{C}_{\mathrm{SR}}(\sigma, x_0), \tag{II.55}$$

where $\mathcal{D}_{\mathrm{SR}}(\sigma, x_0), \mathcal{C}_{\mathrm{SR}}(\sigma, x_0)$ are given in (II.42).

Lemma II.9. *Given system classes* $\mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}, \mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}$ *in* (II.46)*, the sets of sufficiently rich input* (II.55) *satisfy*

$$\begin{aligned} \{\mathrm{u} : Q^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm}\} \subset \mathcal{D}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}) \subset \{\mathrm{u} : Q^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm,n}\} \\ \{\mathrm{u} : D^n(\mathrm{u}) \in \Omega^{\mathrm{C}}_{nm}\} \subset \mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}) \subset \{\mathrm{u} : D^n(\mathrm{u}) \in \Omega^{\mathrm{C}}_{nm,n}\}. \end{aligned} \tag{II.56}$$

♗

Proof. It is sufficient to apply Theorems II.6, II.4, II.7, II.5 to notice that for all $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}$ and $x_0 \in \mathbb{R}^n$, it holds

$$\begin{aligned} Q^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm} &\implies \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n \\ Q^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm,n} &\impliedby \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n. \end{aligned} \tag{II.57}$$

※

The difficulties of obtaining a complete characterization of the sets $\mathcal{D}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}), \mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1})$ are introduced by the fact that in (II.55) we are intersecting different sets (in other words, there exists inputs which are SR for a certain $\sigma_1 \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}$ but not for $\sigma_2 \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x},>1}$). The reason of this needs not to be searched into the initial condition, but into the system matrices A, B and how each input enters into the system. To corroborate this statement notice that, by [228, Lemma 2.2], if (A, B) is reachable, then it is possible to build a feedback gain F such that $(A + BF, b)$ is reachable, with $b \in \mathrm{im}(B)$. With such a reshape of the system, the conditions for single input systems hold, which are independent on A, b, F and x_0 (so, it is only the *structure* of the system that influences the ability to obtain PE trajectories).

II.3.4 Tightness of the sufficient and necessary condition

We show the tightness of the presented result by providing examples which show that the obtained results are tight and cannot be improved without considering more specific classes of systems. Notice that in Definition II.13, PE is an infinite-dimensional notion, and cannot be properly checked via a numerical example. To deal with this problem, in both the following examples we inject periodic inputs and gather sufficiently large data batches to obtain a representative behavior for the whole infinite-dimensional time window.

Sufficient condition

We pick here a condition about u weaker than the one claimed to be sufficient by Theorem II.6, and we show it is not enough to guarantee PE of (x, u). We consider a discrete-time LTI system in form (II.43), with $n = 7, m = 3$ and matrices

$$A := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0.024 & -0.26 & 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.21 & -1.07 & 1.8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.8 \end{bmatrix},$$
$$B := \begin{bmatrix} 0 & 2 & 1 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0.4 & 7 & 4 & 0 & 0 \\ 5 & 2 & 0.9 & 4 & 6 & 2 & 1 \end{bmatrix}^\top. \tag{II.58}$$

It can be verified that the pair (A, B) is reachable and A is Schur. We want to show that $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm}$ (instead of $\mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(n+1)m}$) does not ensure $(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{C}}_{n+m}$. By choosing initial condition $x_0 = 0$, and input described by the dynamics $u_{t+1} = K_x x_t + K_u u_t + v_t^1 + v_t^2$, with $u_0 = 0$ and

$$v_t^1 = \begin{bmatrix} -0.4082 \\ 0.9082 \\ 0.0918 \end{bmatrix} (\sin(t) + \sin(2t) + \sin(3t) + \sin(4t)),$$

$$v_t^2 = \begin{bmatrix} 0.4082 \\ 0.0918 \\ 0.9082 \end{bmatrix} (\sin(5t) + \sin(6t) + \sin(7t) + \sin(8t)),$$

$$K_x = 10^{-3} \begin{bmatrix} -8 & 8.67 & -300 & -70 & 356.7 & -600 & -266.7 \\ -4 & 4.33 & -150 & -350 & 178.3 & -300 & -133.3 \\ 4 & -4.33 & 150 & 350 & -178.3 & 300 & 133.3 \end{bmatrix},$$

$$K_u = \begin{bmatrix} -0.6667 & -0.1333 & -1.3 \\ -0.3333 & -0.0667 & -0.65 \\ 0.3333 & 0.0667 & 0.65 \end{bmatrix},$$

simulating for $t = 1, \dots, 1000$, it can be checked that the input verifies $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm}$. However the resulting state-input trajectory (x, u) is not PE, so we have shown that

$$\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm} \not\Rightarrow \quad (\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}. \tag{II.59}$$

In Figure II.1, we plot the functions

$$\begin{aligned} r_1(\mathrm{x}, \mathrm{u}, T) &:= \operatorname{rank}\left(\sum_{t=0}^{T} (x_t, u_t)(x_t, u_t)^\top\right) \\ r_n(\mathrm{u}, T) &:= \operatorname{rank}\left(\sum_{t=0}^{T} \mathrm{Q}^n(\mathrm{u})_t \mathrm{Q}^n(\mathrm{u})_t^\top\right) \end{aligned} \tag{II.60}$$

as the time window T of considered samples increases.

Remark II.14. This example is a counterexample also for the sufficiency conjecture in [143, Pag. 4], which we may state (with some abuse of notation) as

$$\mathrm{Q}^{\nu+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(\nu+1)m} \implies (\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}, \tag{II.61}$$

where ν is the controllability index [228, Pag. 121] of the pair (A, B). In this case, the controllability index of the pair (II.58) is $\nu = 3$, and $\nu + 1 < n$, so we may write $\mathrm{Q}^{\nu+1}(\mathrm{u}) = P\mathrm{Q}^n(\mathrm{u})$ for some full row rank matrix P, and $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm}$ ensures that our input verifies also $\mathrm{Q}^{\nu+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(\nu+1)m}$ [98, Lemma 4.8.3]. However, we have shown that $(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}$ is not achieved by this input, thus this counterexample demonstrates also that

$$\mathrm{Q}^{\nu+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(\nu+1)m} \not\Rightarrow \quad (\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}. \tag{II.62}$$

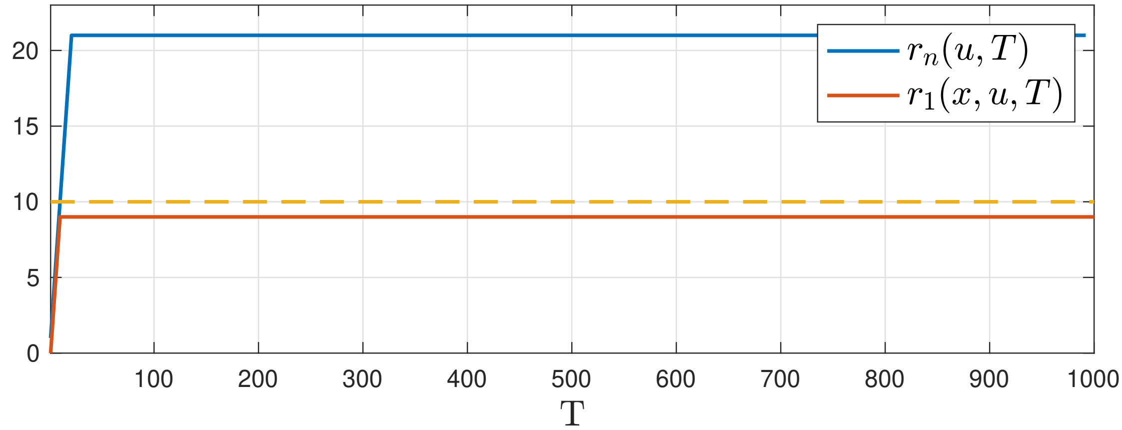

Figure II.1: Directions spanned in time by the signals (x, u) and $\mathrm{Q}^n(\mathrm{u})$.

◇

Necessary condition

We pick here a condition about u stronger than the one claimed to be necessary by Theorem II.4, and we show it is not guaranteed by PE of (x, u). We consider a discrete-time LTI system in form (II.43), with $n = 7, m = 3$, matrix B given in (II.58) and

$$A := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -0.3 & 0.2 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -0.3 & 0.2 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.7324 \end{bmatrix}.$$

It can be verified that the pair (A, B) is reachable and A is Schur. We want to show that $(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}$ does not ensure $\mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(n+1)m,n+m+1}$. By choosing initial condition $x_0 = 0$, and input

$$u_t = \begin{bmatrix} \sin(t) + \sin(2t) \\ \sin(3t) + \sin(4t) \\ \sin(5t) \end{bmatrix}, \tag{II.63}$$

simulating for $t = 1, \ldots, 1000$, it can be verified that $(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m}$. However, it can be verified also that $\mathrm{Q}^{n+1}(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{(n+1)m,n+m+1}$. With this example, we have shown that

$$(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m} \quad \not\Longrightarrow \quad \mathrm{Q}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(n+1)m,n+m+1}. \tag{II.64}$$

In Figure II.1, we plot the functions $r_1(x, u, T)$ as defined in (II.60) and

$$r_{n+1}(\mathrm{u}, T) := \operatorname{rank}\left(\sum_{t=0}^{T} \mathrm{Q}^{n+1}(\mathrm{u})_t \mathrm{Q}^{n+1}(\mathrm{u})_t^\top\right) \tag{II.65}$$

as the time window T of considered samples increases. Notice these plots show that the necessary condition given in Theorem II.4 holds.

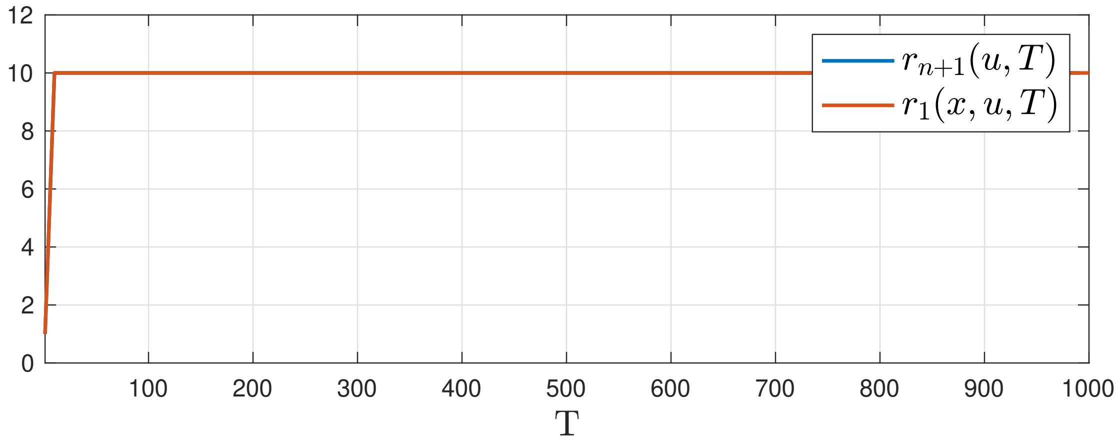

Figure II.2: Directions spanned in time by the signals (x, u) and $\mathrm{Q}^{n+1}(\mathrm{u})$.

Remark II.15. This example is a counterexample also for the necessary conjecture in [143, Pag. 4], which we may state (with some abuse of notation) as

$$(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m} \implies \mathrm{Q}^{\nu+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(\nu+1)m}, \tag{II.66}$$

where ν is the controllability [228, Pag. 121] of the pair (A, B). In this case, the controllability index of the pair (II.58) is $\nu = 3$, and $\nu + 1 < n$, so we may write $\mathrm{Q}^{\nu+1}(\mathrm{u}) = P\mathrm{Q}^{n}(\mathrm{u})$ for some full row rank matrix P. However, since we have shown that $\mathrm{Q}^{n}(\mathrm{u})$ spans only $n + m = 10$ directions, $\mathrm{Q}^{\nu+1}(\mathrm{u})$ can span at most 10 directions. Being $\mathrm{Q}^{\nu+1}(\mathrm{u})$ $(\nu + 1)m = 12$-dimensional, this means it is not PE, and thus this counterexample demonstrates also that

$$(\mathrm{x}, \mathrm{u}) \in \Omega^{\mathrm{D}}_{n+m} \not\Rightarrow \mathrm{Q}^{\nu+1}(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(\nu+1)m}. \tag{II.67}$$

◊

II.3.5 Future work

In this section, we have addressed the problem of guaranteeing persistence of excitation of state and input signals in the context of LTI systems via the application of a sufficiently rich input. Throughout the derivation of the results, we assumed an infinite-dimensional perspective allowing to make robustness considerations on both the persistence of excitation and sufficient richness properties. Exploiting the analogies between time shifts for discrete-time and derivatives for continuous-time, we are able to develop a unifying notation to state necessary and sufficient conditions to obtain PE of commonly used regressors for both the frameworks. Leveraging on these conditions, we analyzed the shape of the set of sufficiently rich signals for stable controllable LTI systems. Future work will be be done in exploring the following direction: given the class of multivariable linear systems sharing output and system dimensions, we have provided necessary and sufficient conditions to achieve PE which do not coincide. However, it has not been proven that, given a *single* linear system, it is not possible to find a unique necessary and sufficient condition characterizing all the sufficiently rich inputs for that system. Future work will thus be in this direction.

II.4 Sufficient richness in the special case of sinusoids

In this section, we find some tighter sufficient conditions for PE which shed light on why sufficient conditions on multi-input systems do not coincide with the necessary one.

II.4.1 Preliminaries: the geometric structure of linear systems

Before approaching the sufficient richness given in case of sinusoids, we give some preliminary notions on a powerful decomposition for linear systems.

Definition II.18. *[Rational canonical structure [228], Thm. 0.1] Let $A \in \mathbb{R}^{n\times n}$. There exists a full rank $T : \mathbb{R}^{n\times n}$ and a unique $k \in \mathbb{N}$ such that*

$$TAT^{-1} = \operatorname{diag}(A_1, \dots, A_k), \tag{II.68}$$

where each $A_i \in \mathbb{R}^{n_i\times n_i}$ is in the companion form

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{i,1} & a_{2,i} & \dots & \dots & a_{n_i} \end{bmatrix}. \tag{II.69}$$

♔

Definition II.19. *[Cyclic index [228], pag. 17] The cyclic index of A, hereby denoted as* $\operatorname{cyc}(A)$*, is the (unique) number of companion matrix blocks in its rational canonical form.* ♔

Notice that if the number of inputs m is smaller than the cyclic number of a matrix A, then it is not possible to have controllability. For simplicity reasons, we restrict the following results for the case of $m = \operatorname{cyc}(A)$. We stress however that this restriction could be easily avoided, and we refer the reader to [156, 228, 229] for more insight into these decompositions.

Theorem II.8. *[Thm. 1.2, [228]] Every controllable pair (A, B) such that $m = \mathrm{cyc}(A)$ admits a matrix representation*

$$A = \mathrm{diag}(A_1, \dots, A_m), \quad B = \begin{bmatrix} b_{11} & \dots & b_{1m} \\ 0 & \ddots & \vdots \\ 0 & 0 & b_{mm} \end{bmatrix}, \tag{II.70}$$

where A is in rational canonical form and each pair (A_i, b_{ii}) is controllable. ♛

Definition II.20. *[Controllability indices [228], pag. 121, [156] Prop. 2.1] Let (A, B) controllable and $m = \mathrm{cyc}(A)$. The controllability indices $n_1, \dots, n_m$ of A are the (unique) dimensions of the blocks A_i on the diagonal of its rational canonical form. The controllability index of A is defined as*

$$\bar{n} = \max\{n_1, \dots, n_m\}. \tag{II.71}$$

Alternatively, the controllability index of (A, B) is the smallest natural $\bar{n}$ for which the matrix

$$R(\bar{n}) = [A^{\bar{n}-1}B, \dots, B] \tag{II.72}$$

is full rank. ♔

II.4.2 Preliminaries: spectral lines and PE

One of the main concepts behind the definitions and results given in this section is the notion of *spectral line* [41, Def. 3.2], which we provide here for signals defined over $\mathbb{R}_{\geq 0}$.

Definition II.21. *[Spectral line [41] 🕮] A signal $w : \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ is said to have a spectral line at frequency $\omega \in [-\infty, \infty]$ of amplitude $\hat{w}(\omega) \in \mathbb{C}^d \neq 0$ if and only if*

$$\hat{w}(\omega) \coloneqq \lim_{T\to\infty} \frac{1}{T} \int_s^{s+T} w(\tau) e^{-j\omega\tau} d\tau \tag{II.73}$$

exists uniformly in $s \in \mathbb{R}_{\geq 0}$. ♔

Example II.1. A very easy way to introduce spectral lines in a signal is through sinusoids. Consider, e.g., the real signal

$$u(t) = U\cos(\omega t + \varphi) = \frac{U}{2}\left(e^{i(\omega t+\varphi)} + e^{-i(\omega t+\varphi)}\right). \tag{II.74}$$

It can be checked that this signal contains two spectral lines at frequencies ω and $-\omega$, and their amplitudes are given by

$$\hat{u}(\omega) = \frac{U}{2}e^{i\varphi}, \qquad \hat{u}(-\omega) = \frac{U}{2}e^{-i\varphi}. \tag{II.75}$$

Consider now the signal $u(t) = U(\sin(\omega t), \cos(\omega t))$. Containing only a single frequency, the new signal has again two spectral lines, and applying (II.74), we can find them to be linearly independent in $\mathbb{C}^2$; furthermore, their expression is given by:

$$\hat{u}(\omega) = \frac{U}{2}\begin{bmatrix}-i\\1\end{bmatrix} \qquad \hat{u}(-\omega) = \frac{U}{2}\begin{bmatrix}i\\1\end{bmatrix}. \tag{II.76}$$

◇

More generally, a real-valued signal containing a spectral line at ω_0 with amplitude $\hat{w}(\omega_0)$ has always another spectral line at $-\omega_0$, for which it holds that $\hat{w}(\omega_0) = \hat{w}(-\omega_0)^*$. This type of description is particularly well-suited for periodic and almost-periodic signals. In fact, it is shown in [34, V] that an almost-periodic signal $w(t)$ has spectral lines only in the so-called set of characteristic exponents $\{\omega_i\}_{i\in\mathbb{N}}$, which is a countable and unique set for any almost periodic signal. Moreover [34, XV], any almost-periodic signal $w(t)$ can be arbitrarily approximated by a trigonometric polynomial consisting of its spectral lines $\hat{w}(\omega_i)$ at its characteristic exponents (the so called almost-periodic Fourier series). Similar approximation capabilities hold trivially also for periodic signals [207, Pag. 79].
In the following lemma, we report the result of [41, Lemma 3.4] that establishes the connection between the spectral lines of a signal and persistency of excitation by showing the condition is not only sufficient but also necessary.

Lemma II.10. *[[41], Lemma 3.4] Let $w \in C_b^\infty(\mathbb{R}^d)$ have at least d spectral lines whose amplitudes $\hat{w}(\omega_1), \dots, \hat{w}(\omega_d)$ are linearly independent in $\mathbb{C}^d$. Then, $w(t) \in \Omega_d^{\mathrm{C}}$* ♗

For completeness, we report also [41, Prop. 5.1] since it is the key result to understand why the above presented decompositions are fundamental.

Proposition II.1. *[[41], Prop. 5.1] Let $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x},1}$, and let the scalar input* u *to the system have n spectral line. Then,* $\mathrm{u} \in C_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},1})$. ♖

Whilst the above result holds only for single-input system, we show in the next section how to leverage on the above presented decomposition for finding tighter conditions for sufficiently rich signals.

II.4.3 Sufficiently rich sinusoids

We start our analysis by considering multivariable linear systems in rational canonical structure as per Definition II.18.

Theorem II.9. *Let $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}$ be in form* (II.70)*, and let the input* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$ *be such that each i−th input component, $i = 1, \dots, m$, contains at least n_i spectral lines at frequencies ω_{ij}, $j = 1, \dots, n_i$, where all $\omega_{ij} \in \mathbb{R}$ are distinct and n_i are the controllability indices of (A, B). Then, for all $x(0) \in \mathbb{R}^n$,* $\mathrm{u} \in C_{\mathrm{SR}}(\sigma, x(0))$. ♕

Whilst we leave the proof in Appendix V.4.10 for readability reasons, we give an high-level overview on the involved mechanism.

i) Decomposition (II.70) allows us to obtain $\mathrm{cyc}(A)$ linear systems which are coupled only in the input signal. Notice however that the m−th system is a single-input system. It is thus sufficient to inject n_m spectral lines into the m-th component of the input to obtain PE in the last components of $x(t)$, applying [41, Prop. 5.1].

ii) Passing to the $(m-1)$-th subsystem, we inject n_{m-1} spectral lines via the $(m-1)$-th input component. However, notice that this time we do not have a single input system, since B is in (block) upper triangular form. However, it is possible to show that the previously injected spectral lines "do not interfere" with those in the $(m-1)$-th component, and that they are all linearly independent.

iii) By repeating the same reasoning for all the subsystems, we obtain the given result.

Notice that, whilst the above result is similar to [232, Prop. 4], in our theorem it is not necessary to inject n sinusoid at different frequencies to achieve sufficient richness; it is enough to inject (overall) only n spectral lines, which can be obtained from (see Example II.1) $n/2$ different frequencies. Furthermore, we are not restricting the input signal to be periodic or almost periodic.

Remark II.16. Notice that an input signal respecting the conditions of Theorem II.9 is not necessarily sufficiently rich for all systems $\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1}$ of fixed dimension. In fact, the controllability indices of two matrices $A_1, A_2 \in \mathbb{R}^{n\times n}$ may be different, and thus the number of required spectral lines may vary. ◇

The above presented theorem is insightful from a theoretical point of view; however, its application for design purposes is not adapt to uncertain systems since it requires the knowledge of the similarity transformation bringing (A,B) in rational canonical form (II.70). In the following lemma, we establish the importance of the knowledge of the controllability index of a matrix A, since it allows the application of the above theorem without requiring the knowledge of the full decomposition.

Lemma II.11. *Let $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}$ and $m = \mathrm{cyc}(A)$. Let the input $\mathrm{u} \in C^\infty_b(\mathbb{R}^m)$ be such that each i−th input component, $i = 1,\dots,m$, contains at least $\bar{n}$ spectral lines at frequencies ω_{ij}, $j = 1,\dots,\bar{n}$, where all $\omega_{ij} \in \mathbb{R}$ are distinct and $\bar{n}$ is the controllability index of (A,B). Then, for all $x(0) \in \mathbb{R}^n$, $\mathrm{u} \in C_{\mathrm{SR}}(\sigma, x(0))$.* ♙

Proof. The proof is straightforward realizing that i) there exists a similarity transformation that brings A, B into form (II.70), and ii) that under this similarity transformation, since $\bar{n} \geq n_i$ for all $i = 1,\dots,m$, by injecting $\bar{n}$ spectral lines on each component we ensure that all the hypotheses of Theorem II.9 are verified. ※

Notice that, even if $\bar{n}$ was not known, it always holds $n \geq \bar{n}$, so one could inject n spectral lines into each component (and thus recover [98, Thm. 5.2.3]). To conclude, we state the above lemma in a more generic version, inspired by [232, Thm. 3]

Theorem II.10. *Let the input $\mathrm{u} \in C^\infty_b(\mathbb{R}^m)$ contain at least m sets of $\bar{n}$ spectral lines at different frequencies $\hat{u}(\omega^i_1),\dots,\hat{u}(\omega^i_{\bar{n}})$, $i = 1,\dots,m$ such that* *i) for all* $i = 1, \ldots, m$, $j = 1, \ldots, \bar{n}$

$$\hat{u}(\omega_j^i) = k_{ij}\hat{u}_i \tag{II.77}$$

for some $\hat{u}_i \in \mathbb{C}^m$ *and* $k_{ij} \in \mathbb{C}$

ii) $\hat{u}_1, \ldots, \hat{u}_m$ *are linearly independent in* $\mathbb{C}^m$. *Then, for all* $x(0) \in \mathbb{R}^n$, $\mathrm{u} \in C_{\mathrm{SR}}(\sigma, x(0))$.

♛

Proof. Denote as $\hat{U} \in \mathbb{C}^{m\times m}$ the full rank matrix

$$\hat{U} := \begin{bmatrix} \hat{u}_1 & \ldots & \hat{u}_m \end{bmatrix}. \tag{II.78}$$

Notice that any spectral line in $u(t)$ can be written as

$$\hat{u}(\omega_j^i) = \hat{U}\mathrm{e}_i k_{ij}, \tag{II.79}$$

where $\mathrm{e}_i \in \mathbb{C}^m$ is a zero vector with a 1 in the i-th entry. Define $v(t) := \hat{U}^{-1}u(t)$. In the new coordinates, the dynamic reads

$$\dot{x} = Ax + B\hat{U}v. \tag{II.80}$$

Notice that the spectral lines of u are related to the spectral lines of v by the relation

$$\begin{aligned} \hat{v}(\omega_j^i) &= \hat{U}^{-1}\hat{u}(\omega_j^i) = \hat{U}^{-1}\hat{U}\mathrm{e}_i k_{ij} \\ &= \mathrm{e}_i k_{ij}, \end{aligned} \tag{II.81}$$

namely, in the new coordinates we obtain an input $v(t)$ which has $\bar{n}$ spectral lines on each components at distinct frequencies. Since the change of coordinate in the input space does not lose controllability of the new pair $(A, B\hat{U})$, we can then apply Theorem II.11 to conclude the proof. ※

Chapter III

Data-driven stabilization via filtering for continuous-time linear systems

Knowledge of the system state and its derivatives is fundamental in both the estimation of unknown parameters and in the direct data-driven control. In this chapter, we prepare at first the theoretical framework for a new type of observer, which we call "gazer". Differently from other approaches, we are not interested in knowing the true system state: we are interested only in "representing" the plant dynamics, according to some requirements, here informally listed:

i) *Model-free*: only rough knowledge about the plant dimensions should be necessary for the gazer implementation.

ii) *Representation*: we seek for the existence of a surjective, time-invariant map Π between the state of the gazer ζ and the state of the plant x, mapping gazer trajectories into plant trajectories. Its knowledge is not required.

iii) *Attractivity*: we require that, for any initialization of the gazer and plant state $\zeta(0), x(0)$, it holds asymptotically that $\Pi\zeta(t) \to x(t)$.

iv) *Stability*: the gazer should not introduce instabilities.

The idea of directly looking for surjective maps from the gazer state to the plant state somehow shortcuts the established design procedure of finding at first an injective map T from the plant state to the gazer state, and then inverting it [26, 45, 106, 140], with advantages and disadvantages which are still not entirely clear (to me).
After having explored gazers for SISO and MIMO systems, we show how to use them for solving control problems by developing a data-driven stabilization method for CT-LTI systems with theoretical

guarantees and no reliance on signal derivatives. The framework is based on linear matrix inequalities (LMIs) and is illustrated in the state-feedback and single-input single-output output-feedback scenarios. To avoid the need for differentiation, we exploit filters that, rather than approximating the derivatives, reconstruct a non-minimal realization of the plant where the state and its derivative are accessible. Using batches of input and filtered data, LMIs are then employed to compute a dynamic stabilizer of the plant. The effectiveness of the framework is validated through numerical examples.

This chapter is organized as follows. In Section III.1 we give an overview of the scientific literature pertaining adaptive observers and data-driven LMI-based data-driven control, highlighting the typical assumptions and requirements in the continuous-time framework. In Section III.2, we state the design requirement for gazers. In Section III.3, we give sufficient criterias for the gazer design, and we propose some possible solutions leveraging some new matrix results. In Section III.4, we provide a stabilization numerical algorithm for the state-feedback and the SISO scenarios, describing its properties and validating the approach with numerical results. The main proofs are given contextually; the others can be found in the Appendix V.5. Part of the results of this chapter have been submitted as a conference paper, [38].

III.1 Literature review

Over the past decades, the paradigm of data-driven learning has gained increasing attention in control theory. A key factor to control a system, both in the model-based and in the data-driven context, is the availability of measurements of the system state, which enables the application of state-feedback techniques. This is particularly important when the controlled system has an uncertain dynamics, since the *information* needed to achieve control has to be retrieved mainly from the collected data, and not from the system knowledge. The extraction of the state information from output measurements becomes thus a very important aspect in every data-driven control technique. From the earliest works by Kalman [103–105] and by Luenberger [137–140], it was clear that certain dynamical systems allow for the extraction of this information from the measured output. Since then, a great number of observers, namely, dynamical systems able to asymptotically reconstruct the state from measurements, have been developed (see [26] for a comprehensive review). Implementing an observer when the dynamics of the system is uncertain increases significantly the challenges in its design. In the so-called adaptive observers, the proposed solution is to leverage canonical forms in which a limited number of dynamics parameters are updated by a suitable update law, at first considering SISO plants [48, 111, 135, 136] and then passing to the MIMO framework [5, 97, 163, 230]. Among these works, [5] is particularly interesting since it presents an adaptive observer which satisfy the previously mentioned criteria design, namely, neither model knowledge or an update law are necessary for achieving convergence to a nonminimal representation of the state.

Once clarified how to extract full-state information from the collected data, several fields such as system identification [123], adaptive control [98] and, more recently, reinforcement learning [202] proposed algorithm to use it in different ways. Recently, inspired by the results in [40], the dominant paradigm in data-driven control has become to compute controllers directly from data using linear matrix inequalities (LMIs) or other optimization problems, without even requiring an intermediate identification step [62]. In this work, we focus on LMI-based methods for the stabilization of continuous-time linear time-invariant (LTI) systems.

Fundamental contributions to data-driven control of discrete-time systems include [62] and [212], which introduced two distinct data-based LMI formulations for state-feedback stabilization. These methodologies have since been used to address the stabilization of bilinear systems [31], linear time-varying systems [166], and the linear quadratic regulator problem [62], also accounting for the effects of noise [63, 68]. The integration of partial model knowledge into these approaches was explored in [24], and in [215, 216] the authors study how to give probabilistic guarantees on exploration and robust stabilization. Moreover, necessary and sufficient conditions for data informativity have been thoroughly investigated [214].

In the continuous-time scenario, the discrete-time state-feedback stabilization paradigm can be recovered via suitable sampling techniques. However, this comes at the cost of requiring samples of the state derivatives [62], causing robustness issues in the presence of noise. In [25], LMIs inspired by [212] were derived for the design of a stabilizing gain with non-periodic sampling and noisy state-derivative

estimates. Similarly, derivative estimates were employed in [146], which proposed quadratic matrix inequalities for stabilizing linear parameter-varying systems in both discrete and continuous time. Recent contributions include the study of the impact of sampling on data informativity [72] and the stabilization of continuous-time switched and constrained systems [30].
To avoid differentiation in the state-feedback scenario, [61] proposed to construct datasets based on integrals and temporal differences of the available signals. To the best of the authors' knowledge, no other algorithm in the continuous-time literature completely removes the need for state derivatives. Furthermore, no output-feedback approach has been developed thus far, where the sensitivity to noise is even more pronounced due to the need for multiple differentiations.

III.2 Problem setup

III.2.1 Preliminaries: some geometric properties of linear systems

The solutions of $\Pi\Lambda = A\Pi$

Consider the square matrices $A \in \mathbb{R}^{n\times n}, \Lambda \in \mathbb{R}^{m\times m}$ Given any similarity transformations $U \in \mathbb{R}^{n\times n}, V \in \mathbb{R}^{m\times m}$ which bring A, Λ into their Jordan forms $\tilde{A}$ and $\tilde{\Lambda}$, namely

$$\begin{aligned} A &= U\tilde{A}U^{-1} \\ \Lambda &= V\tilde{\Lambda}V^{-1}, \end{aligned} \tag{III.1}$$

all solutions $\Pi \in \mathbb{R}^{n\times m}$ to the equation $\Pi\Lambda = A\Pi$ are given by [76, Pag. 219]

$$\Pi = U\tilde{\Pi}V^{-1}, \tag{III.2}$$

where $\tilde{\Pi}$ can be divided in blocks

$$\tilde{\Pi} = \begin{bmatrix} \tilde{\Pi}_{11} & \dots & \tilde{\Pi}_{1v} \\ \dots & \dots & \dots \\ \tilde{\Pi}_{u1} & \dots & \tilde{\Pi}_{uv} \end{bmatrix} \tag{III.3}$$

with $u, v \in \mathbb{N}$ the number of Jordan blocks in A, Λ. By choosing the same ordering for the eigenvalues of A and Λ given by the chosen Jordan forms $\tilde{A}, \tilde{\Lambda}$, we denote the $i-$th eigenvalue of A as λ_i^A (and the same for Λ). Each block $\tilde{\Pi}_{ij}$ in (III.3) is zero if $\lambda_i^A \neq \lambda_j^\Lambda$. In case $\lambda_i^A = \lambda_j^\Lambda$, $\tilde{\Pi}_{ij}$ is given by

$$\tilde{\Pi}_{ij} = \begin{bmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix} \in \mathbb{R}^{n_i \times n_j}, \tag{III.4}$$

where n_i, n_j are the algebraic multiplicity of eigenvalues $\lambda_i^A = \lambda_j^\Lambda$ in that Jordan block, and $a, b, c, d, \dots \in \mathbb{R}$ are free parameters whose number is determined by $\min(n_i, n_j)$. In case $n_i \neq n_j$, the blocks may have the rectangular forms

$$\tilde{\Pi}_{ij} = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{\Pi}_{ij} = \begin{bmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix}. \tag{III.5}$$

A Jordan controllability decomposition for (Λ, ℓ)

Given (Λ, ℓ) controllable, there exists a similarity transform V [95, Example 3.4] such that

$$\Lambda = V\tilde{\Lambda}V^{-1}, \quad \ell = V \begin{bmatrix} e_{n_1^\Lambda} \\ \dots \\ e_{n_v^\Lambda} \end{bmatrix}, \tag{III.6}$$

where $\tilde{\Lambda}$ is in Jordan form and where each $e_{n_i^\Lambda} \in \mathbb{R}^{n_i}$ is a vector of zeros but for a 1 in the last entry, with n_i the algebraic multiplicity of λ_i. Notice that there is a row correspondence between each $i-$th Jordan block of $\tilde{\Lambda}$ and the $i-$th vector $e_{n_i^A}$, namely, for all $i = 1, \dots, u$

$$\begin{bmatrix} \lambda_i^\Lambda & 1 & \dots & 0 \\ 0 & \ddots & \ddots & \dots \\ \dots & \dots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i^\Lambda \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 0 \\ \dots \\ \dots \\ 1 \end{bmatrix}, \tag{III.7}$$

where u is the number of distinct eigenvalues in Λ (and also the number of Jordan blocks, since Λ is cyclic [228, Pag. 15] being (Λ, ℓ) controllable).

How is controllability modified by output injection?

Since, in the subsequent discussion, we will often deal with output injections and feedback gains, it is interesting to study how the controllability properties of a pair (A, b) are influenced by an output injection $\psi c^\top$. More specifically, consider $A \in \mathbb{R}^{n\times n}$, $b, c \in \mathbb{R}^n$ such that $(c^\top, A)$ is observable. We have the following result.

Lemma III.1. *Let $(c^\top, A)$ observable with $A \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^n$. Choose any symmetric set of n complex numbers $\sigma^\star$. Choose $\psi \in \mathbb{R}^n$ such that $\sigma(A - \psi c^\top) = \sigma^\star$. Then,*

i) for all $\lambda \in \sigma^\star \cap \sigma(A)$, $\operatorname{rank} \begin{bmatrix} A - \psi c^\top - \lambda I & b \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A - \lambda I & b \end{bmatrix}$

ii) for all $\lambda \in \sigma^\star \setminus \sigma(A)$, $\operatorname{rank} \begin{bmatrix} A - \psi c^\top - \lambda I & \psi \end{bmatrix} = n$.

♙

The proof of Lemma III.1 is provided in Appendix V.5.1.
Trivially, Lemma III.1 says that if $\sigma^\star \cap \sigma(A) = \emptyset$ then $(A - \psi c^\top, \psi)$ is controllable. If this is not the case, if (A, b) is stabilizable (controllable), then $(A - \psi c^\top, [b\ \psi])$ is stabilizable (controllable).

III.2.2 Problem statement: gazing the state of a linear system

Consider a continuous-time linear time-invariant system in the form

$$\begin{aligned} \dot{x} &= Ax + By \\ y &= Cx \end{aligned} \tag{III.8}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $y \in \mathbb{R}^p$ is the output, $A \in \mathbb{R}^{n\times n}$ is the state matrix, $B \in \mathbb{R}^{n\times m}$ is the input matrix, while $C \in \mathbb{R}^{p\times n}$ is the output matrix.
Given the dynamical system (III.8), consider an auxiliary system in the form

$$\dot{\zeta} = \mathscr{A}\zeta + \mathscr{L}y + \mathscr{B}u, \tag{III.9}$$

with state $\zeta \in \mathbb{R}^z$ and system matrices $\mathscr{A} \in \mathbb{R}^{z\times z}$, $\mathscr{B} \in \mathbb{R}^{z\times m}$, $\mathscr{L} \in \mathbb{R}^{z\times p}$, driven by the same input and output of system (III.8).

Definition III.1. *We say system* (III.9) *is a* gazer *of system* (III.8) *if it satisfies the following requirements.*

i) ***Submersion:*** *There exists a linear surjective map* $\Pi : \mathbb{R}^z \to \mathbb{R}^n$ *such that, for all* $x(0) \in \mathbb{R}^n$, *there exists* $\zeta(0) \in \mathbb{R}^z$ *for which*

$$x(t) = \Pi\zeta(t) \tag{III.10}$$

for all $t \geq 0$ *and for any input signal* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$ *entering dynamics* (III.8) *and* (III.9). *Notice that pre-multiplication of both sides of equation* (III.10) *by C defines a map* $\mathscr{C} : \mathbb{R}^z \to \mathbb{R}^p$ *which satisfies*

$$\mathscr{C} = C\Pi \tag{III.11}$$

for which, if $x(t) = \Pi\zeta(t)$, *then* $y(t) = \mathscr{C}\zeta(t)$. *In other words, any trajectory of the plant* (III.8) is represented *by a trajectory of the gazer.*

ii) ***Attractivity:*** *For any* $x(0) \in \mathbb{R}^n$, $\zeta(0) \in \mathbb{R}^z$ *and* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$, *it holds*

$$\lim_{t\to\infty} |\Pi\zeta(t) - x(t)| = 0. \tag{III.12}$$

In other words, any trajectory of the gazer converges in time to a trajectory representation of the current plant trajectory.

iii) ***Stabilizability:*** *The map $\mathscr{C}$ in* (III.11) *defines the closed-loop system*

$$\begin{aligned}\dot{\zeta} &= (\mathscr{A} + \mathscr{L}\mathscr{C})\zeta + \mathscr{B}u \\ y &= \mathscr{C}\zeta.\end{aligned} \tag{III.13}$$

We want that, under stabilizability or controllability of the true pair (A, B), at least stabilizability is preserved for the pair $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$.

♕

When we say that a gazer is "in filter form", we refer to the dynamical system (III.9) driven by the plant output y. When we say that a gazer is "in feedback form", we refer to the dynamical system (III.13). Since in the framework in which we are interested we have little or no knowledge of the true system matrices A, B, C, a property that we seek for the design of $\mathscr{A}, \mathscr{B}, \mathscr{L}$ is their independence on the knowledge of A, B, C. Namely, we want to implement the update law (III.9) by requiring only the measurements of y and u and designing $\mathscr{A}, \mathscr{B}, \mathscr{L}$ a priori, leveraging the existence of the unknown -but unnecessary- map $\mathscr{C}$ (which is conditioned by the existence of the map Π) to match the two dynamical systems. Furthermore, another desirable feature is the possibility of choosing z without the exact knowledge of n. A few *observations* are in order.

Remark III.1. Full row rankess of Π and $z \geq n$ are necessary conditions to fulfill the first requirement. Given the linearity of the framework, the map Π defines a submersion between the trajectories of the gazer (III.9) and the trajectories of system (III.8). ◇

Remark III.2. The stabilizability requirement ensures that the internal dynamics of the filter system are stable (if not controllable). In other words, attractivity ensures that the information in $x(t)$ is contained also in $\zeta(t)$ for any initialization; in addition to this we require that, if the plant possesses good controllability properties, then the filter (III.9) must not blow up. ◇

Remark III.3. Differently from an observer, we are not interested in finding the true plant trajectory $x(t)$, we are just interested in reaching a *representation* $\zeta(t)$ of that trajectory. Differently from adaptive observer, a gazer does not require any additional adaptation law nor persistent excitation of $\zeta(t)$ to reach this representation. Adaptive laws may be implemented to obtain an estimate of the output map $\mathscr{C}$; however, we will show in Section III.4 that this is not a necessary step for control purposes. ◇

III.3 Gazer design for LTI systems

III.3.1 Sufficient criterias for the design of a gazer

The submersion problem

Consider the matrix representation of Π defined in (III.10). Pre-multiplying by Π the equation (III.9), and using (III.10) - (III.11), we have that any Π satisfying the submersion requirement satisfies the following system of matrix equations

$$\begin{cases} \Pi(\mathscr{A} + \mathscr{L}\mathscr{C}) = A\Pi \\ \mathscr{C} = C\Pi \\ \Pi\mathscr{B} = B. \end{cases} \tag{III.14}$$

Indeed, notice how these equations are similar - but not the same - as those arising in classical observer design [140, Eqns. (4), (16)]. The theoretical difference consists in directly looking for a surjective map $\Pi : \zeta \mapsto x$ instead of looking for an injective $T : x \mapsto \zeta$ and then left-inverting it. Having a clear understanding of the practical differences between the two approaches will be the subject of future work; however, equations of the type (III.14) seems to pose less constraints on the design of the input matrix $\mathscr{B}$. In the next lemma, we characterize a necessary condition for solving (III.14) for a full rank Π which already gives insight on some properties of the pair $(\mathscr{A}, \mathscr{L})$.

Lemma III.2. *Any full rank solution to* (III.14) *is such that* $\sigma(A) \subseteq \sigma(\mathscr{A} + \mathscr{L}\mathscr{C})$. ♙

Proof. The proof (by contraposition) is straightforward by recalling from Section III.2.1 that, if $\mathscr{A}+\mathscr{L}\mathscr{C}$ and A are in Jordan form, if there exists an eigenvalue $\lambda \in \sigma(A)$, $\lambda \notin \sigma(\mathscr{A} + \mathscr{L}\mathscr{C})$ then there is at least one row of zeros in Π. ※

Lemma III.2 requires the output map $\mathscr{C}$ to be able to place at least n eigenvalues in the correct place from $\mathscr{A}$ to $\mathscr{A} + \mathscr{L}\mathscr{C}$. This, in turn, can be achieved only levaraging the controllability properties of the pair $(\mathscr{A}, \mathscr{L})$. In other words, the existence of solutions to equations (III.14) will depend on the capability of $\mathscr{C}$ i) to match the eigenvalues A - if seen, via $\mathscr{L}\mathscr{C}$, as an output injection term-, and ii) to match the output map (III.11). We conclude this section with the following proposition, which sums up the discussion (we omit the proof since it is straightforward).

Proposition III.1. *Any full rank matrix* $\Pi \in \mathbb{R}^{n\times z}$ *satisfying equations* (III.14) *solves the submersion problem, namely, for all* $x(0) \in \mathbb{R}^n$, *there exists* $\zeta(0) \in \mathbb{R}^z$ *for which* $x(t) = \Pi\zeta(t)$ *for all* $t \geq 0$ *and for any input signal* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$ *entering dynamics* (III.8) *and* (III.9). ♖

Remark III.4. Notice that the existence of a full rank matrix $\Pi \in \mathbb{R}^{n\times z}$ satisfying (III.14) can be used to define a coordinate change

$$\zeta \longmapsto \begin{bmatrix} \xi \\ \eta \end{bmatrix} := \begin{bmatrix} \Pi \\ \Psi \end{bmatrix} \zeta, \tag{III.15}$$

(where the rows in $\Psi \in \mathbb{R}^{(z-n)\times z}$ are free but linearly independent from those of Π) such that the gazer is in Kalman observability form:

$$\begin{aligned} \begin{bmatrix} \dot{\xi} \\ \dot{\eta} \end{bmatrix} &= \begin{bmatrix} A & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} B \\ B_2 \end{bmatrix} u \\ y &= C\xi, \end{aligned} \tag{III.16}$$

proving that the observable part of the gazer and the original system coincide. This change of coordinates underlines the presence of an internal model of the plant into the gazer. ◊

The attractivity problem

Next, we consider the attractivity problem. Supposing that a full rank Π satisfying (III.14) exists, we define the observation error

$$\epsilon \coloneqq x - \Pi\zeta, \tag{III.17}$$

and we obtain that its dynamics, induced by systems (III.8) and (III.9), is given by

$$\begin{aligned} \dot{\epsilon} &= \dot{x} - \Pi\dot{\zeta} \\ &= Ax + Bu - \Pi(\mathscr{A}\zeta + \mathscr{L}Cx + \mathscr{B}u) \\ &= (A - \Pi\mathscr{L}C)\epsilon + (A\Pi - \Pi(\mathscr{A} + \mathscr{L}C\Pi))\zeta + (B - \Pi\mathscr{B}u) \\ &= (A - LC)\epsilon, \end{aligned} \tag{III.18}$$

where we have defined $L \coloneqq \Pi\mathscr{L}$. By construction, if Π satisfies the system (III.14), then it solves also (by substituting the second equation into the first)

$$\begin{cases} \Pi\mathscr{A} = (A - LC)\Pi \\ \Pi\mathscr{B} = B. \end{cases} \tag{III.19}$$

Since Π is a full row rank solution of (III.19), it holds necessarily that

$$\sigma(A - LC) \subseteq \sigma(\mathscr{A}). \tag{III.20}$$

From this consideration, we derive the following result (we omit the proof since it follows trivially from the above discussion).

Proposition III.2. *Let a full rank solution* Π *of* (III.14) *exist. If* $\mathscr{A}$ *is Hurwitz, then the attractivity problem is solved, namely, for any* $x(0) \in \mathbb{R}^n$, $\zeta(0) \in \mathbb{R}^z$ *and* $\mathrm{u} \in C_b^\infty(\mathbb{R}^m)$, *it holds* $\lim_{t\to\infty} |\Pi\zeta(t) - x(t)| = 0$. ♖

The stabilizability problem

We start with a brief discussion on why we require only the stabilizability of the pair $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$ (and not its controllability). In fact, one may be interested the estimation of the matrix $\mathscr{C}$ leveraging the algebraic relation $y = \mathscr{C}\zeta$. For this estimation to be successful, persistency of excitation of ζ is in general required, and it is known that this property can be enforced through an input signal only if the pair $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$ is controllable [161, Lemma 5] (at least, in stable systems).
Nevertheless, we now leverage Remark III.4 to show that the matrix $\mathscr{C}$ can be entirely estimated through an appropriate initialization of the filters also in absence of controllability (since PE on the full state

ζ is not needed). Consider, for example, an estimate $\hat{\mathscr{C}} \in \mathbb{R}^{p\times z}$ which is updated via a gradient algorithm [98, §4.3.5], namely,

$$\dot{\hat{\mathscr{C}}} = (y - \hat{\mathscr{C}}\zeta)\zeta^\top. \tag{III.21}$$

Defining the error coordinate $\tilde{\mathscr{C}} := \hat{\mathscr{C}} - \mathscr{C}$, its dynamics is given by

$$\begin{aligned}\dot{\tilde{\mathscr{C}}} &= (y - \hat{\mathscr{C}}\zeta)\zeta^\top \\ &= (\mathscr{C}\zeta - \hat{\mathscr{C}}\zeta)\zeta^\top \\ &= -\tilde{\mathscr{C}}\zeta\zeta^\top.\end{aligned} \tag{III.22}$$

From the decomposition in III.16, notice that if (A, B) is controllable then any uncontrollable subspace must lie in $\ker(\Pi)$, and thus also in $\ker(\mathscr{C})$, since $\mathscr{C} = C\Pi$. This means we can always find a full rank change of coordinates $T \in \mathbb{R}^{z\times z}$ such that

$$T\zeta = \begin{bmatrix}\zeta_\text{c}\\ \zeta_\text{nc}\end{bmatrix} \qquad \mathscr{C}T^{-1} = \begin{bmatrix}\mathscr{C}_\text{c} & 0\end{bmatrix}, \tag{III.23}$$

where ζ_c spans the controllable subspace and ζ_nc the uncontrollable one. Since these subspaces are orthogonal, we pick T orthogonal, namely, $T^{-1} = T^\top$, and by post-multiplying (III.22) by T^{-1}, it holds that

$$\begin{aligned}\dot{\tilde{\mathscr{C}}}T^{-1} &= -\tilde{\mathscr{C}}\zeta\zeta^\top T^{-1} \\ \begin{bmatrix}\dot{\tilde{\mathscr{C}}}_\text{c} & \dot{\tilde{\mathscr{C}}}_\text{nc}\end{bmatrix} &= -\tilde{\mathscr{C}}T^{-1}T\zeta\zeta^\top T^\top T^{-\top}T^{-1} \\ &= -\begin{bmatrix}\tilde{\mathscr{C}}_\text{c} & \tilde{\mathscr{C}}_\text{nc}\end{bmatrix}\begin{bmatrix}\zeta_\text{c}\\ \zeta_\text{nc}\end{bmatrix}\begin{bmatrix}\zeta_\text{c}\\ \zeta_\text{nc}\end{bmatrix}^\top T^{-\top}T^{-1} \\ &= -\begin{bmatrix}\tilde{\mathscr{C}}_\text{c} & \tilde{\mathscr{C}}_\text{nc}\end{bmatrix}\begin{bmatrix}\zeta_\text{c}\\ \zeta_\text{nc}\end{bmatrix}\begin{bmatrix}\zeta_\text{c}\\ \zeta_\text{nc}\end{bmatrix}^\top.\end{aligned} \tag{III.24}$$

By choosing $\zeta(0) = 0$, it holds $\zeta_\text{nc} = 0$ for all $t \geq 0$, and thus the gradient update (III.24) becomes

$$\begin{aligned}\begin{bmatrix}\dot{\tilde{\mathscr{C}}}_\text{c} & \dot{\tilde{\mathscr{C}}}_\text{nc}\end{bmatrix} &= -\begin{bmatrix}\tilde{\mathscr{C}}_\text{c} & \tilde{\mathscr{C}}_\text{nc}\end{bmatrix}\begin{bmatrix}\zeta_\text{c}\zeta_\text{c}^\top & 0\\ 0 & 0\end{bmatrix}^\top \\ &= -\begin{bmatrix}\tilde{\mathscr{C}}_c\zeta_\text{c}\zeta_\text{c}^\top & 0\end{bmatrix},\end{aligned} \tag{III.25}$$

which, if $\hat{\mathscr{C}}(0) = 0$ requires only PE of the controllable $\zeta_\text{c}(t)$ to converge to zero. In other words, if the plant output matrix C is identifiable, then also $\mathscr{C}$ must be identifiable.

Remark III.5. Notice furthermore that if the gazer matrices $\mathscr{A}, \mathscr{L}, \mathscr{B}$ are known by design, and if $\mathscr{C}$ is successfully estimated, then it is possible to bring the system in Kalman observability form, thus reconstructing the true state $x = \Pi\zeta$ (or, at least, a full rank coordinate change of x). ◇

We conclude this section with the following sufficient condition to achieve the stabilizability requirement.

Proposition III.3. *Let a full rank solution* Π *of* (III.14) *exist. If* $(\mathscr{A}, \mathscr{B})$ *is stabilizable, the stabilizability problem is solved, namely, if* (A, B) *is stabilizable then* $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$ *is stabilizable.* ♖

The proof of Proposition III.3 is given in Appendix V.5.2.
Collecting Propositions III.1, III.2 and III.3, and using them as guidelines for the design of a system gazer, the main issue that still needs to be solved is an in-depth study of the solvability of the system of equations (III.14). In the following, we propose two "prototype equations" which will be used to guarantee the feasibility of the subsequent design of gazers.

III.3.2 Special matrix equations and their solution

The system $\Pi\Lambda = A\Pi$**,** $\theta^\top = c^\top\Pi$

We consider now a simplified framework, which may be interpreted as (III.14) when the plant is autonomous with a single output. In fact, it is useful to drop (for the moment) the presence of the input to better understand the importance of the observability properties of $(c^\top, A)$.
Letting $A, \Lambda \in \mathbb{R}^{n\times n}$ and $c, \theta \in \mathbb{R}^n$, we are interested in solving in Π the system of equations

$$\begin{cases} \Pi\Lambda = A\Pi \\ \theta^\top = c^\top\Pi. \end{cases} \tag{III.26}$$

At first, we study under which conditions a solution to the equations (III.26) exists.

Lemma III.3. *Let pair* $(c^\top, A)$ *be observable, with* $A, \Lambda \in \mathbb{R}^{n\times n}$ *cyclic and* $c, \theta \in \mathbb{R}^n$*. Then, equations* (III.26) *have a solution for any* θ *in the unknown* $\Pi \in \mathbb{R}^{n\times n}$ *if and only if* A, Λ *are similar.* ♙

The proof of Lemma III.3 is given for its dual Lemma III.5, in Appendix V.5.3.
Notice that, by substituting Λ with $\mathscr{A} + \mathscr{L}\mathscr{C}$, the above lemma underlines the importance of the plant observability. In fact, since we require $\sigma(A) \subset \sigma(\mathscr{A} + \mathscr{L}\mathscr{C})$ to obtain a full row rank Π (see Lemma III.2), the map $\mathscr{C}$ has to be chosen to place the eigenvalues of $\mathscr{A} + \mathscr{L}\mathscr{C}$. Observability of the true plant ensures that this choice for the output map does not compromise the solvability of the problem - at least, in the simplified system (III.26).
Next, we prove uniqueness of the solution, and we give an explicit expression for it.

Lemma III.4. *If a solution to* (III.26) *exists, it is unique and it is given by*

$$\Pi = O^{-1}_{c^\top, A} O_{\theta^\top, \Lambda}, \tag{III.27}$$

where $O_{c^\top,A}, O_{\theta^\top,\Lambda}$ *are the observability matrices of pairs* $(c^\top, A)$ *and* $(\theta^\top, \Lambda)$. ♙

The proof of Lemma III.4 is given for its dual Lemma III.6, in Appendix V.5.4.

The system $\Pi\Lambda = A\Pi$, $\Pi\ell = b$

Now, we pass to the dual simplified framework, namely, we drop the requirement to match the output map and we focus on single-input systems. This analysis is useful to characterize the importance of the controllability properties of the gazer matrices $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$. Letting $A, \Lambda \in \mathbb{R}^{n\times n}$ and $b, \ell \in \mathbb{R}^n$, we are interested in solving in Π the system of equations

$$\begin{cases} \Pi\Lambda = A\Pi \\ \Pi\ell = b. \end{cases} \tag{III.28}$$

At first, we study under which conditions a solution to the equations (III.28) exists.

Lemma III.5. *Let pair (Λ, ℓ) be controllable, with $A, \Lambda \in \mathbb{R}^{n\times n}$ cyclic and $b, \ell \in \mathbb{R}^n$. Then, equations* (III.28) *have a solution for any b in the unknown $\Pi \in \mathbb{R}^{n\times n}$ if and only if A, Λ are similar.* ♗

The proof of Lemma III.5 is in Appendix V.5.3.
Notice that, by substituting Λ with $\mathscr{A} + \mathscr{L}\mathscr{C}$ and ℓ with $\mathscr{B}$, the above lemma underlines the importance of the controllability properties of the pair $(\mathscr{A} + \mathscr{L}\mathscr{C}, \mathscr{B})$. In fact, this controllability ensures the possibility of representing *any* input entering the plant.
We remark that, whilst similarity of A, Λ guarantees the existence of a solution to system (III.28) for any b, this condition is not necessary for the existence of a solution for some specific b, and we present here a counterexample.

Example III.1. Choose the matrices

$$\Lambda = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \tag{III.29}$$

with $b = (0, 1)$ and $\ell = (1, 1)$. It can be checked that the two spectra are given by $\sigma(A) = \{1, 2\}$, $\sigma(\Lambda) = \{0, 2\}$ and

$$\Pi = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{III.30}$$

is a solution to (III.28). ◇

Next, we prove uniqueness of the solution, and we give an explicit expression for it.

Lemma III.6. *If a solution to* (III.28) *exists, it is unique and it is given by*

$$\Pi = R_{A,b} R_{\Lambda,\ell}^{-1}, \tag{III.31}$$

where $R_{A,b}$, $R_{\Lambda,\ell}$ are the reachability matrices of pairs (A, b) and (Λ, ℓ). ♗

The proof of Lemma III.6 is in Appendix V.5.4.
At last, we characterize an interesting set of pairs (A, b) which *require* the similarity of two matrices Λ, A for which similarity is necessary.

Lemma III.7. *Let the pair* (A, b) *in Lemma III.5 be controllable. The system* (III.28) *has a solution if and only if* A, Λ *are similar. Furthermore, the unique solution is full rank.* ♙

The proof of Lemma III.7 is in Appendix V.5.5.

Remark III.6. In case $\Lambda \in \mathbb{R}^{r\times r}$, $\ell \in \mathbb{R}^r$ with $r > n$, in order to guarantee the existence of a solution to (III.28) it is sufficient to ask for $\sigma(A) \subset \sigma(\Lambda)$. The expression of the solution is then modified into

$$\Pi = \begin{bmatrix} b & Ab & \dots & A^{r-1}b \end{bmatrix} R_{\Lambda,\ell}. \tag{III.32}$$

◇

In the following, we will show how to combine these result to propose design gazers at first for the SISO case, and then for the MIMO case.

III.3.3 A gazer for SISO systems

We begin the discussion by showing how filters which have classically been seen as adaptive observers (see, e.g., [162, Eq. (4.29)] or [5, Eq. (6)] for a more specific design) do actually implement gazers in filter form. Consider the SISO system

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= c^\top x, \end{aligned} \tag{III.33}$$

with $x \in \mathbb{R}^n$ and scalar input u and output y. Pick any $\Lambda \in \mathbb{R}^{r\times r}$, $\ell \in \mathbb{R}^r$ with $r \geq n$ and such that the pair (Λ, ℓ) is controllable. Consider the gazer in filter form with state $\zeta = (\zeta_y, \zeta_u) \in \mathbb{R}^{2r}$ and dynamics given by

$$\begin{aligned} \dot{\zeta}_y &= \Lambda\zeta_y + \ell y \\ \dot{\zeta}_u &= \Lambda\zeta_u + \ell u. \end{aligned} \tag{III.34}$$

In other words, the matrices $\mathscr{A}, \mathscr{B}, \mathscr{L}$ of the gazer (III.9) are

$$\mathscr{A} = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \in \mathbb{R}^{2r\times 2r}, \quad \mathscr{B} = \begin{bmatrix} 0 \\ \ell \end{bmatrix} \in \mathbb{R}^{2r}, \quad \mathscr{L} = \begin{bmatrix} \ell \\ 0 \end{bmatrix} \in \mathbb{R}^{2r} \quad \mathscr{C} = \begin{bmatrix} \theta_y^\top & \theta_u^\top \end{bmatrix} \in \mathbb{R}^{2r}, \tag{III.35}$$

where $\mathscr{A}, \mathscr{B}, \mathscr{L}$ are known and $\mathscr{C}$ is unknown. We are now interested in finding a full row rank matrix $\Pi \in \mathbb{R}^{n\times 2r}$ satisfying equations (III.14).

Theorem III.1. *Let pairs* (Λ, ℓ) *be controllable and* $(c^\top, A)$ *observable, with* $A \in \mathbb{R}^{n\times n}$, $b, c \in \mathbb{R}^n$ *and*

$\Lambda \in \mathbb{R}^{r\times r}$, $\ell \in \mathbb{R}^r$, $r \geq n$. *Consider* $\mathscr{A}, \mathscr{B}, \mathscr{L}$ *given in* (III.35), *and the equations* (III.14), *namely,*

$$\Pi \begin{bmatrix} \Lambda + \ell\theta_y^\top & \ell\theta_u^\top \\ 0 & \Lambda \end{bmatrix} = A\Pi, \quad \Pi \begin{bmatrix} 0 \\ \ell \end{bmatrix} = b \\ \begin{bmatrix} \theta_y^\top & \theta_u^\top \end{bmatrix} = c^\top \Pi, \tag{III.36}$$

in the unknowns $\Pi = [\Pi_y \ \Pi_u] \in \mathbb{R}^{n\times 2r}$. *Let* $l \in \mathbb{R}^n$ *be a vector for which* $\sigma(A - lc^\top) \subseteq \sigma(\Lambda)$*; then,*

$$\begin{aligned} \Pi_y &= \begin{bmatrix} l & (A - lc^\top)l & \dots & (A - lc^\top)^{r-1} l \end{bmatrix} R_{\Lambda,\ell}^{-1} \\ \Pi_u &= \begin{bmatrix} b & (A - lc^\top)b & \dots & (A - lc^\top)^{r-1} b \end{bmatrix} R_{\Lambda,\ell}^{-1} \end{aligned} \tag{III.37}$$

is a solution of equation (III.36). *Furthermore, if* (A, b) *is controllable, or if* $\sigma(\Lambda) \cap \sigma(A) = \emptyset$, Π *is of full row rank.* ♕

The proof of Theorem III.1 is in Appendix V.5.6.
Notice that Λ, ℓ are known a priori and no system knowledge is required for their design (but for an upper bound on the dimension n of the system). Given Propositions III.2 and III.3, if Λ is also Hurwitz then $\mathscr{A}, \mathscr{B}, \mathscr{L}$ in (III.35) implement a gazer of (III.33) as per Definition III.1.
Algorithm 1 recaps the design procedure described so far.

III.3.4 A gazer for MIMO systems

To the authors' knowledge, the following design for MIMO system is new. An example of a similar, but less generic, design, can be found in [5, Eqns. $(8) - (9)$]. Consider the MIMO system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned} \tag{III.40}$$

with $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^m$ and output $y \in \mathbb{R}^p$. Pick any $\Lambda \in \mathbb{R}^{r\times r}, \ell \in \mathbb{R}^r$ with $r \geq n$ and such that the pair (Λ, ℓ) is controllable. Consider the gazer in filter form with state $\zeta = (\zeta_y^1, \dots, \zeta_y^p, \zeta_u^1, \dots, \zeta_u^m) \in \mathbb{R}^{(p+m)r}$ and dynamics given by

$$\begin{aligned} \dot{\zeta}_y^i &= \Lambda\zeta_y^i + \ell y^i, \quad i = 1, \dots, p \\ \dot{\zeta}_u^j &= \Lambda\zeta_u^j + \ell u^j, \quad j = 1, \dots, m. \end{aligned} \tag{III.41}$$

Algorithm 1 Gazer design for SISO systems

Plant

Dynamics:

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= c^\top x. \end{aligned} \tag{III.38}$$

System matrices:

- $A \in \mathbb{R}^{n\times n}, b, c \in \mathbb{R}^n$,
- $(c^\top, A)$ observable.

Gazer construction

Filters dimension: $r \geq n$
Filters matrices:

- $\Lambda \in \mathbb{R}^{r\times r}, \ell \in \mathbb{R}^r$,
- Λ Hurwitz,
- (Λ, ℓ) controllable.

Gazer implementation:

Filters state: $\zeta = (\zeta_y, \zeta_u) \in \mathbb{R}^{2r}$
Initialization: $\zeta(0) \in \mathbb{R}^{2r}$
Dynamics:

$$\begin{aligned} \dot{\zeta}_y &= \Lambda\zeta_y + \ell y \\ \dot{\zeta}_u &= \Lambda\zeta_u + \ell u. \end{aligned} \tag{III.39}$$

In other words, the matrices $\mathscr{A}, \mathscr{B}, \mathscr{L}$ of (III.9) are given by

$$\begin{aligned}
\mathscr{A} &\coloneqq \operatorname{diag}(\Lambda, \dots, \Lambda) \in \mathbb{R}^{(p+m)r\times(p+m)r} \\
\mathscr{B} &\coloneqq \begin{bmatrix} 0_{np\times m} \\ \operatorname{diag}(\ell, \dots, \ell) \end{bmatrix} \in \mathbb{R}^{(p+m)r\times m} \\
\mathscr{L} &\coloneqq \begin{bmatrix} \operatorname{diag}(\ell, \dots, \ell) \\ 0_{nm\times p} \end{bmatrix} \in \mathbb{R}^{(p+m)r\times p} \\
\mathscr{C} &\coloneqq \begin{bmatrix} \theta_y^{11,\top} & \dots & \theta_y^{1p,\top} & \theta_u^{11,\top} & \dots & \theta_u^{1m,\top} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \theta_y^{p1,\top} & \dots & \theta_y^{pp,\top} & \theta_u^{p1,\top} & \dots & \theta_u^{pm,\top} \end{bmatrix} \in \mathbb{R}^{p\times(p+m)r}.
\end{aligned} \tag{III.42}$$

We are now interested in finding a full row rank matrix $\Pi \in \mathbb{R}^{n\times(p+m)r}$ satisfying equations (III.14).

Theorem III.2. *Let pairs (Λ, ℓ) be controllable and (C, A) observable, with $A \in \mathbb{R}^{n\times n}$, $B = [b_1, \dots, b_m] \in$*

$\mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ *and* $\Lambda \in \mathbb{R}^{r\times r}$, $\ell \in \mathbb{R}^r$, $r \geq n$. *Consider* $\mathscr{A}, \mathscr{B}, \mathscr{L}$, *given in* (III.42), *and the equations* (III.14) *in the unknowns*

$$\Pi = \begin{bmatrix} \Pi_y^1 & \dots & \Pi_y^p & \Pi_u^1 & \dots & \Pi_u^m \end{bmatrix} \in \mathbb{R}^{n\times(p+m)r}, \tag{III.43}$$

$\Pi_y^i \in \mathbb{R}^{n\times r}$, $\Pi_u^j \in \mathbb{R}^{n\times r}$ *for* $i = 1, \dots, p$ *and* $j = 1, \dots, m$. *Let* $L = [l_1, \dots, l_p] \in \mathbb{R}^{n\times p}$ *be a vector for which* $A - LC$ *is cyclic and* $\sigma(A - LC) \subseteq \sigma(\Lambda)$; *then*

$$\begin{aligned} \Pi_y^i &= \begin{bmatrix} l_i & (A-LC)l_i & \dots & (A-LC)^{r-1}l_i \end{bmatrix} R_{\Lambda,\ell}^{-1}, \quad i = 1, \dots, p, \\ \Pi_u^j &= \begin{bmatrix} b_j & (A-LC)b_j & \dots & (A-LC)^{r-1}b_j \end{bmatrix} R_{\Lambda,\ell}^{-1}, \quad j = 1, \dots, m, \end{aligned} \tag{III.44}$$

is a solution of equation (III.14). *Furthermore, if* (A, B) *is controllable, or if* $\sigma(\Lambda) \cap \sigma(A) = \emptyset$, Π *is of full row rank.* ♕

The proof of Theorem III.2 is in Appendix V.5.7.
Also in this case, Λ, ℓ are known a priori and no system knowledge is required for their design (but for an upper bound on the dimension n of the system). Given Propositions III.2 and III.3, if Λ is also Hurwitz then $\mathscr{A}, \mathscr{B}, \mathscr{L}$ in (III.42) implement a gazer of (III.40) as per Definition III.1.
We conclude with a recap of the design procedure described so far in Algorithm 2.

III.4 Robust filtering for data-driven LMIs

We now show how to leverage the previously studied gazers to solve control problems in a purely data-driven way. Although the next sections will deal with both state and output feedback, for convenience, we illustrate the problem for the state-feedback scenario. Consider a linear time-invariant system of the form

$$\dot{x} = Ax + Bu, \tag{III.47}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, and A and B are unknown matrices of appropriate dimensions. For a given initial condition $x(0)$ and some input $u(t)$, suppose that the resulting input-state trajectory of (III.47) has been collected over an interval $[0, T]$, with $T > 0$. More specifically, suppose that the continuous-time dataset

$$(u(t), x(t)), \qquad \forall t \in [0, T], \tag{III.48}$$

is available. We are interested in finding an algorithm that uses (III.48) to compute a stabilizing controller for (III.47), without any prior knowledge of A and B. We present some preliminary notions related to the existing approaches in the literature. To recover the results of data-driven stabilization of discrete-time systems, algorithms developed in a continuous-time setting are based on collecting a finite batch of data of u, x, and $\dot{x}$ with a suitable sampling mechanism [25, 62, 146]. Given a fixed sampling time

Algorithm 2 Gazer for MIMO systems

Plant

Dynamics:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx. \end{aligned} \tag{III.45}$$

System matrices:

- $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times m}$,
- (C, A) observable.

Gazer construction

Filters dimension: $r \geq n$

Filters matrices:

- $\Lambda \in \mathbb{R}^{r\times r}, \ell \in \mathbb{R}^{r}$,
- Λ Hurwitz,
- (Λ, ℓ) controllable.

Gazer implementation:

Filters state: $\zeta = (\zeta_y^1, \dots, \zeta_y^p, \zeta_u^1, \dots, \zeta_u^m) \in \mathbb{R}^{(p+m)r}$

Initialization: $\zeta(0) \in \mathbb{R}^{(p+m)r}$

Dynamics:

$$\begin{aligned} \dot{\zeta}_y^i &= \Lambda\zeta_y^i + \ell y^i, \quad i = 1, \dots, p \\ \dot{\zeta}_u^j &= \Lambda\zeta_u^j + \ell u^j, \quad j = 1, \dots, m. \end{aligned} \tag{III.46}$$

$T_s \coloneqq T/N$, with $N \in \mathbb{N}$, the following batch is obtained:

$$\begin{aligned} U &\coloneqq \begin{bmatrix} u(0) & u(T_s) & \cdots & u((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{m\times N} \\ X &\coloneqq \begin{bmatrix} x(0) & x(T_s) & \cdots & x((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{n\times N} \\ \dot{X} &\coloneqq \begin{bmatrix} \dot{x}(0) & \dot{x}(T_s) & \cdots & \dot{x}((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{n\times N}. \end{aligned} \tag{III.49}$$

We introduce a key definition used in this chapter.

Definition III.2. *A data batch of the form* (III.49) *is excited if*

$$\operatorname{rank} \begin{bmatrix} X \\ U \end{bmatrix} = n + m. \tag{III.50}$$

♔

Provided that the gathered data (III.49) are excited, they can be used to construct a stabilizing feedback law of the form $u = Kx$ for system (III.47). In particular, it is possible to make the closed-loop system matrix $A + BK$ is Hurwitz by choosing

$$K = UQ(XQ)^{-1}, \tag{III.51}$$

where $Q \in \mathbb{R}^{N\times n}$ is any solution of the following LMI:

$$\begin{cases} \dot{X}Q + Q^\top \dot{X}^\top < 0 \\ XQ = Q^\top X^\top > 0. \end{cases} \tag{III.52}$$

This result follows mutatis mutandis from the discrete-time case; see [62, Thms. 2 and 3], [214, Thm. 17]. We remark that, in the discrete-time scenario, only the data X and U are needed to compute K [62]. Instead, the continuous-time framework requires $\dot{X}$, which cannot be reliably inferred from (III.48) if the data are corrupted by noise. Also, even in a noise-free scenario, approximation via finite differences leads to persistent errors in the dataset.

The direct data-driven control framework introduced in this chapter addresses the above issue. In particular, we define a new methodology that does not require the measurement of $\dot{x}$. The results are presented both in the state-feedback scenario of system (III.47) and the case of output feedback for SISO systems.

III.4.1 Filters and LMIs for state-feedback

In this section, we are interested in designing a stabilizing controller for system (III.47) under the following assumption.

Assumption III.1. *The pair* (A, B) *is controllable.*

To avoid the central challenge of having to measure $\dot{x}$, we propose a strategy that involves the design of a filter of x and u. This filter is not used to approximate $\dot{x}$ but to reconstruct a non-minimal realization of the plant (III.47) whose state and state derivative are accessible. Thus, our approach avoids the robustness issues originating from the computation of derivatives from noisy data.

Non-Minimal Realization of the Dynamics

Consider the following dynamical system, having input u and output $\xi \in \mathbb{R}^n$:

$$\begin{aligned} \dot{\zeta} &= \begin{bmatrix} A & B \\ 0 & -\lambda I_m \end{bmatrix} \zeta + \begin{bmatrix} 0 \\ \gamma I_m \end{bmatrix} u \\ \xi &= \frac{1}{\gamma} \begin{bmatrix} A + \lambda I_n & B \end{bmatrix} \zeta, \end{aligned} \tag{III.53}$$

where $\zeta \in \mathbb{R}^{n+m}$ is the state and λ and γ, with $\gamma \neq 0$, are constant scalar tuning gains. System (III.53) will be proved to be a gazer in feedback form obtained by compactly rewriting the filter form

$$\dot{\zeta} = -\lambda\zeta + \gamma \begin{bmatrix} \xi \\ u \end{bmatrix} \tag{III.54}$$

together with the output map

$$\xi = \frac{1}{\gamma} \begin{bmatrix} A + \lambda I_n & B \end{bmatrix} \zeta. \tag{III.55}$$

This filter, for $\lambda > 0$, acts as a low-pass filter of ξ and u. The relationship between systems (III.47) and (III.53) is provided in the next lemma.

Lemma III.8. *Under Assumption III.1, for all λ and all $\gamma \neq 0$, the matrix $\Pi = \gamma^{-1}[A + \lambda I \;\; B]$ is a full rank solution of the gazer equations* (III.14) *between the plant* (III.47) *and the system* (III.53). ♙

The proof of Lemma III.8 is provided in Appendix V.5.8.
In other words, Lemma III.8 states that all input-state trajectories of (III.47) are input-output trajectories of (III.53). It is useful to recognize the structural controllability properties of system (III.53), which are stronger than simple stabilizability.

Lemma III.9. *Under Assumption III.1, for all λ and all $\gamma \neq 0$, the pair*

$$\left(\begin{bmatrix} A & B \\ 0 & -\lambda I_m \end{bmatrix}, \begin{bmatrix} 0 \\ \gamma I_m \end{bmatrix} \right) \tag{III.56}$$

is controllable. ♙

The proof of Lemma III.9 is provided in Appendix V.5.9.

Controller Design

The proposed procedure, described in Algorithm 3, is based on the following key ideas:

i) Consider an input-state trajectory of system (III.47) of the form (III.57) (in the following page). Choose gains λ and γ such that, in addition to $\gamma \neq 0$ as before, also $\lambda > 0$. By Lemma III.8, data (III.57) can be seen as an input-output trajectory of system (III.53), with $\xi(t) = x(t)$.

ii) Since (III.53) is equivalent to (III.54), its behavior is simulated with (III.58), which is a low-pass filter of the data due to $\lambda > 0$ and can be interpreted as a gazer of (III.53) (see Definition III.1).

iii) The estimated non-minimal state $\hat{\zeta}$ and its derivative $\dot{\hat{\zeta}}$ can be used for a data-driven control strategy that exploits the realization (III.53) to stabilize the original plant (III.47).

Algorithm 3 Control Design from Input-State Data

Initialization:

Dataset:

$$(u(t), x(t)), \qquad \forall t \in [0, T]. \tag{III.57}$$

Gains for tuning: $\lambda > 0, \gamma \neq 0, T_s > 0$.

Data Batches Construction:

Filter of the data:

$$\dot{\hat{\zeta}}(t) = -\lambda \hat{\zeta}(t) + \gamma \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad \forall t \in [0, T]. \tag{III.58}$$

Initialization: $\hat{\zeta}(0) = 0$.

Sampled data batches:

$$\begin{aligned} U &:= \begin{bmatrix} u(0) & u(T_s) & \cdots & u((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{m \times N} \\ Z &:= \begin{bmatrix} \hat{\zeta}(0) & \hat{\zeta}(T_s) & \cdots & \hat{\zeta}((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{(n+m) \times N} \\ \dot{Z} &:= \begin{bmatrix} \dot{\hat{\zeta}}(0) & \dot{\hat{\zeta}}(T_s) & \cdots & \dot{\hat{\zeta}}((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{(n+m) \times N} \\ E &:= \begin{bmatrix} x(0) \; e^{-\lambda T_s} x(0) & \cdots & e^{-\lambda (N-1) T_s} x(0) \end{bmatrix} \in \mathbb{R}^{n \times N}. \end{aligned} \tag{III.59}$$

Controller design:

LMI: find $Q \in \mathbb{R}^{N \times (n+m)}$ such that:

$$\begin{cases} \left(\dot{Z} - \left[\begin{smallmatrix} \gamma I_n \\ 0 \end{smallmatrix}\right] E\right) Q + Q^\top \left(\dot{Z} - \left[\begin{smallmatrix} \gamma I_n \\ 0 \end{smallmatrix}\right] E\right)^\top < 0 \\ ZQ = Q^\top Z^\top > 0. \end{cases} \tag{III.60}$$

Gain computation:

$$K = UQ(ZQ)^{-1}. \tag{III.61}$$

Control law (arbitrary initial conditions):

$$\dot{\hat{\zeta}}_c = -\lambda \hat{\zeta}_c + \gamma \begin{bmatrix} x \\ u \end{bmatrix}, \qquad u = K \hat{\zeta}_c. \tag{III.62}$$

Notice that, since A and B are unknown, $\hat{\zeta}(0)$ cannot be chosen such that $x(0) = \gamma^{-1} \begin{bmatrix} A + \lambda I_n & B \end{bmatrix} \hat{\zeta}(0)$. In other words, as it will be formalized later, for a generic initialization of $\hat{\zeta}$, it will hold that

$$x(t) \neq \gamma^{-1} \begin{bmatrix} A + \lambda I_n & B \end{bmatrix} \hat{\zeta}(t) \qquad \forall t \in [0, T]. \tag{III.63}$$

Therefore, the algorithm needs to account for the fact that the filtrations $\hat{\zeta}(t)$ converge only asymptotically to trajectories $\zeta(t)$ representation of the true state trajectories $x(t)$, and that T may not be a degree of freedom to reduce the error between $\zeta(t)$ and $\hat{\zeta}(t)$. This type of problem is not new to the field of signal processing, and we deal with it introducing a "denoising" term in the LMIs (III.60). We now make the above arguments precise. Consider the interconnection of plant (III.47) and filter (III.58), having states $(x, \hat{\zeta})$. To characterize the filter transient, define the mismatch error

$$\epsilon := x - \frac{1}{\gamma}\begin{bmatrix} A + \lambda I_n & B \end{bmatrix}\hat{\zeta}, \tag{III.64}$$

whose evolution can be computed from (III.47), (III.58) as follows:

$$\begin{aligned} \dot{\epsilon} &= Ax + Bu - \frac{1}{\gamma}\begin{bmatrix} A + I_n & B \end{bmatrix}\left(-\lambda\hat{\zeta} + \gamma\begin{bmatrix} x \\ u \end{bmatrix}\right) \\ &= -\lambda x + \frac{\lambda}{\gamma}\begin{bmatrix} A + I_n & B \end{bmatrix}\hat{\zeta} = -\lambda\epsilon. \end{aligned} \tag{III.65}$$

Using the change of coordinates (III.64), the interconnection can be represented with states $(\epsilon, \hat{\zeta})$ as:

$$\begin{aligned} \dot{\epsilon} &= -\lambda\epsilon \\ \dot{\hat{\zeta}} &= \underbrace{\begin{bmatrix} A & B \\ 0 & -\lambda I_m \end{bmatrix}}_{=:F}\hat{\zeta} + \underbrace{\begin{bmatrix} 0 \\ \gamma I_m \end{bmatrix}}_{=:G}u + \underbrace{\begin{bmatrix} \gamma I_n \\ 0 \end{bmatrix}}_{=:D}\epsilon, \end{aligned} \tag{III.66}$$

where the $\hat{\zeta}$-subsystem is a system with the same structure of (III.53) and subject the perturbation $D\epsilon$, which converges to 0 exponentially.
From (III.64) and choosing $\hat{\zeta}(0) = 0$ for simplicity, it holds that $\epsilon(0) = x(0)$. Thus, $\epsilon(t) = e^{-\lambda t}x(0)$ can be computed for every $t \in [0, T]$. The proposed procedure involves collecting N samples of u, $\hat{\zeta}$, $\dot{\hat{\zeta}}$, and ϵ as shown in (III.59), then solving LMI (III.60) and computing a control gain K from (III.61). The resulting controller (III.62) is a dynamic feedback law that incorporates the filter dynamics. Furthermore, when (III.62) is used online, its state $\hat{\zeta}_c$ can be initialized arbitrarily. We are ready to present the main result for Algorithm 3.

Theorem III.3. *Consider Algorithm 3 and let Assumption III.1 hold. Then:*

1. *LMI* (III.60) *is feasible if the batch* (III.59) *is excited, i.e.:*

$$\text{rank}\begin{bmatrix} Z \\ U \end{bmatrix} = n + 2m. \tag{III.67}$$

2. *For any solution Q of* (III.60)*, the gain K computed from* (III.61) *is such that $F + GK$ is Hurwitz. As a consequence, the origin $(x, \hat{\zeta}_c) = 0$ of the closed-loop interconnection of plant* (III.47) *and controller* (III.62) *is globally exponentially stable.*

♛

Proof. 1): Under Assumption III.1, (F, G) is controllable by Lemma III.9. Therefore, there exist matrices P, K satisfying:

$$\begin{cases} (F+GK)P + P^\top (F+GK)^\top \prec 0 \\ P = P^\top \succ 0. \end{cases} \tag{III.68}$$

Given any P, K satisfying (III.68), condition (III.67) implies that there exists a matrix M_K such that [62, Thm. 2]:

$$\begin{bmatrix} I_{n+m} \\ K \end{bmatrix} = \begin{bmatrix} Z \\ U \end{bmatrix} M_K. \tag{III.69}$$

Notice that $\dot{Z} = FZ + GU + DE$ from (III.66). Then, using (III.69), it holds that:

$$\begin{aligned} F + GK &= \begin{bmatrix} F & G \end{bmatrix} \begin{bmatrix} I_n \\ K \end{bmatrix} = \begin{bmatrix} F & G \end{bmatrix} \begin{bmatrix} Z \\ U \end{bmatrix} M_K \\ &= (\dot{Z} - DE) M_K. \end{aligned} \tag{III.70}$$

Combining (III.68) and (III.70), we obtain:

$$\begin{cases} (\dot{Z} - DE) M_K P + P^\top M_K^\top (\dot{Z} - DE)^\top \prec 0 \\ P = P^\top \succ 0. \end{cases} \tag{III.71}$$

Let $Q = M_K P$ and notice that (III.69) implies that $P = Z M_K P = ZQ$. Replacing these identities in (III.71), we obtain (III.60).

2): Suppose that there exists Q that satisfies (III.60). Since ZQ is symmetric and positive definite, Z has full row rank. Also, $Z^\dagger := Q(ZQ)^{-1}$ is a right inverse of Z. Using $\dot{Z} = FZ + GU + DE$ and the above properties, we have that

$$F = (\dot{Z} - DE - GU) Z^\dagger. \tag{III.72}$$

Using (III.61) and (III.72), we obtain:

$$F + GK = (\dot{Z} - DE - GU) Z^\dagger + GUZ^\dagger = (\dot{Z} - DE) Z^\dagger. \tag{III.73}$$

Let $P = ZQ$. Then, the first inequality of (III.60) reads as:

$$(\dot{Z} - DE) Q (ZQ)^{-1} P + P (ZQ)^{-1} Q^\top (\dot{Z} - DE)^\top \prec 0, \tag{III.74}$$

implying that $(\dot{Z} - DE) Q (ZQ)^{-1} = (\dot{Z} - DE) Z^\dagger = F + GK$ is Hurwitz.
To conclude the proof, we pass to the online implementation of the designed controller, namely, we study the interconnection of (III.47) and (III.62). Using the change of coordinates

$$\begin{bmatrix} x \\ \hat{\zeta}_c \end{bmatrix} \longmapsto \begin{bmatrix} \epsilon \\ \hat{\zeta}_c \end{bmatrix} := \begin{bmatrix} I_n & -\frac{1}{\gamma} \begin{bmatrix} (A + \lambda I_n) & B \end{bmatrix} \\ 0 & I_{n+m} \end{bmatrix} \begin{bmatrix} x \\ \hat{\zeta}_c \end{bmatrix} \tag{III.75}$$

as done in (III.64) and (III.66), and applying $u = K\hat{\zeta}_c$, this interconnection can be represented as

$$\begin{aligned}\dot{\epsilon} &= -\lambda\epsilon \\ \dot{\hat{\zeta}}_c &= (F + GK)\hat{\zeta}_c + D\epsilon.\end{aligned} \tag{III.76}$$

Since $F + GK$ is Hurwitz and $\lambda > 0$, global exponential stability of (III.76) follows from standard results for cascaded linear systems. ※

Remark III.7. The formal study on how to guarantee the excitation condition (III.67) is left for future studies. However, we provide some informal design guidelines:

i) System (III.53) is controllable by Lemma III.9, so if the input $u(t)$ can be chosen sufficiently rich as per Chapter II, to obtain a persistently exciting $(u(t), \zeta(t))$ (having a stable system is not necessary for the guaranteeing only a spanning conditions, see Remark II.12). Since $\hat{\zeta}(t) - \zeta(t) \to 0$ exponentially, for a dataset length $T > 0$ sufficiently high, there exists $\mu > 0$ such that:
$$\int_0^T \begin{bmatrix} u(\tau) \\ \hat{\zeta}(\tau) \end{bmatrix} \begin{bmatrix} u(\tau) \\ \hat{\zeta}(\tau) \end{bmatrix}^\top \mathrm{d}\tau \geq \mu I_{n+2m}. \tag{III.77}$$

ii) Under sufficient smoothness of the involved signals and sufficiently small sampling time $T_s > 0$ (see, e.g., [72, Lemma IV.3]), (III.77) implies (III.67).

◇

III.4.2 Filters and LMIs for output-feedback

To highlight the parallelism with the state-feedback scenario, we slightly abuse the notation of Section III.4.1 by adopting similar symbols. Consider a single-input single-output system of the form

$$\begin{aligned}\dot{x} &= Ax + bu \\ y &= c^\top x\end{aligned} \tag{III.78}$$

where $x \in \mathbb{R}^n$ is the unmeasured state, $u \in \mathbb{R}$ is the control input, $y \in \mathbb{R}$ is the measured output, and A, b, and c are matrices of appropriate dimensions whose values are unknown but satisfy the following assumption.

Assumption III.2. *The pair (A, b) is controllable and the pair $(c^\top, A)$ is observable. Furthermore, we assume n to be known.*

Remark III.8. We require the knowledge of n only for simplicity of presentation, and this is not a restrictive assumption for the following reasons. First, several techniques are available to estimate it from data [211]. Second, the knowledge of an upper bound $\bar{n} \geq n$ would be enough for the subsequent design of the filters, see Remark III.6. ◇

Compared to Section III.4.1, the additional challenge of this scenario is that only the output y is available instead of the state x. However, once again, it is possible to introduce a filter of the data that reconstructs a non-minimal realization of the plant, thus enabling the application of data-driven control without measuring derivatives.

Non-Minimal Realization of the Dynamics

Consider the SISO gazer designed in Algorithm 1 in feedback form, having input u and output $\xi \in \mathbb{R}$:

$$\begin{aligned}\dot{\zeta} &= \begin{bmatrix}\Lambda + \ell\theta_1^\top & \ell\theta_2^\top \\ 0 & \Lambda\end{bmatrix}\zeta + \begin{bmatrix}0 \\ \ell\end{bmatrix}u \\ \xi &= \begin{bmatrix}\theta_1^\top & \theta_2^\top\end{bmatrix}\zeta,\end{aligned} \tag{III.79}$$

where $\zeta \in \mathbb{R}^{2n}$ is the state , $\Lambda \in \mathbb{R}^{n\times n}$ and $\ell \in \mathbb{R}^n$ are constant tuning gains, and $\theta_1, \theta_2 \in \mathbb{R}^n$ are constant vectors (depending on A, b, c, Λ, and ℓ) whose values will be chosen to match the input-output behavior of systems (III.78) and (III.79). In filter form, system (III.79), also used in the literature of adaptive observers [162, Ch. 4], reads

$$\begin{aligned}\dot{\zeta}_1 &= \Lambda\zeta_1 + \ell\xi \\ \dot{\zeta}_2 &= \Lambda\zeta_2 + \ell u\end{aligned} \tag{III.80}$$

with output map

$$\xi = \theta_1^\top\zeta_1 + \theta_2^\top\zeta_2. \tag{III.81}$$

Recalling Theorem III.1 from the previous section, under Assumption III.2 it is possible to show that the gazer equations (III.14), namely, equations (III.36) for the above chosen matrices, have a full rank solution Π given in (III.37) (in the Arxiv version of this work [38], due to space limitations, we presented a more restrictive proof requiring distinct eigenvalues in Λ).
The next result is the equivalent of III.9 for the case of SISO plant.

Lemma III.10. *Under Assumption III.2, and given* $[\theta_1^\top, \theta_2^\top] = c^\top\Pi$, *where* Π *is the solution of equations* (III.36)*, the pair*

$$\left(\begin{bmatrix}\Lambda + \ell\theta_1^\top & \ell\theta_2^\top \\ 0 & \Lambda\end{bmatrix}, \begin{bmatrix}0 \\ \ell\end{bmatrix}\right) \tag{III.82}$$

is controllable. ♙

The proof of Lemma III.10 is provided in Appendix V.5.10.

Controller design

The procedure presented in Algorithm 4 follows similar ideas to those presented for the state-feedback scenario. Define a system of the form (III.84) replicating dynamics (III.80). Our main concern is,

Algorithm 4 Control Design from Input-Output Data

Initialization:

Dataset:

$$(u(t), y(t)), \qquad \forall t \in [0, T]. \tag{III.83}$$

Tuning: $\Lambda = \operatorname{diag}(-\lambda_1, \ldots, -\lambda_n)$, with $0 < \lambda_1 < \ldots < \lambda_n$, $\ell = (\gamma_1, \ldots, \gamma_n)$, with $\gamma_1, \ldots, \gamma_n \neq 0, T_s > 0$.

Data Batches Construction:

Filters of the data:

$$\dot{\hat{\zeta}}(t) = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \hat{\zeta}(t) + \begin{bmatrix} \ell & 0 \\ 0 & \ell \end{bmatrix} \begin{bmatrix} y(t) \\ u(t) \end{bmatrix}, \quad \forall t \in [0, T]. \tag{III.84}$$

Initialization: $\hat{\zeta}(0) = 0$.

Auxiliary dynamics:

$$\dot{\chi}(t) = \Lambda \chi(t), \qquad \forall t \in [0, T]. \tag{III.85}$$

Initialization: $\chi(0) = [1 \ \cdots \ 1]^\top$.

Sampled data batches:

$$\begin{aligned} U &:= \begin{bmatrix} u(0) & u(T_s) & \cdots & u((N-1)T_s) \end{bmatrix} \in \mathbb{R}^{1 \times N} \\ Z_a &:= \begin{bmatrix} \begin{bmatrix} \chi(0) \\ \hat{\zeta}(0) \end{bmatrix} & \cdots & \begin{bmatrix} \chi((N-1)T_s) \\ \hat{\zeta}((N-1)T_s) \end{bmatrix} \end{bmatrix} \in \mathbb{R}^{3n \times N} \\ \dot{Z}_a &:= \begin{bmatrix} \begin{bmatrix} \dot{\chi}(0) \\ \dot{\hat{\zeta}}(0) \end{bmatrix} & \cdots & \begin{bmatrix} \dot{\chi}((N-1)T_s) \\ \dot{\hat{\zeta}}((N-1)T_s) \end{bmatrix} \end{bmatrix} \in \mathbb{R}^{3n \times N}. \end{aligned} \tag{III.86}$$

Controller design:

LMI: find $Q \in \mathbb{R}^{N \times 3n}$ such that:

$$\begin{cases} \dot{Z}_a Q + Q^\top \dot{Z}_a^\top < 0 \\ Z_a Q = Q^\top Z_a^\top > 0. \end{cases} \tag{III.87}$$

Gain computation:

$$K = UQ(Z_a Q)^{-1} \begin{bmatrix} 0 \\ I_{2n} \end{bmatrix}. \tag{III.88}$$

Control law (arbitrary initial conditions):

$$\dot{\hat{\zeta}}_c = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \hat{\zeta}_c + \begin{bmatrix} \ell & 0 \\ 0 & \ell \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix}, \quad u = K \hat{\zeta}_c. \tag{III.89}$$

similarly to the state-feedback case, to find a way to compensate for the transient in which $\hat{\zeta}(t)$, obtained from a "generic" initialization, converges to $\zeta(t)$ such that $\Pi\zeta(t) = x(t)$. For doing this, we choose Λ diagonal with negative, different entries and ℓ such that the pair (Λ, ℓ) is controllable. As a consequence, Λ is Hurwitz and thus (III.84) is a low-pass filter of the input-output data (III.83). The proposed denoising procedure can be derived for more general choices of Λ and ℓ, although at the expense of increased notational burden. Consider Π, θ_1, θ_2 from Theorem III.1, then define:

$$\epsilon := x - \Pi\hat{\zeta}, \tag{III.90}$$

and note that $\epsilon(0) = x(0)$ since we choose $\hat{\zeta}(0) = 0$. The dynamics of ϵ are computed from (III.78), (III.36), and (III.84) as follows:

$$\begin{aligned} \dot{\epsilon} &= Ax + bu - \Pi \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} \hat{\zeta} + \Pi \begin{bmatrix} \ell & 0 \\ 0 & \ell \end{bmatrix} \begin{bmatrix} c^\top x \\ u \end{bmatrix} \\ &= (A - \Pi_1 \ell c^\top)\epsilon + \left(A\Pi - \Pi \begin{bmatrix} \Lambda + \ell\theta_1^\top & \ell\theta_2^\top \\ 0 & \Lambda \end{bmatrix} \right) \hat{\zeta} \\ &= H\Lambda H^{-1}\epsilon, \end{aligned} \tag{III.91}$$

where H is a non-singular matrix that exists due to Λ and $A - \Pi_1\ell c^\top$ being similar by construction (see the proof of Theorem III.1 in Appendix V.5.6). We can write the interconnection of plant (III.78) and filters (III.84) using the change of coordinates (III.90), leading to

$$\begin{aligned} \dot{\epsilon} &= H\Lambda H^{-1}\epsilon \\ \dot{\hat{\zeta}} &= \underbrace{\begin{bmatrix} \Lambda + \ell\theta_1^\top & \ell\theta_2^\top \\ 0 & \Lambda \end{bmatrix}}_{=:F} \hat{\zeta} + \underbrace{\begin{bmatrix} 0 \\ \ell \end{bmatrix}}_{=:g} u + \underbrace{\begin{bmatrix} \ell c^\top \\ 0 \end{bmatrix}}_{=:D} \epsilon, \end{aligned} \tag{III.92}$$

which shares the same structure of (III.66).

Contrary to Section III.4.1, $D\epsilon$ is not available in the output-feedback scenario. Since $\epsilon \to 0$ exponentially, a simple approach would be to sample u, $\hat{\zeta}$, and $\dot{\hat{\zeta}}$ after a sufficiently long time to make the perturbation $D\epsilon$ small enough. Instead, we propose an approach that compensates $D\epsilon$ exactly without any need for a waiting time.

From (III.92) and $\epsilon(0) = x(0)$, $\epsilon(t)$ can be computed as

$$\epsilon(t) = e^{H\Lambda H^{-1}t}\epsilon(0) = He^{\Lambda t}H^{-1}x(0) = L\chi(t), \tag{III.93}$$

where $L := ((H^{-1}x(0))^\top \otimes H)\,\mathrm{diag}(\mathrm{e}_1, \ldots, \mathrm{e}_n) \in \mathbb{R}^{n\times n}$ is an unknown matrix depending on H and $x(0)$ (where $(\mathrm{e}_1, \ldots, \mathrm{e}_n)$ is the standard orthonormal basis of $\mathbb{R}^n$), while $\chi(t) := [e^{-\lambda_1 t} \ \cdots \ e^{-\lambda_n t}]^\top \in \mathbb{R}^n$. Instead, note that $\chi(t)$ obeys dynamics $\dot{\chi} = \Lambda\chi(t)$, with $\chi(0) = [1 \ \cdots \ 1]^\top$, which is entirely known and implementable (since both $\chi(0)$ and Λ are known). The sequence $(u(t), \hat{\zeta}(t))$ obtained from (III.83), (III.84) satisfies for all $t \in [0, T]$ the following differential equation:

$$\begin{bmatrix} \dot{\chi} \\ \dot{\hat{\zeta}} \end{bmatrix} = \begin{bmatrix} \Lambda & 0 \\ DL & F \end{bmatrix} \begin{bmatrix} \chi \\ \hat{\zeta} \end{bmatrix} + \begin{bmatrix} 0 \\ g \end{bmatrix} u, \tag{III.94}$$

with initial conditions $\chi(0) = [1 \ \cdots \ 1]^\top$, $\hat{\zeta}(0) = 0$.
Note that the state and the state derivative of (III.94) are available. As a consequence, by implementing (III.84) and (III.85), we can sample u, $(\chi, \hat{\zeta})$, and $(\dot{\chi}, \dot{\hat{\zeta}})$, to obtain the batches (III.86). Then, LMI (III.87) can be used to compute a feedback law for plant (III.78). Suppose that there exists a matrix Q that solves LMI (III.87), then we can use the same arguments of Theorem III.3 to show that $[K_\chi \ K] = UQ(Z_\mathrm{a}Q)^{-1}$ is such that

$$\begin{bmatrix} \Lambda & 0 \\ DL + gK_\chi & F + gK \end{bmatrix} \tag{III.95}$$

is Hurwitz. Thus, K computed from (III.88) is such that $F + gK$ is Hurwitz. The resulting controller is given in (III.89).
We are ready to state the result corresponding to Theorem III.3. As the proof is identical, we omit it for brevity.

Theorem III.4. *Consider Algorithm 4 and let Assumption III.2 hold. Then:*

1. *LMI* (III.87) *is feasible if the batch* (III.86) *is excited, i.e.:*

$$\operatorname{rank} \begin{bmatrix} Z_a \\ U \end{bmatrix} = 3n + 1. \tag{III.96}$$

2. *For any solution Q of* (III.87)*, the gain K computed from* (III.88) *is such that $F + gK$ is Hurwitz. As a consequence, the origin $(x, \hat{\zeta}_\mathrm{c}) = 0$ of the closed-loop interconnection of plant* (III.78) *and controller* (III.89) *is globally exponentially stable.*

♕

Remark III.9. $V := [\chi(0) \ \cdots \ \chi((N-1)T_\mathrm{s})]$ is a Vandermonde matrix with roots $e^{-\lambda_1 T_\mathrm{s}}, \ldots, e^{-\lambda_n T_\mathrm{s}}$, so it has full row rank when $N \geq n$. Define $Z := [\hat{\zeta}(0) \ \cdots \ \hat{\zeta}((N-1)T_\mathrm{s})]$. Then, (III.96) holds if $[Z^\top \ U^\top]^\top$ has full row rank and is linearly independent from V. For the full rankness of $[Z^\top \ U^\top]^\top$, we refer to Remark III.7. To give an intuition on the second requirement, let $\hat{\zeta}(t)$ and $u(t)$ be the sum of p distinct sinusoids. From $\sin(\omega t) = (e^{\mathrm{i}\omega t} - e^{-\mathrm{i}\omega t})/(2\mathrm{i})$, $\cos(\omega t) = (e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t})/2$:

$$\begin{bmatrix} Z \\ U \end{bmatrix} = \Psi W = \Psi \begin{bmatrix} 1 & e^{\mathrm{i}\omega_1 T_\mathrm{s}} & \ldots & e^{\mathrm{i}\omega_1 (N-1) T_\mathrm{s}} \\ 1 & e^{-\mathrm{i}\omega_1 T_\mathrm{s}} & \ldots & e^{-\mathrm{i}\omega_1 (N-1) T_\mathrm{s}} \\ \vdots & \vdots & & \vdots \\ 1 & e^{\mathrm{i}\omega_p T_\mathrm{s}} & \ldots & e^{\mathrm{i}\omega_p (N-1) T_\mathrm{s}} \\ 1 & e^{-\mathrm{i}\omega_p T_\mathrm{s}} & \ldots & e^{-\mathrm{i}\omega_p (N-1) T_\mathrm{s}} \end{bmatrix} \tag{III.97}$$

where $\Psi \in \mathbb{C}^{(2n+1)\times 2p}$ and W is a Vandermonde matrix. Since W is linearly independent from V if $N \geq n + 2p$, each non-zero row of ΨW is linearly independent from V. ◇

Remark III.10. In order to pass to MIMO systems, one may substitute the SISO filters in (III.84) with MIMO ones (see Section III.3). However, in the MIMO case, the resulting nonminimal representation of the dynamics (the analogous of (III.79)) may be stabilizable but not controllable. So, further studies are necessary in order to guarantee the existence of solutions to the corresponding LMIs (given that certain states cannot be excited by the input but only through the initial condition). ◇

III.4.3 Numerical simulations

Algorithms 3 and 4 have been implemented in MATLAB using YALMIP [125] and MOSEK [12] to solve the LMIs. The developed code is available at the linked repository[1].

Design with Input-State Data

We consider the continuous-time linearized model of an unstable batch reactor given in [219], also used in [62] after time discretization. The system matrices are

$$
\begin{aligned}
A &= \begin{bmatrix} 1.38 & -0.2077 & 6.715 & -5.676 \\ -0.5814 & -4.29 & 0 & 0.675 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & -2.104 \end{bmatrix}, \\
B &= \begin{bmatrix} 0 & 0 \\ 5.679 & 0 \\ 1.136 & -3.146 \\ 1.136 & 0 \end{bmatrix},
\end{aligned}
\tag{III.98}
$$

where (A, B) is controllable and the eigenvalues of A are $\{-8.67, -5.06, 0.0635, 1.99\}$. We consider an exploration interval of length $T = 1.5$ s, where we apply a sum of 4 sinusoids to both entries of u and select 8 distinct frequencies to ensure informative data. We choose filter gains $\lambda = \gamma = 1$ and collect the data with sampling time $T_s = 0.1$ s.

Algorithm 3 has been extensively tested for random initial conditions $x(0)$, with each entry extracted from the uniform distribution $\mathcal{U}(-1, 1)$, returning a stabilizing controller each time. As an example, for the case $x(0) = [0.311 \;\; -0.6576 \;\; 0.4121 \;\; -0.9363]^\top$, we obtain the gain

$$
K = \begin{bmatrix} -1.507 & -18.69 & 0.155 & -0.681 & 2.925 & 0.79 \\ 17.45 & 0.224 & 44.06 & -36.37 & 1.09 & -3.518 \end{bmatrix}
\tag{III.99}
$$

which places the eigenvalues of the closed-loop matrix $F+GK$ in $\{-5.107 \pm 10.729\mathrm{i}, -1.238, -1.024 \pm 9.654\mathrm{i}, -0.759\}$.

[1] https://github.com/IMPACT4Mech/continuous-time_data-driven_control

Design with Input-Output Data

We consider a non-minimum phase SISO system having input-output behavior specified by the transfer function

$$c^\top(sI - A)^{-1}b = \frac{s-1}{s(s^2+2)}, \tag{III.100}$$

for which we choose a minimal realization in controllability canonical form. We perform exploration of $T = 2$ s with an input u given by the sum of 4 sinusoids at distinct frequencies. We choose filter parameters $\Lambda = \mathrm{diag}(-1, -2, -3)$, $\ell = (1, 2, 3)$ and sampling time $T_s = 0.1$ s.
Similar to the previous case, Algorithm 4 has been extensively tested with random initial conditions so that each entry of $x(0)$ is extracted from the uniform distribution $\mathcal{U}(-5, 5)$. In each test, the procedure returned a stabilizing controller. For the case $x(0) = [-3.9223 \quad 4.0631 \quad 3.7965]^\top$, we obtain

$$K = \begin{bmatrix} -0.508 & 3.208 & -2.392 & 0.001 & -0.577 & 1.055 \end{bmatrix} \tag{III.101}$$

which places the eigenvalues of $F + gK$ in $\{-2.028, -0.723 \pm 0.647\mathrm{i}, -0.22, -0.147 \pm 2.09\mathrm{i}\}$.

III.4.4 Future work

In this chapter, we have addressed the problem of obtaining a representation of the unavailable state of a linear system in a model-free fashion. First, we have proposed a kind of device, the "gazer", which, differently from classical observers, is not focused in obtaining the true state of the linear system, but it is only interested in *representing* the true state trajectory via a constant surjective map from the gazer and the plant trajetories. Sufficient conditions on gazer hyperparameters have been given to guarantee by design the satisfaction of the proposed requirements. Next, structural properties such as observability of the plant and controllability of the gazer have been studied in their role of guaranteeing the existence of a solution to the gazing problem. Leveraging all of these results, we have shown how classical filters can be seen as gazers in filter form, and we have derived a new design for MIMO systems. Future research will be focused on extending the proposed framework to nonlinear systems and to better understand the tight differences with the existing procedures for observer design.
Next, we addressed the problem of data-driven stabilization for continuous-time unknown linear systems by proposing a framework that combines signal filtering with LMIs. By filtering the input and state/output data, we obtained a non-minimal realization of the plant where the derivatives become accessible. Using LMIs inspired by those in [62], we computed stabilizing gains for dynamic filter-based controllers. This approach circumvents the need for state/output derivatives without resorting to noise-sensitive numerical techniques like finite differences. Future research will focus on extending the method to multi-input multi-output and nonlinear systems, to study the LMI feasibility conditions when the system is only stabilizable and to investigate the conditions that ensure exciting data.

Chapter IV

Robustly stable on-policy data-driven linear quadratic regulation

The problem of guaranteeing stability of an unknown system whilst controlling it has been historically addressed by the field of adaptive control [98, 112, 162, 191]. In particular, in the field of Model Reference Adaptive Control, adaptive control laws were designed for tracking desired *reference models*, namely, exosystems representing an ideal behavior. Whilst these approaches have nowadays somehow lost part of their relevance (in favour of a major interest in optimal control, reinforcement learning and other topics), their robustness has so far not been matched by new techniques, which are more concerned about reaching some form of optimality and often require large amounts of data. In this chapter, we try to combine the robustness guarantees of MRAC techniques with the optimality guarantees given by RL techniques. In particular, we address this problem for the most relevant benchmark in optimal control, namely, the Linear Quadratic Regulator, and we summarize in Table IV the generic problem.

This chapter is organized as follows. In Section IV.1 we give an overview of the scientific literature pertaining data-driven LQR techniques, highlighting the typical assumptions and requirements. Next, in Section IV.2, we give mathematical preliminaries on the ARE and DRE stability properties, we formalize the problem we want to solve (namely, the on-policy data-driven LQR) and a generic framework for the design of controller that solve the problem in a robust way. We also state the chapter contribution to the scientific community. In Section IV.3, we present MR-ARL, the architecture proposed for solving the robustly stable on-policy data-driven LQ problem in the case of structured uncertainties. We give a controller in the form of an algorithm; then we analyze its stabilization properties and we provide numerical simulations on a doubly-fed induction motor. All the proofs of the intermediate results can be found in Appendix V.6, whilst the proofs of the main proposition and theorems are given within the chapter. A preliminary conference paper [36] containing the results of this chapter has been presented to the Conference on Decision and Control (CDC) 2023, and a more comprehensive work [37] is currently available on Arxiv.

TABLE I **Robustly stable on-policy data-driven LQR**

Plant:

$$\dot{x} = Ax + Bu, \tag{IV.1}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^m$, and matrices A and B.
Infinite-horizon LQR: find an optimal control policy $\pi^\star : \mathbb{R}^n \to \mathbb{R}^m$ such that $u(t) = \pi^\star(x(t))$ minimizes, for all initial conditions $x_0 \in \mathbb{R}^n$, the following cost functional along the solutions of the plant:

$$J(x_0, u(\cdot)) := \int_0^\infty x(\tau)^\top Q x(\tau) + u(\tau)^\top R u(\tau) \mathrm{d}\tau, \tag{IV.2}$$

with tuning matrices $Q = Q^\top \geq 0$, $R = R^\top > 0$.
Problem: with A and B partially or totally unknown, find a dynamic controller of the form

$$\begin{aligned} \dot{z} &= \varphi(x, z, d) && \textit{learning dynamics} \\ u &= \pi(x, z, d) && \textit{applied control policy} \end{aligned} \tag{IV.3}$$

with $z \in \mathcal{Z} \subset \mathbb{R}^\ell$ are additional states and $d \in \mathbb{R}^q$ is a dither signal, such that the following properties are achieved.

1. ***Exploration:*** d probes the uncertainties of A and B.
2. ***Exploitation:*** map $x \mapsto \pi(x, z, 0)$ converges to the optimal policy $x \mapsto \pi^\star(x)$.
3. ***Robust stability:*** learning is enforced through robust asymptotic stability of the closed-loop system.

IV.1 Literature review

Reinforcement Learning is a machine learning field that emerged to perform optimization and decision-making by interacting with an environment without or with limited knowledge of its mathematical model [109, 177]. Over the past years, this field has been successfully applied to multiple domains, including computer games, biology, and economics and finance. RL has garnered the attention of the control engineering community, where it has been used to address optimal control in uncertain or model-free scenarios. Learning from system data aligns RL with principles found in adaptive control literature [10], which seeks to design dynamic controllers for regulation and tracking in the presence of model uncertainties. This work systematically investigates the connection between the fields of optimal and adaptive control, paving the way for a new RL paradigm that provides formal certificates of *robust closed-loop learning and control*, thereby leading to effective performance in real-world applications.
In particular, we focus on solving the infinite-horizon linear quadratic regulator (LQR) problem by developing an *on-policy data-driven algorithm* where data collection and optimization are done simultaneously by applying the learned policy to the actual system. The requirements of our framework are schematically presented above in Table I and later formalized in Section IV.2. A key distinctive feature of our proposed framework is the requirement of robust asymptotic stability for the whole closed-loop system including both the learning and control dynamics. This requirement, as elucidated in the subsequent sections, encapsulates the notion that the proven learning features in nominal cases must persist in perturbed scenarios, encompassing disturbances, measurement noise, slowly varying parameters, and sample-and-hold implementations. With a priori guarantees of effective closed-loop controller implementation, our framework is particularly tailored for safety-critical applications, such as collaborative robotics and aircraft control.
Motivated by the above discussion, we provide an overview of the literature pertaining to data-driven LQR, distinguishing between both so-called off-policy and on-policy approaches. Then, we recall model reference adaptive control, one of the main inspirations of the approach of this article.

Off-Policy and Offline Data-Driven LQR

Off-policy approaches involve finding the optimal policy whilst using a different policy for controlling the system. In this context, we find iterative methods, often inspired by the Kleinman algorithm, involving either parameter identification or direct estimate of the policy [102, 110, 127, 150, 169, 170, 222]. Typically, in these methods, the stabilization of the controlled system during the evolution of the learning algorithm is circumvented by assuming an initial stabilizing policy. However, as discussed in [235], there are situations where this assumption may be unrealistic due to plant uncertainties. The algorithm [29] does not need this assumption. On the other hand, offline approaches involve identifying the optimal policy from data batches [49, 57, 62, 63, 67, 182] or via system-level synthesis [65]. All these approaches differ from our setup, since they clearly separate the phase of data-collection and learning, where stability issues are simplified or not addressed, from the phase where the estimated policy is applied online.

On-Policy Data-Driven LQR

In case of several types of perturbations, offline, off-policy and model based-approaches either require to repeat in time multiple experiments to keep the controller updated, suppose the stabilization is already solved by some known policy or are intrinsecally suboptimal due to uncertainties and modeling difficulties. As compared to off-policy approaches, the on-policy paradigm solves these problems by adding the significant challenge of ensuring stability of the interconnection between the plant and the control/learning algorithm. Early attempts to address this setup are [43, 74, 102, 108, 149, 179, 180, 218], where the stabilizing feedback gain is updated at discrete iterations. However, the stability of the whole closed-loop system is not analyzed and an initial stabilizing policy is required, similar to off-policy approaches. More recently, in [88] the knowledge of an initial stabilizing policy is not needed; however, closed-loop stability and persistency of excitation are assumed. In [116], even though the need of an initial stabilizing policy remains, the authors carefully analyze in three different algorithms how to ensure closed-loop stability of input-affine systems during the learning phase. A data-driven approach to compute the initial gain is presented in [172]. To the best of our knowledge, [173] is the only work in the literature that provides stability guarantees without a stabilizing initialization, although the focus is on the learning dynamics and not the overall closed-loop system.

Model Reference Adaptive Control

We finally review the literature dealing with model reference adaptive control. The principle of this technique is to match the unknown system dynamics to a reference model with desired properties [98, 162, 205]. To ensure design feasibility, this stabilization technique requires the plant to satisfy constraints named *matching conditions*. A recent work combining MRAC and RL is [87], where RL techniques are used to find the optimal controller for a reference model based on nominal plant parameters. Then, MRAC is applied to assign the reference model to the real system. However, convergence to the desired policy is not proved and can only be ensured to a policy that is optimal for the reference model and not the true system.

IV.2 Problem setup

IV.2.1 Preliminaries: linear quadratic regulation

We start by introducing the basic concepts of LQR for system (IV.1) and the cost functional (IV.2). The infinite-horizon LQR problem involves finding a *control policy* $\pi^\star : \mathbb{R}^n \to \mathbb{R}^m$ such that applying $u(t) = \pi^\star(x(t))$ for all $t \in \mathbb{R}_{\geq 0}$ solves, for all initial conditions $x_0 \in \mathbb{R}^n$, the following optimal control problem:

$$\begin{aligned} \min_{u(\cdot)} J(x_0, u(\cdot)) &:= \int_0^\infty x(\tau)^\top Q x(\tau) + u(\tau)^\top R u(\tau) \mathrm{d}\tau \\ \text{subject to:} \quad \dot{x}(t) &= Ax(t) + Bu(t), \quad \forall t \in \mathbb{R}_{\geq 0}, \\ x(0) &= x_0. \end{aligned} \tag{IV.4}$$

Under the assumption that pair (A, B) be stabilizable and $(\sqrt{Q}, A)$ be detectable, the LQR problem (IV.4) is solved by the linear stabilizing policy:

$$\pi^\star(x) := K^\star x, \qquad K^\star := -R^{-1} B^\top P^\star, \tag{IV.5}$$

where $P^\star \in \mathcal{S}_0^n$ is the unique solution in $\mathcal{S}_0^n$ of the algebraic Riccati equation (ARE) [113, Thm. 13]:

$$A^\top P^\star + P^\star A - P^\star B R^{-1} B^\top P^\star + Q = 0. \tag{IV.6}$$

Additionally, $P^\star \in \mathcal{S}_+^n$ if pair $(\sqrt{Q}, A)$ is observable [227, Thm. 4.1]. We also recall that $P^\star$ specifies the *value function*, which is defined, for a given initial condition x, as the minimum of $J(x, u(\cdot))$:

$$V^\star(x) := \min_{u(\cdot)} J(x, u(\cdot)) := x^\top P^\star x. \tag{IV.7}$$

Consider the differential Riccati equation (DRE) with initial condition in the symmetric positive semidefinite cone:

$$\dot{P} = A^\top P + PA - PBR^{-1}B^\top P + Q, \quad P \in \mathcal{S}_0^n. \tag{IV.8}$$

As shown in [47, Thm. 1], under observability of $(\sqrt{Q}, A)$, the cone $\mathcal{S}_0^n$ is forward invariant under the motion induced by (IV.8). Namely, $P(0) \in \mathcal{S}_0^n \implies P(t) \in \mathcal{S}_0^n$, for all $t \in \mathbb{R}_{\geq 0}$. Next, if (A, B) is stabilizable and $(\sqrt{Q}, A)$ is detectable, the equilibrium point $P^\star$ is uniformly globally asymptotically stable (UGAS) for the differential equation (IV.8) [113, Thm. 15]. Furthermore, if (A, B) is controllable and $\sqrt{Q} > 0$, $P^\star$ is uniformly locally exponentially stable (ULES) [47, Thm. 4]. Formal results describing the stability properties of DRE (IV.8) are provided in [47, 113].

IV.2.2 Problem statement: robustly stable on-policy data-driven LQR

Following the discussion in the introduction, we now provide a rigorous formulation of the on-policy data-driven linear quadratic regulation (LQR) problem addressed in the chapter. Solving the LQR problem (IV.4) involves computing the solution $P^\star$ of ARE (IV.5), which depends on the matrices A and B of plant (IV.1). Therefore, if A and B are unknown or only partially known, it is necessary to resort to data-driven approaches based on acquiring the data of the state-input sequences $(x(t), u(t))$ (continuously or via batches).

In this work, we are interested in finding an on-policy data-driven algorithm, where data collection and learning are performed simultaneously by applying the learned policy. We now provide a novel rigorous

framework to formalize this problem so that its solution guarantees, by design, desirable learning and robust stability properties.
As anticipated in the introduction, the class of controllers that we seek are described by continuous-time dynamical systems of the form

$$\begin{aligned} \dot{z} &= \varphi(x, z, d) \\ u &= \pi(x, z, d) \end{aligned} \qquad z \in \mathcal{Z}, \tag{IV.9}$$

where z is the controller state, $\mathcal{Z} \subset \mathbb{R}^{\ell}$ is a closed set, $d \in \mathbb{R}^q$ is a uniformly bounded signal, named hereafter *dither*, while φ and π are maps that are locally Lipschitz in their arguments. To the algorithm (IV.9), we associate the *learning set* $\mathcal{L}$, defined as:

$$\mathcal{L} := \{z \in \mathcal{Z} : \pi(x, z, 0) = K^{\star}x, \ \forall x \in \mathbb{R}^n\}, \tag{IV.10}$$

which denotes the set of all controller states such that the control policy π coincides with the optimal policy $\pi^{\star}$ in (IV.5) whenever the dither d is turned off.
It is known that persistency of excitation of certain internal signals is necessary and sufficient for convergence of learning schemes [191]. Therefore, we take into account this problem by introducing the exogenous signal d in the control scheme (IV.9) with the specific role of guaranteeing PE of the signals of interest. To this regard, in the Arxiv version of this work [37], we use [98, Def. 5.2.1] to guarantee this property. However, in this thesis (since new results are available) we prefer to use the more generic results presented in Chapter II. Notice however that d may be used also to represent references for tracking or external disturbances. To ensure well-posedness of the problem formulation, from now on we consider a general class of signals d that can be generated by an autonomous dynamical system (*exosystem*) of the form

$$\begin{aligned} \dot{w} &= s(w) \\ d &= h(w) \end{aligned} \qquad w \in \mathcal{W}, \tag{IV.11}$$

where w is the exosystem state, $\mathcal{W} \subset \mathbb{R}^p$ is the set of admissible initial conditions $w(0)$, while s and h are locally Lipschitz maps. Since d is a bounded signal defined for all $t \in \mathbb{R}_{\geq 0}$, we suppose that the set $\mathcal{W}$ be compact and strongly forward invariant under the flow of (IV.11).

Remark IV.1. Exosystem (IV.11) is not implemented in the control solution but is used to represent the class of admissible signals d. Moreover, the results of the section hold if (IV.11) is replaced by a well-posed hybrid dynamical system [79] to include discontinuous dither signals, since the analysis in Section IV.3.2 does not require continuity of the dither $d(t)$. Here, we use a continuous-time exosystem to avoid an additional notational burden. ◇

The closed-loop system resulting from the interconnection of exosystem (IV.11), plant (IV.1), and controller (IV.9) is given by

$$\begin{aligned} \dot{w} &= s(w) \\ \dot{x} &= Ax + B\pi(x, z, h(w)) \qquad (w, x, z) \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}. \\ \dot{z} &= \varphi(x, z, h(w)) \end{aligned} \tag{IV.12}$$

We are ready to precisely state the requirements for controller (IV.9), which include a precise stability characterization for the closed-loop system (IV.12).

Definition IV.1. *We say that controller* (IV.9) solves the robustly stable on-policy data-driven LQR problem *if the learning set $\mathcal{L}$ in* (IV.10) *is non-empty and, for a chosen class of dither signals d generated by exosystem* (IV.11)*, there exists a compact attractor $\mathcal{A}$, satisfying*

$$\mathcal{A} \subset \mathcal{W} \times \mathbb{R}^n \times \mathcal{L}, \tag{IV.13}$$

that is asymptotically stable for the closed-loop system (IV.12). ♔

We show that any algorithm satisfying Definition IV.1 covers all of the design requirements stated in the introduction.

- **Exploration**: choosing the dither d determines the shape and the attractivity properties of $\mathcal{A}$, thus it ensures the necessary probing to estimate the optimal policy.
- **Exploitation:** since the projection of $\mathcal{A}$ in the z direction is a subset of the learning set $\mathcal{L}$, uniform attractivity of $\mathcal{A}$ (encoded in asymptotic stability) ensures $z \to \mathcal{L}$ and, thus, correct estimation of the optimal policy.
- **Robust stability:** under the regularity properties required for the controller and assumed for the exosystem, asymptotic stability of the attractor $\mathcal{A}$ is preserved (practically and semiglobally) under a broad range of non-vanishing perturbations [79, Ch. 7.2] such as disturbances, measurement noise, unmodeled dynamics, sample-and-hold implementations of the controller [186], and actuator dynamics [187].

Remark IV.2. In Definition IV.1, we do not specify the restrictions on the knowledge of A and B to cover a broad range of applications and solutions. However, the prior knowledge on the parametric uncertainties determines the design of φ, π, and $\mathcal{Z}$. Note that, if controller (IV.9) is not appropriately chosen, the learning set may be empty. ◊

Remark IV.3. The convergence of z to the learning set $\mathcal{L}$ in Definition IV.1 implies that the controlled plant becomes asymptotically:

$$\dot{x} = (A + BK^\star)x + \Delta(x, z, d), \tag{IV.14}$$

where $\Delta(x, z, d) := B(\pi(x, z, d) - K^\star x)$ vanishes in $d = 0$. Moreover, since $\Delta(x, z, d)$ is locally Lipschitz, it is bounded for all (x, z, w) in the compact attractor $\mathcal{A}$. As a consequence, the input-to-state stability of (IV.14) implies that

$$\limsup_{t\to\infty} |x(t)| \le \alpha(\limsup_{t\to\infty} |d(t)|), \tag{IV.15}$$

where α is a class $\mathcal{K}$ function depending on the system, the cost matrices Q and R, and the attractor $\mathcal{A}$. In other words, the ultimate bound of x directly depends on the amplitude of the injected dither d. ◇

IV.3 MR-ARL: Model Reference Adaptive Reinforcement Learning

IV.3.1 The algorithm

We now present a new control and learning approach, named *Model Reference Adaptive Reinforcement Learning* (MR-ARL), which satisfies Definition IV.1 in a scenario where $Q > 0$, B is known and A is subject to *structured uncertainties* characterized by the following assumptions.

Assumption IV.1. *There exists a closed convex set $C \subset \mathbb{R}^{n\times n}$ that is known for control design and is such that:*

1. *C has a non-empty interior and $A \in \mathrm{Int}(C)$;*
2. *$(\hat{A}, B)$ is controllable for all $\hat{A} \in C$.*

Remark IV.4. From [145, 173], it is known that if there exists A_0 such that (A_0, B) is controllable and $(\sqrt{Q}, A_0)$ is observable, then there exists a scalar $\rho > 0$ such that (A, B) is controllable and $(\sqrt{Q}, A)$ is observable for all A such that $|A - A_0| \le \rho$. Therefore, during implementation, a possible choice for the set C is a ball centered in a nominal value A_0 of A with radius ρ chosen to include all possible uncertainties, provided however that controllability is preserved in the set. ◇

Assumption IV.2. *Consider the linear map $\mathcal{B} : K \in \mathbb{R}^{m\times n} \longmapsto BK \in \mathbb{R}^{n\times n}$, where B is the input matrix in* (IV.1). *For some known $A_0 \in C$, it holds that*

$$A_0 - A \in \mathrm{Im}(\mathcal{B}), \tag{IV.16}$$

where A is the state matrix in (IV.1).

Assumption IV.2 is an alternative formulation of the matching conditions used in the MRAC literature [205]. In other words, this assumption requires the possibility of matching uncertainties in the systems via a feedback gain.
Notice indeed that that for any $\hat{A}$ such that $\hat{A} - A_0 \in \mathrm{Im}(\mathcal{B})$, Assumption IV.2 implies that there exists $K_a \in \mathbb{R}^{m\times n}$ such that

$$\hat{A} - A = BK_a. \tag{IV.17}$$

Given the two sets C and $A_0 + \mathrm{Im}(\mathcal{B})$, we are interested in all matrices $\hat{A} \in \Theta$, where

$$\Theta := C \cap (A_0 + \mathrm{Im}(\mathcal{B})) \tag{IV.18}$$

since they can be used to build reference models. Notice that if B is full row rank, this assumption holds for any $A_0 \in \mathbb{R}^n$.

Remark IV.5. In our framework, we suppose B is known; whilst this assumption can be relaxed by requiring only the knowledge of a certain structure in B [204, Pag. 375], [54] (and without introducing any differences in the Assumptions IV.1 and IV.2), to the authors' knowledge no MRAC technique has so far solved in a purely data-driven way the stabilization problem. Since the subsequent analysis of this algorithm is already quite complicated, we leave for future work possible relaxation of this assumption. ◊

Remark IV.6. This assumption can be checked in a purely data-driven way. Assume, for simplicity, to be able to measure the derivative $\dot{x}$ (if this is not the case, the same reasoning may applied to appropriate filters of x). Build the signal

$$\varepsilon := \dot{x} - (A_0 x + Bu) = (A - A_0)x. \tag{IV.19}$$

Under a sufficiently rich input and if the system is controllable, $x(t)$ spans in time all directions, and collecting $N \geq n$ samples of $\varepsilon(t), x(t)$, we get

$$\underbrace{\begin{bmatrix}\varepsilon(t_1) & \dots & \varepsilon(t_N)\end{bmatrix}}_{=:E} = (A - A_0)\underbrace{\begin{bmatrix}x(t_1) & \dots & x(t_N)\end{bmatrix}}_{X}, \tag{IV.20}$$

from which Assumption IV.2 holds if $EX^{\dagger} \in \mathrm{Im}(\mathcal{B})$, given the full row rankness of X. ◊

Our controller is conceived as an *actor-critic* modular architecture where a *reference model* bridges the two parts of the design. The resulting structure is an MRAC where the reference model is continuously updated with value iteration, thus we aptly name it *Model Reference Adaptive Reinforcement Learning*. We introduce the building blocks of the design.

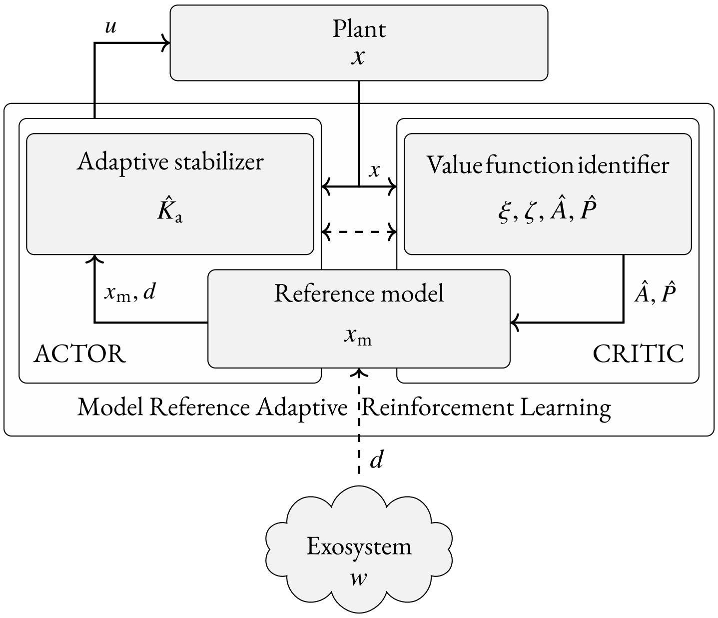

Figure IV.1: Block scheme of the *Model Reference Adaptive Reinforcement Learning*.

- **Critic:** this block performs *data-driven value function identification* to build an optimal and asymptotically stable reference model. In particular, a gradient identifier computes an estimate $\hat{A} \in \Theta$ of A that is used to obtain an estimate $\hat{P}$ of the solution $P^\star$ of ARE (IV.6). In this respect, Assumption IV.1 guarantees that for any estimate $\hat{A}$ the computation of $\hat{P}$ is feasible. Then, $\hat{A}$ and the optimal gain estimate $-R^{-1}B^\top \hat{P}$ are used to build a reference model having state matrix $\hat{A} - BR^{-1}B^\top \hat{P}$. As input to the reference model, we consider a dither d with sufficient richness properties to ensure convergence to the true system parameters and to the optimal policy.

- **Actor:** this block assigns the input to the plant to *adaptively track the reference model*. During the transient, the feedback gain $-R^{-1}B^\top \hat{P}$ may be not stabilizing for the real system. For this reason, the actor introduces in the control law an additional adaptive feedback gain $\hat{K}_a$ to cancel the mismatch between the estimated matrix $\hat{A}$ and the real A. Canceling such a mismatch is possible due to Assumption IV.2.

See Fig. IV.1 for a block scheme of Model Reference Adaptive Reinforcement Learning. The full description of the design is presented in Algorithm 5 and discussed in detail in the next subsections.

Algorithm 5 MR-ARL

Initialization:

$\xi(0), \zeta(0) \in \mathbb{R}^n$ filter initial conditions

$\hat{P}(0) \in \boldsymbol{S}_0^n$, $\hat{A}(0) \in \Theta$, with Θ from Eq. IV.18

$\lambda, \gamma, \nu, g, \mu > 0$ design gains

$d(t) \in C_{\text{SR}}(\mathbb{L}^{\text{C}}_{\text{x},>1})$, e.g. as per Lemma II.9

Repeat:

Swapping filters:

$$\dot{\xi} = -\lambda\xi + x, \qquad \dot{\zeta} = -\lambda(x + \zeta) - Bu \tag{IV.21}$$

Identifier dynamics:

$$\dot{\hat{A}} = \mathcal{P}_{\hat{A}\in C}\left\{-\gamma BB^\dagger \frac{\epsilon\xi^\top}{1 + \nu|\xi||\epsilon|}\right\}, \quad \epsilon := \hat{A}\xi - (x + \zeta) \tag{IV.22}$$

Value iteration:

$$\dot{\hat{P}} = g\left(\hat{A}^\top\hat{P} + \hat{P}\hat{A} - \hat{P}BR^{-1}B^\top\hat{P} + Q\right) \tag{IV.23}$$

Reference model:

$$\dot{x}_{\text{m}} = (\hat{A} - BR^{-1}B^\top\hat{P})x_{\text{m}} + Bd \tag{IV.24}$$

Adaptive gain dynamics:

$$\dot{\hat{K}}_{\text{a}} = -\mu B^\top\hat{P}(x - x_{\text{m}})x^\top + B^\dagger\dot{\hat{A}} \tag{IV.25}$$

System input:

$$u = -R^{-1}B^\top\hat{P}x + \hat{K}_{\text{a}}x + d \tag{IV.26}$$

Critic: Value Function Identifier

In this subsection, we build a continuous-time identifier of $P^\star$ based on the estimation of matrix A. Given the structure of system (IV.1), we compute an estimate $\hat{A} \in \Theta$ of A by designing a swapping filter [112, Ch. 6] of the form (IV.21), with $\lambda > 0$ a scalar gain for tuning the filter time constant. The role of this filter is to substitute the knowledge of systems derivatives by constructing, with available quantities and asymptotically, an algebraic relation that can be used to estimate A. Using the filter states, define the *prediction error*

$$\epsilon := \hat{A}\xi - (x + \zeta), \tag{IV.27}$$

which we can rewrite as $\epsilon = (\hat{A} - A)\xi + \tilde{\epsilon}$, where

$$\tilde{\epsilon} := A\xi - (x + \zeta) \tag{IV.28}$$

is an error signal that is shown in Section IV.3.2 to converge exponentially to zero. Notice that if $\tilde{\epsilon} = 0$, we obtain the algebraic relation $x + \zeta = A\xi$, which can be used to identify A. For doing so, we use the normalized projected gradient descent algorithm (IV.22) to update the estimate $\hat{A}$ according to the prediction error $(\hat{A} - A)\xi + \tilde{\epsilon}$. In (IV.22), parameters $\gamma > 0$ and $\nu > 0$ are scalar gains, while the multiplicative term $BB^{\dagger}$ is a projection onto $\mathrm{Im}(\mathcal{B})$ [80, Sec. 5.5.4]. This projection is needed to ensure, given any initialization $\hat{A}(0) \in A_0 + \mathrm{Im}(\mathcal{B})$, that the estimate $\hat{A}(t)$ never leaves this hyperplane. Finally, $\mathcal{P}_{\hat{A}\in C}\{\cdot\}$ is a Lipschitz continuous parameter projection operator, whose expression is provided, e.g., in [112, Appendix E] and depends on the shape of the set C. For convenience, we report in the Appendix V.6.7 an expression for $\mathcal{P}_{\hat{A}\in C}\{\cdot\}$ in case C is a closed ball about a nominal parameter.
Given the estimate $\hat{A}$, we are interested in computing the matrix $\hat{P} \in S_+^n$ that solves the ARE

$$\mathcal{R}(P, \hat{A}) := \hat{A}^\top P + P\hat{A} - PBR^{-1}B^\top P + Q = 0. \tag{IV.29}$$

From [36], such a matrix could be obtained by computing the map $\mathcal{P}(\hat{A})$ that solves (IV.29) for each $\hat{A}$, i.e., such that:

$$\mathcal{R}(\mathcal{P}(\hat{A}), \hat{A}) = 0, \quad \text{for all } \hat{A} \in \Theta. \tag{IV.30}$$

For simplicity in the implementation and inspired by [173], in Algorithm 5, we compute $\hat{P}$ via the dynamical system (IV.23), which is a DRE rescaled by the tuning gain $g > 0$. Notice that, if $\hat{A}$ is constant, then the solution $\hat{P}$ of (IV.23) converges to $\mathcal{P}(\hat{A})$.

Remark IV.7. Assumption IV.1 guarantees that, for each $\hat{A} \in \Theta$, $\mathcal{P}(\hat{A})$ exists, is unique, and positive definite for any $\hat{A} \in \Theta$ (see Section IV.2.1). Although stabilizability of $(\hat{A}, B)$ would be sufficient in Assumption IV.1 for the solvability of $\mathcal{R}(P, \hat{A}) = 0$, controllability is essential to guarantee convergence of the identifier under sufficient richness of the dither $d(t)$ [161, Lemma 5]. Moreover, $Q > 0$ ensures that $\mathcal{P}(\hat{A})$ has exponential stability properties for DRE dynamics (IV.23) (see Section IV.2.1). ◇

Remark IV.8. From Assumption IV.1 and the parameter projection in (IV.22), matrix $\hat{A} - BR^{-1}B^\top\mathcal{P}(\hat{A})$ is Hurwitz by design, being the optimal closed-loop. Therefore, if $\hat{A}$ converges to a constant matrix, (IV.23) ensures that $\hat{A} - BR^{-1}B^\top\hat{P}$ converges to a Hurwitz matrix. ◇

Reference Model

Given the estimate $\hat{P}$ of $P^\star$, we design a reference model for system (IV.1). The reference model has to embed all the properties required for the plant, i.e., robust stability, optimality and persistency of excitation. To these aims, consider system (IV.24), where $x_{\mathrm{m}} \in \mathbb{R}^n$ is the reference model state. We embed the stability and optimality properties through $\hat{A} - BR^{-1}B^\top\hat{P}$, which is designed to converge to a Hurwitz matrix. Different from classic MRAC, the state matrix $\hat{A} - BR^{-1}B^\top\hat{P}$ of the reference

model (IV.24) is not constant but time-varying as it depends on the estimates $\hat{A}$ and $\hat{P}$. This property leads to an adaptive design where the known-plant stabilizing gains are time-varying.
Finally, we embed the persistency of excitation properties through dither $d \in \mathbb{R}^m$, which is chosen through any characterization of sufficient richness, e.g., $\mathrm{D}^n(\mathrm{d}) \in \Omega^{\mathrm{C}}_{nm}$ (as per Theorem II.7) or via sinusoids (as per Theorem II.10).

Remark IV.9. We model $d(t)$ as the output of an exosystem of the form (IV.11). It is not necessary to actually implement the exosystem as part of the algorithm, as we show in the numerical example. ◇

Actor: Model Reference Adaptive Controller

Given the reference model (IV.24), we design an adaptive controller for system (IV.1). Define the tracking error $e \coloneqq x - x_{\mathrm{m}}$ and compute its time derivative from (IV.1), (IV.24) as

$$\begin{aligned} \dot{e} &= Ax + Bu - (\hat{A} - BR^{-1}B^\top \hat{P})(x - e) - Bd \\ &= (\hat{A} - BR^{-1}B^\top \hat{P})e \\ &\quad + (A - \hat{A})x + B(u + R^{-1}B^\top \hat{P}x - d). \end{aligned} \tag{IV.31}$$

To ensure that the plant (IV.1) asymptotically copies the behavior of the reference model (IV.24), i.e., $e(t) \to 0$, we exploit the fact that $\hat{A} \in \Theta$, which is guaranteed by Assumption IV.2 and by the projection $BB^\dagger$ in the update law. In particular, from (IV.17), for each $\hat{A} \in \Theta$ there exists $K_{\mathrm{a}}(\hat{A}) \in \mathbb{R}^{m\times n}$ such that

$$\hat{A} - A = BK_{\mathrm{a}}(\hat{A}). \tag{IV.32}$$

More specifically, we can ensure that map $K_{\mathrm{a}}(\hat{A})$ is smooth in $\hat{A}$ by choosing:

$$K_{\mathrm{a}}(\hat{A}) \coloneqq B^\dagger(\hat{A} - A). \tag{IV.33}$$

This way, (IV.31) becomes

$$\dot{e} = (\hat{A} - BR^{-1}B^\top \hat{P})e + B(u + R^{-1}B^\top \hat{P}x - K_{\mathrm{a}}(\hat{A})x - d), \tag{IV.34}$$

suggesting a control law of the form

$$u \coloneqq -R^{-1}B^\top \hat{P}x + K_{\mathrm{a}}(\hat{A})x + d \tag{IV.35}$$

if the plant dynamics were known. However, $K_{\mathrm{a}}(\hat{A})$ is unavailable for design as it depends also on A, as highlighted in (IV.32), thus we consider the certainty-equivalence-based adaptive controller given in (IV.26), where $K_{\mathrm{a}}(\hat{A})$ is replaced by the adaptive gain $\hat{K}_{\mathrm{a}}$, driven by the adaptive law (IV.25) where $\mu > 0$ is a scalar gain. The first term in the adaptive law (IV.25) is a standard update to ensure the error e goes asymptotically to zero in a framework where the model mismatch is constant. However, since $\hat{A}$ is continuously updated by identifier (IV.22), the second term in the update law takes into account the time-varying mismatch.

Main Result

We now provide the main results of this work, where we show that Model Reference Adaptive Reinforcement Learning solves the robustly stable on-policy data-driven LQR problem as per Definition IV.1. To simplify the algorithm analysis, we follow a singular perturbation approach [206], namely, we give a first result supposing to have a DRE dynamics in (IV.23) infinitely faster than the rest of the system (*reduced-order system*), i.e., supposing $\hat{P}(t) = \mathcal{P}(A(t))$ as in (IV.30) at each t. For this reason, we mark the results for the reduced-order system with a subscript s to highlight its slow dynamics. We then leverage on this intermediate result to prove that Algorithm 5 solves the robustly stable on-policy data-driven LQR problem.
Consider the Model Reference Adaptive Reinforcement Learning with ARE implementation of $\hat{P}$ as in (IV.30). Following the notation of Section IV.2.1, the controller obtained by combining the value function identifier (IV.21), (IV.22), (IV.30), reference model (IV.24), and the adaptive stabilizer (IV.25), (IV.26) is in the form (IV.9), with state

$$z_{\mathrm{s}} \coloneqq (\xi, \zeta, \hat{A}, x_{\mathrm{m}}, \hat{K}_{\mathrm{a}}) \in \mathcal{Z}_{\mathrm{s}} \coloneqq \mathbb{R}^{2n} \times \Theta \times \mathbb{R}^{n} \times \mathbb{R}^{n\times n}, \tag{IV.36}$$

and output policy

$$\pi(x, z_{\mathrm{s}}, d) \coloneqq (\hat{K}_{\mathrm{a}} - BR^{-1}B^{\top}\mathcal{P}(\hat{A}))x + d, \tag{IV.37}$$

from which it follows that the learning set $\mathcal{L}$ in (IV.10) is non-empty and given by

$$\mathcal{L}_{\mathrm{s}} \coloneqq \{z_{\mathrm{s}} \in \mathcal{Z}_{\mathrm{s}} : \hat{K}_{\mathrm{a}} - BR^{-1}B^{\top}\mathcal{P}(\hat{A}) = -BR^{-1}B^{\top}P^{\star}\}. \tag{IV.38}$$

In particular, we recall that our algorithm aims at reaching the learning set by ensuring $\hat{A} \to A$ (hence $\hat{P} \equiv \mathcal{P}(\hat{A}) \to P^{\star}$ by continuity of the map $\mathcal{P}(\hat{A})$) and $\hat{K}_{\mathrm{a}} \to 0$. The following result shows that, with $\gamma > 0$ sufficiently small, the Model Reference Adaptive Reinforcement Learning with ARE implementation of $\hat{P}$ solves the robustly stable on-policy data-driven LQR problem.

Theorem IV.1. *Consider the closed-loop system given by the interconnection of plant* (IV.1) *and the controller of Algorithm 5, with $\hat{P}(t) = \mathcal{P}(\hat{A}(t))$ for all t and $\mathcal{P}(\hat{A})$ satisfying* (IV.30). *Let the stationary dither d be generated by an exosystem of the form* (IV.11) *and let its entries be sufficiently rich of order $n+1$ and uncorrelated. Then, there exists $\gamma^{\star} > 0$ such that, for all $\gamma \in (0, \gamma^{\star}]$, there exists a compact set $\mathcal{A}_{\mathrm{s}}$ satisfying*

$$\mathcal{A}_{\mathrm{s}} \subset \{(w, x, z_{\mathrm{s}}) \in \mathcal{W} \times \mathbb{R}^{n} \times \mathcal{L}_{\mathrm{s}} : \hat{A} = A, \hat{K}_{\mathrm{a}} = 0, x = x_{\mathrm{m}}, \epsilon = 0\} \tag{IV.39}$$

that is uniformly globally asymptotically stable. ♕

The proof of Theorem IV.1 is provided in Appendix V.6.8.

Consider now the Model Reference Adaptive Reinforcement Learning (MR-ARL) algorithm with DRE implementation of $\hat{P}$ as in (IV.23) (Algorithm 5). Following the notation of Section IV.2.1, the controller obtained by combining the value function identifier (IV.21), (IV.22), (IV.23), reference model (IV.24), and the adaptive stabilizer (IV.25), (IV.26) is in the form (IV.9), with state

$$z := (z_s, \hat{P}) \in \mathcal{Z} := \mathcal{Z}_s \times S_0^n, \tag{IV.40}$$

where z_s and $\mathcal{Z}_s$ are given in (IV.36). The output policy then becomes

$$\pi(x, z, d) := (\hat{K}_a - BR^{-1}B^\top \hat{P})x + d, \tag{IV.41}$$

and the learning set $\mathcal{L}$ is given by

$$\mathcal{L} := \{z \in \mathcal{Z} : \hat{K}_a - BR^{-1}B^\top \hat{P} = -BR^{-1}B^\top P^\star\}. \tag{IV.42}$$

In this case, the learning set is reached via $\hat{A} \to A$ (hence $\hat{P} \to P^\star$ by asymptotic stability of the DRE (IV.8), see Section IV.2.1) and $\hat{K}_a \to 0$. The next result, which is the main result of this work, shows that with $\gamma > 0$ sufficiently small and $g > 0$ sufficiently large, the Model Reference Adaptive Reinforcement Learning in Algorithm 5 solves the robustly stable on-policy data-driven LQR problem.

Theorem IV.2. *Consider the closed-loop system given by the interconnection of plant* (IV.1) *and the controller of Algorithm 5. Pick d and γ as in Theorem IV.1. Then, the compact set*

$$\begin{aligned}\mathcal{A} := \mathcal{A}_s \times P^\star \subset \{&(w, x, z) \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{L} : \\ &\hat{A} = A, \hat{K}_a = 0, x = x_m, \epsilon = 0, P = P^\star\}\end{aligned} \tag{IV.43}$$

is semiglobally uniformly asymptotically stable in the tuning parameter $g > 0$, where $\mathcal{A}_s$ is given in (IV.39) *and g is the one in* (IV.23)*. Namely, for any compact set $\mathcal{K} \subset \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}$ of initial conditions for the closed-loop system, there exists $g > 0$ such that $\mathcal{A}$ is uniformly asymptotically stable with domain of attraction containing $\mathcal{K}$.* ♕

The proof of Theorem IV.2 is provided in Appendix V.6.9.
Notice that solving the robustly stable on-policy data-driven LQR problem as per Definition IV.1 allows to characterize the evolution of $x(t)$ as done in Remark IV.3, and thus to achieve arbitrarily small $|x(t)|$ by injecting a sufficiently small dither $|d(t)|$.

IV.3.2 Algorithm analysis

In the following, we will only study the properties of the reduced-order version of the algorithm, i.e., with $\hat{P}(t) = \mathcal{P}(A(t))$ for all t and $\mathcal{P}(\hat{A})$ satisfying (IV.30). The second result, i.e., the stability of Algorithm 5 (implementing the DRE), is obtained by invoking singular perturbations techniques. We now show boundedness and forward completeness of the solutions of the closed-loop system obtained from the interconnection of the identifier dynamics (IV.46), (IV.47), the reference model (IV.48), and the adaptive error system (IV.49), (IV.51). The overall analysis entails proving uniform bounds on the solutions of the main involved subsystems, then combining the results using arguments similar to [112, Thm. 6.3] (see the proof of Proposition IV.1). We begin by showing uniform boundedness of $\hat{A}$ and $\dot{\hat{A}}$.

Identifier Dynamics

Consider the error coordinate $\tilde{\epsilon}$ in (IV.28), which can be written as

$$\tilde{\epsilon} := A\xi - (x + \zeta). \tag{IV.44}$$

Then, from (IV.1), (IV.21), it holds that

$$\begin{aligned} \dot{\tilde{\epsilon}} &= A(-\lambda\xi + x) - (Ax + Bu - \lambda(x + \zeta) - Bu) \\ &= -\lambda(A\xi - (x + \zeta)) = -\lambda\tilde{\epsilon}, \end{aligned} \tag{IV.45}$$

which ensures that the prediction error $\epsilon := \hat{A}\xi - (x + \zeta) = (\hat{A} - A)\xi + \tilde{\epsilon}$ converges to $(\hat{A} - A)\xi$ exponentially. Define $\tilde{A} := \hat{A} - A$. Then, from (IV.28), (IV.45), we can rewrite the identifier dynamics (IV.21), (IV.22), (IV.27) in error coordinates as the following cascaded system

$$\begin{aligned} \dot{\tilde{\epsilon}} &= -\lambda\tilde{\epsilon} \\ \dot{\tilde{A}} &= \mathcal{P}_{\hat{A}\in C}\left\{-\gamma BB^{\dagger}\frac{\tilde{A}\xi\xi^{\top} + \tilde{\epsilon}\xi^{\top}}{1 + \nu|\xi||\epsilon|}\right\}, \end{aligned} \tag{IV.46}$$

driven by $\xi(t)$, solution of the filter

$$\dot{\xi} = -\lambda\xi + x. \tag{IV.47}$$

In the following lemma, we establish the boundedness properties of the identifier subsystem.

Lemma IV.1. *Let the maximal interval of solutions of* (IV.46), (IV.47), (IV.48), (IV.49), (IV.51) *be* $[0, t_f)$. *Then, it holds that*

i) $\tilde{\epsilon}(\cdot), \tilde{A}(\cdot)$ *are uniformly bounded in the interval* $[0, t_f)$

ii) $\hat{A}(t) \in \Theta$ *for all* $t \in [0, t_f)$

iii) $|\dot{\hat{A}}(t)| \leq \gamma$ *for all* $t \in [0, t_f)$.

Furthermore, if $t_f = \infty$, *the origin* $(\tilde{\epsilon}, \tilde{A}) = 0$ *of system* (IV.46), *driven by input* $\xi(t)$, *is uniformly globally stable (UGS).* ♙

The proof of Lemma IV.1 is provided in Appendix V.6.1.

The above results hold even if the input $\xi(t)$, obtained from (IV.21), of the identifier (IV.22) escapes to infinity as $t \to t_f$. We analyze only arbitrarily small intervals $[0, t_f)$ since, being the closed-loop system locally Lipschitz, all the solutions are guaranteed to exists in $[0, t_f)$ for some $t_f > 0$ [107, Thm. 3.1]. Although the overall boundedness analysis entails also the study of $\xi(t)$, system (IV.47) ISS with respect to input $x(t)$, thus its behavior will be analyzed directly in Proposition IV.1. To ensure $\hat{A}(t) \to A$, it is known from the adaptive control literature that vector $\xi(t)$ must be a persistently exciting (PE) signal [98]. However, notice that $\xi(t)$ is a filtered version of $x(t)$, which is generated in closed-loop by interconnecting the plant and the controller. For this reason, special care will be dedicated to its analysis.

Reference Model Dynamics

From (IV.30), when $\hat{P} = \mathcal{P}(\hat{A})$, system (IV.24) can be written highlighting the dependence on the estimate $\hat{A}$ of the identifier:

$$\dot{x}_{\mathrm{m}} = (\hat{A} - BR^{-1}B^{\top}\mathcal{P}(\hat{A}))x_{\mathrm{m}} + Bd, \tag{IV.48}$$

where from (IV.22), (IV.30), the pointwise-in-time value of $\mathcal{P}(\hat{A})$ is provided implicitly as the solution of a parameter-varying ARE. By [175, Thm. 4.1], $\mathcal{P}(\hat{A})$ is an analytic function of $\hat{A}$, being all matrices of ARE $\mathcal{R}(P, \hat{A}) = 0$ in (IV.29) analytic functions of $\hat{A} \in \Theta$. From this fact, matrix $\hat{A} - BR^{-1}B^{\top}\mathcal{P}(\hat{A})$ is an analytic function of $\hat{A}$ and it is Hurwitz at any time t. (see Remark IV.8).
We show now that the reference model (IV.48) is bounded as long as $|\dot{\hat{A}}(t)|$ is sufficiently small.

Lemma IV.2. *Let the maximal interval of solutions of* (IV.46), (IV.47), (IV.48), (IV.49), (IV.51) *be* $[0, t_f)$. *There exists* $\gamma_b^\star > 0$ *such that, if* $|\dot{\hat{A}}(t)| \leq \gamma_b^\star$ *for all* $t \in [0, t_f)$, *then* $x_m(\cdot)$ *is uniformly bounded over the interval* $[0, t_f)$. *Furthermore, if* $t_f = \infty$, *then the reference model* (IV.48) *with input* $d(t)$ *is input-to-state stable.* ♙

The proof of Lemma IV.2 is provided in Appendix V.6.2.

Adaptive Tracking Dynamics

We conclude this overview by studying the interconnection of the error dynamics (IV.34) and the adaptive controller (IV.25), (IV.26). We define $\tilde{K}_{\mathrm{a}} := \hat{K}_{\mathrm{a}} - K_{\mathrm{a}}(\hat{A})$. By choosing (IV.26) as input for (IV.31), we obtain:

$$\dot{e} = (\hat{A} - BR^{-1}B^{\top}\mathcal{P}(\hat{A}))e + B(\hat{K}_{\mathrm{a}}x - K(\hat{A})x) \tag{IV.49}$$

By choosing expression (IV.33) for $K_{\mathrm{a}}(\hat{A})$, we can explicitly calculate the variation in time of $K_{\mathrm{a}}(\hat{A})$ due to the movement of $\hat{A}$. This is out of the standard framework of model reference adaptive control, and thus particular attention is required. We can calculate the time derivative of $K_{\mathrm{a}}(\hat{A})$ by deriving (IV.33):

$$\dot{K}_{\mathrm{a}} = B^{\dagger}\dot{\hat{A}}. \tag{IV.50}$$

Since both B and $\dot{\hat{A}}$ are known, we can use their knowledge to implement adaptive law (IV.25), which takes into account this drift. Given equations (IV.25) and (IV.50), the induced dynamics for $\tilde{K}_\mathrm{a}$ is:

$$\begin{aligned}\dot{\tilde{K}}_\mathrm{a} &= \dot{\hat{K}}_\mathrm{a} - \dot{K}_\mathrm{a} \\ &= -\mu B^\top \mathcal{P}(\hat{A})(x - x_\mathrm{m})x^\top + B^\dagger \dot{\hat{A}} - B^\dagger \dot{\hat{A}} \\ &= -\mu B^\top \mathcal{P}(\hat{A}) e x^\top .\end{aligned} \tag{IV.51}$$

Next, we provide a statement for system (IV.49), (IV.51).

Lemma IV.3. *Let the maximal interval of solutions of* (IV.46), (IV.47), (IV.48), (IV.49), (IV.51) *be* $[0, t_f)$. *Pick* $\gamma_b^\star > 0$ *from Lemma IV.2 and let* $|\dot{\hat{A}}(t)| \le \gamma_b^\star$ *for all* $t \in [0, t_f)$. *Then, signals* $e(\cdot)$, $\tilde{K}_a(\cdot)$ *are uniformly bounded in the interval* $[0, t_f)$. *Furthermore, if* $t_f = \infty$, *the origin* $(e, \tilde{K}_a) = 0$ *of system* (IV.49), (IV.51), *with input* $\hat{A}(t)$, *is UGS.* ♙

The proof of Lemma IV.3 is provided in Appendix V.6.3.
Finally, we combine the previous results to obtain that solutions are globally bounded and forward complete.

Proposition IV.1. *Consider the closed-loop system obtained from the interconnection of the identifier dynamics* (IV.46), (IV.47), *the reference model* (IV.48), *and the adaptive error system* (IV.49), (IV.51). *Pick* $\gamma_b^\star$ *from Lemma IV.1. If* $\gamma \in (0, \gamma_b^\star]$, *then the closed-loop solutions are bounded and forward complete.* ♖

Proof. Suppose that the maximal interval of existence of the solution of (IV.46), (IV.47), (IV.48), (IV.49), and (IV.51) is $[0, t_f)$. Then, from Lemma IV.1, $\tilde{A}(\cdot)$ and $\tilde{\epsilon}(\cdot)$ are uniformly bounded. From Lemma IV.1, $|\dot{\hat{A}}(\cdot)|$ is uniformly bounded by γ. Consider any $\gamma \in (0, \gamma_b^\star]$, then Lemmas IV.2 and IV.3 ensure that $x_\mathrm{m}(\cdot)$, $e(\cdot)$, and $\tilde{K}_\mathrm{a}(\cdot)$ are uniformly bounded, thus also $\xi(\cdot)$ is uniformly bounded from (IV.47) and standard ISS results.
We have thus shown that all signals of the closed-loop system are bounded, with bounds that do not depend on t_f. By contradiction, we conclude that $t_f = \infty$, thus the solutions are forward complete. Namely, if t_f were finite, the solutions would leave any compact set as $t \to t_f$, contradicting the independence of the bounds on t_f [112, Thm. 6.3]. ※

Exponential Convergence to the Optimal Policy

We now focus on the uniform asymptotic stability properties of the closed-loop system (IV.46), (IV.47), (IV.48), (IV.49), (IV.51). First, we show that $x_\mathrm{m}(t)$ is persistently exciting as long as $|\dot{\hat{A}}|$ is sufficiently small.

Lemma IV.4. *Let* $d(t) \in C_\mathrm{SR}(\mathbb{L}^\mathrm{C}_{\mathrm{x},>1})$ *as per* (II.55). *There exists* $\gamma_{PE}^\star \in (0, \gamma_b^\star]$, *with* $\gamma_b^\star$ *from Proposition IV.1, such that, for all* $\gamma \in (0, \gamma_{PE}^\star]$, *the solutions* $x_\mathrm{m}(t)$ *of the reference model* (IV.48) *are persistently exciting (PE).* ♙

The proof of Lemma IV.4 is provided in Appendix V.6.4.
Next, we provide a direct consequence of Lemma IV.4 for the adaptive error dynamics (IV.49), (IV.51).

Lemma IV.5. *Let the hypotheses of Lemma IV.4 hold and let* $\gamma \in (0, \gamma^\star_{PE}]$, *where* $\gamma^\star_{PE}$ *is given in Lemma IV.4. Then, the origin* $(e, \tilde{K}_a) = 0$ *of system* (IV.49), (IV.51) *is uniformly globally asymptotically stable (UGAS) and uniformly locally exponentially stable (ULES).*

The proof of Lemma IV.5 is provided in Appendix V.6.5.
Now that we have established that every solution $e(t)$ converges exponentially to zero, uniformly from compact sets of initial conditions, we can conclude the convergence analysis by studying the identifier dynamics (IV.46).

Lemma IV.6. *Let the hypotheses of Lemma IV.4 hold and let* $\gamma \in (0, \gamma^\star_{PE}]$, *where* $\gamma^\star_{PE}$ *is given in Lemma IV.4. Then, the origin* $(\tilde{\epsilon}, \tilde{A}) = 0$ *of system* (IV.46), *with input* $\xi(t)$, *is uniformly globally exponentially stable (UGES).*

The proof of Lemma IV.6 is provided in Appendix V.6.6.

IV.3.3 Numerical simulations

In this section, we propose two numerical examples to show the effectiveness of Model Reference Adaptive Reinforcement Learning. In the first example, we consider the model of a doubly fed induction motor (DFIM) at constant speed with unknown rotor and stator resistances. In the second example, we test the robustness of the proposed algorithm by considering a DFIM with slowly time-varying unknown resistances, due to the motor heating up, and rotor acceleration. In order to highlight the claimed stability properties of the algorithm, we provide examples where we show the complete transient from the initial condition to the steady state. Notice, however, that MR-ARL is not meant to achieve a fast identification, but to follow (in an online fashion) slowly time-varying changes in the dynamics while preserving stability of the closed loop.

Example 1: Constant Parameters

A DFIM at constant speed can be modeled [117] with a linear system in the form of (IV.1) with state

$$x = (i_{1u}, i_{1v}, i_{2u}, i_{2v}) \in \mathbb{R}^4, \tag{IV.52}$$

where i_{1u}, i_{1v} are the stator currents and i_{2u}, i_{2v} are the rotor currents. The input is

$$u = (u_{1u}, u_{1v}, u_{2u}, u_{2v}) \in \mathbb{R}^4, \tag{IV.53}$$

where u_{1u}, u_{1v} are the stator voltages and u_{2u}, u_{2v}, the rotor voltages. System matrices are defined as

$$\begin{aligned} A &= \frac{1}{\bar{L}} \begin{bmatrix} -L_2R_1 & -\alpha+\beta & L_mR_2 & \beta_2 \\ \alpha-\beta & -L_2R_1 & -\beta_2 & -L_mR_2 \\ L_mR_1 & -\beta_1 & -L_1R_2 & -\alpha-\beta_{12} \\ \beta_1 & L_mR_1 & \alpha+\beta_{12} & -L_1R_2 \end{bmatrix} \\ B &= \frac{1}{\bar{L}} \begin{bmatrix} L_2 & 0 & -L_m & 0 \\ 0 & L_2 & 0 & -L_m \\ -L_m & 0 & L_1 & 0 \\ 0 & -L_m & 0 & L_1 \end{bmatrix}, \end{aligned} \tag{IV.54}$$

where

$$\begin{aligned} \bar{L} &:= L_1L_2 - L_m^2 \\ \alpha &:= \bar{L}\omega_0, \quad \beta := L_m^2\omega_r \\ \beta_{12} &:= L_1L_2\omega_r, \quad \beta_1 := L_1L_m\omega_r, \quad \beta_2 := L_2L_m\omega_r. \end{aligned} \tag{IV.55}$$

Parameters R_1, R_2 are the stator and rotor resistances, while L_1, L_2, L_m are the stator and rotor auto-inductances and the mutual inductance, respectively. Finally, ω_r and ω_0 are the electrical angular speeds of the rotor and the rotating reference frame, which we suppose constant.

Remark IV.10. We suppose to have uncertainties on the parameters R_1 and R_2. This makes the matrix A uncertain in half of its entries. In this example B is such that $\mathrm{Im}(\mathcal{B}) = \mathbb{R}^{n\times n}$, so Assumption IV.2 is fulfilled for any $A_0 \in \mathbb{R}^{n\times n}$. ◊

Denote the true resistances as R_1, R_2. We model our uncertainties specifying nominal values $\bar{R}_1, \bar{R}_2$ and radiuses $r_1, r_2 > 0$ such that

$$\begin{aligned} R_1 &\in [\bar{R}_1 - r_1, \bar{R}_1 + r_1] \\ R_2 &\in [\bar{R}_2 - r_2, \bar{R}_2 + r_2]. \end{aligned} \tag{IV.56}$$

Next, we define C as a ball about the nominal $\bar{A}$ (i.e., having the structure (IV.54) with resistances $\bar{R}_1$ and $\bar{R}_2$) containing all possible parameter variation, i.e.,

$$C := \{\hat{A} \in \mathbb{R}^{n\times n} : |\hat{A} - \bar{A}|_F \le \rho\} \tag{IV.57}$$

with $\rho > 0$ big enough. We report in Table V.1 the physical parameters of the motor. In Table IV.3, we specify the values used for the uncertainties and the desired performances.
The dither $d(t)$ is designed, on each entry $d_i(t)$, according to

$$d_i(t) = 10 \sum_{j=1}^{4} \mathrm{sawtooth}(2\omega_s ijt), \quad i \in \{1, 2, 3, 4\} \tag{IV.58}$$

Table IV.2: Physical parameters of the motor.

Parameter	Value	Parameter	Value
L_1 [H]	0.02645	R_1 [Ω]	0.036
L_2 [H]	0.0264	R_2 [Ω]	0.038
L_m [H]	0.0257	ω_0 [rad/s]	$2\pi 70.8$
p	3	ω_r [rad/s]	$2\pi 62$

Table IV.3: Uncertainty parameters for example 1.

Parameter	Value	Parameter	Value
$\bar{R}_1$ [Ω]	0.03	r_1 [Ω]	0.01
$\bar{R}_2$ [Ω]	0.03	r_2 [Ω]	0.01
ρ	20		

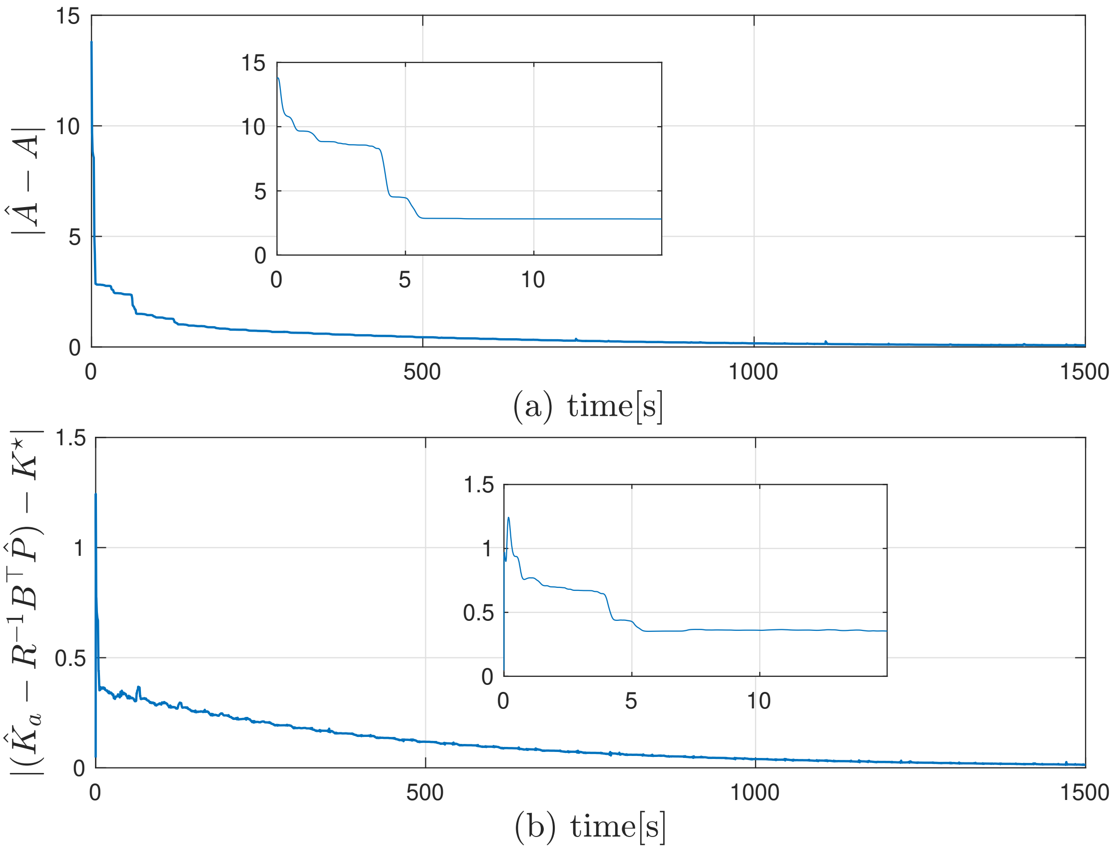

Figure IV.2: Convergence to true A and to optimal gain $K^\star$.

where sawtooth$(\cdot)$ is a triangular wave of unitary amplitude and $\omega_s = 0.2$ rad/s.
The excitation levels in the closed-loop system strongly depend on the excitation levels of the dither signals. Thus, by injecting stronger dither signals, it is possible to achieve faster parameter convergence [7].
In Fig. IV.2-(a), we show the difference between the estimate $\hat{A}(t)$ and the true matrix A. Next, in Fig. IV.2-(b), we show how the error between the optimal feedback gain $K^\star$ and the overall applied feedback

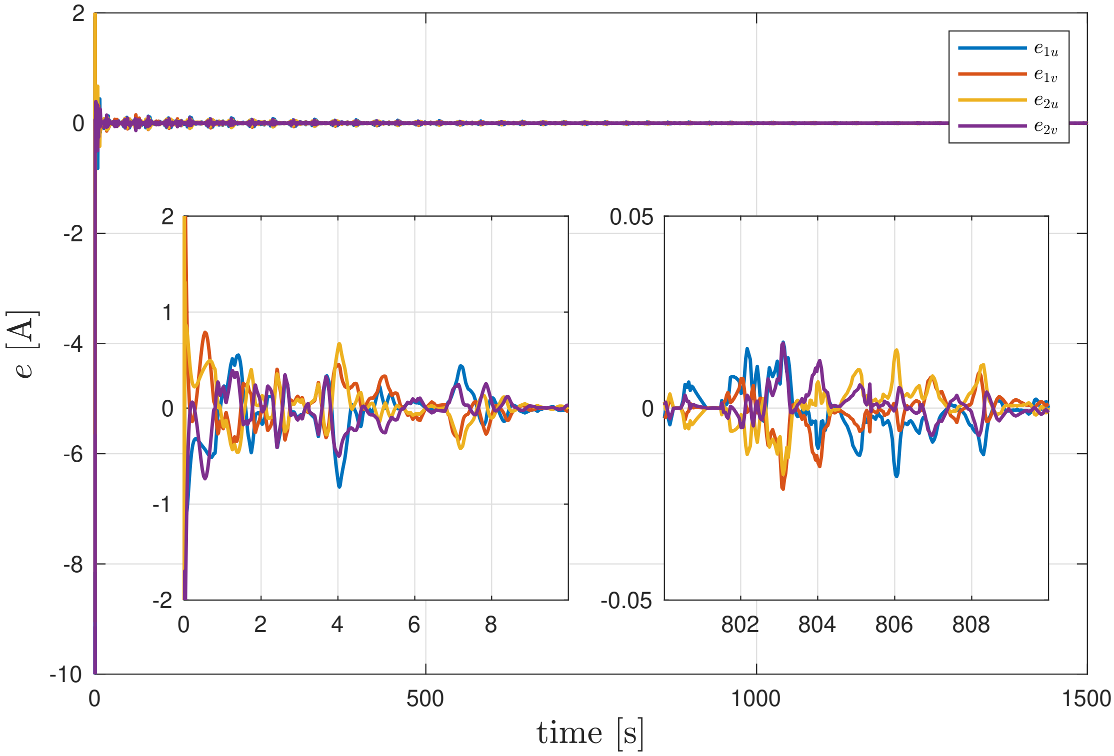

Figure IV.3: Tracking error between plant and reference model.

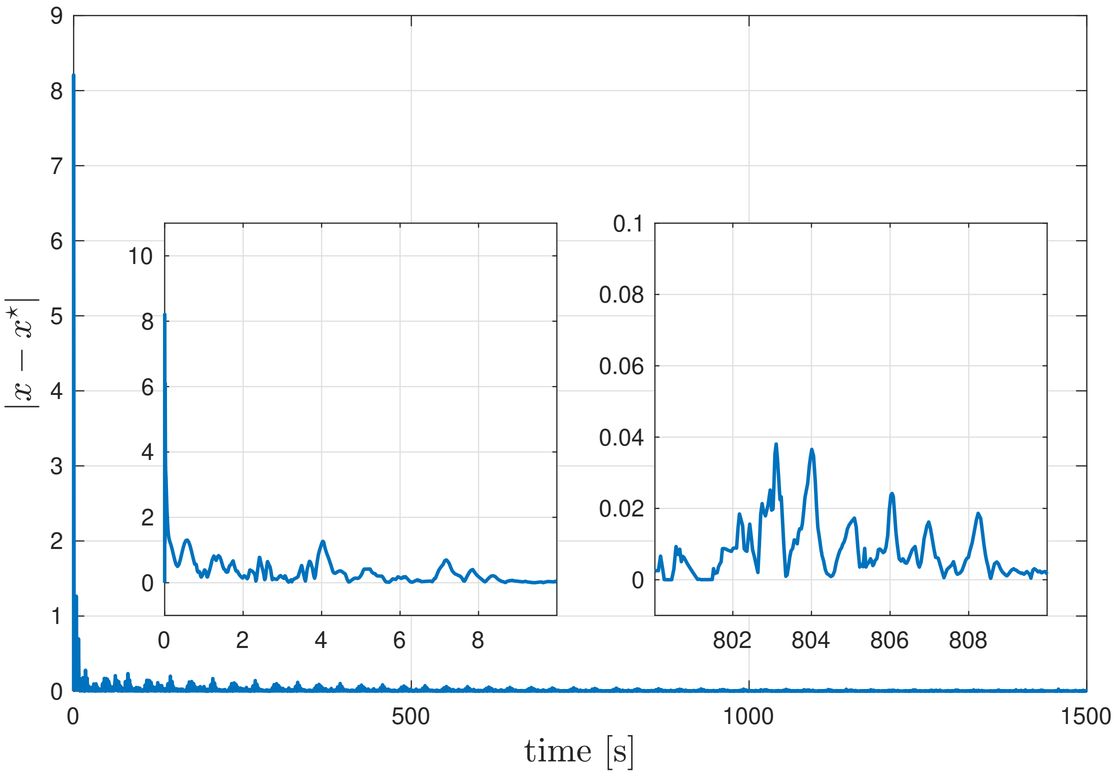

Figure IV.4: Difference between the optimal trajectory and the trajectory generated by MR-ARL.

Table IV.4: Uncertainty parameters for example 2.

Parameter	Value	Parameter	Value
$\bar{R}_1$ [Ω]	0.2	r_1 [Ω]	0.18
$\bar{R}_2$ [Ω]	0.2	r_2 [Ω]	0.18
$\bar{\omega}_r$ [rad/s]	$2\pi 70$	r_ω [rad/s]	$2\pi 15$
ρ	4830		

gain $-R^{-1}B^\top \hat{P} + \hat{K}_a$ approaches zero, thus controlling in an optimal way the system. In Fig. IV.3, we show for completeness the error between the reference model and the real system, which reaches, despite a slower parameter convergence, a small amplitude in a few seconds.
Finally, in Fig. IV.4 we show the difference between the trajectory $x(t)$ generated by MR-ARL and the trajectory obtained by choosing $u^\star(t) = K^\star x^\star(t) + d(t)$, namely, the optimal policy plus the dither signal.

Example 2: Drifting Parameters and Variable Speed

In this example, we apply perturbations to the DFIM with model given in (IV.54) to test the robustness of MR-ARL. We consider two perturbations to the nominal model occurring together: the first one is a time-varying resistance due to motor heating up, while the second one is a time-varying rotor speed due to load changes. We model both disturbances with sigmoid functions and we report them in the plots. The temperature disturbance lasts for about 600 s and brings the temperature from 20 °C to 100 °C, i.e., $\Delta T = 80$ °C. The speed disturbance is a total increase of speed of $2\pi 20$ rad/s occurring in about 60 s. We model the dependence of resistances on temperature with

$$R_i(\Delta T) = R_i + \alpha \Delta T, \qquad i \in \{1, 2\}, \tag{IV.59}$$

where $\alpha_{\mathrm{CU}} = 4.041 \times 10^{-3}$ Ω/°C is the temperature coefficient of resistance of the copper. We set new nominal $\bar{R}_1, \bar{R}_2, \bar{\omega}_r$ with associated range r_1, r_2, r_ω (reported in Table IV.4) to consider these uncertainties. We recalculate C as in the previous example. Besides these time-varying perturbations in the matrix A, we introduce also noise in the measurements of the currents $x(t)$. The measured $\bar{x}(t)$ is given by

$$\bar{x}(t) = (1 + \tilde{I}(t) + \bar{I})x(t) \tag{IV.60}$$

where $\tilde{I}(t)$ is extracted at each t from a uniform distribution in the interval $[-0.5, 0.5]\%$, and $\bar{I} = 1\%$. Finally, we leave the dither as in (IV.58).

Remark IV.11. Due to the parameter variations, the plant becomes a slowly time-varying system. Consistently with the theoretical result, due to the "small" variations, the stability properties of Theorem IV.2 are practically preserved and recovered when the variations vanish. ◊

In Fig. IV.5-(a), we show the difference between the estimate $\hat{A}(t)$ and the true time-varying matrix

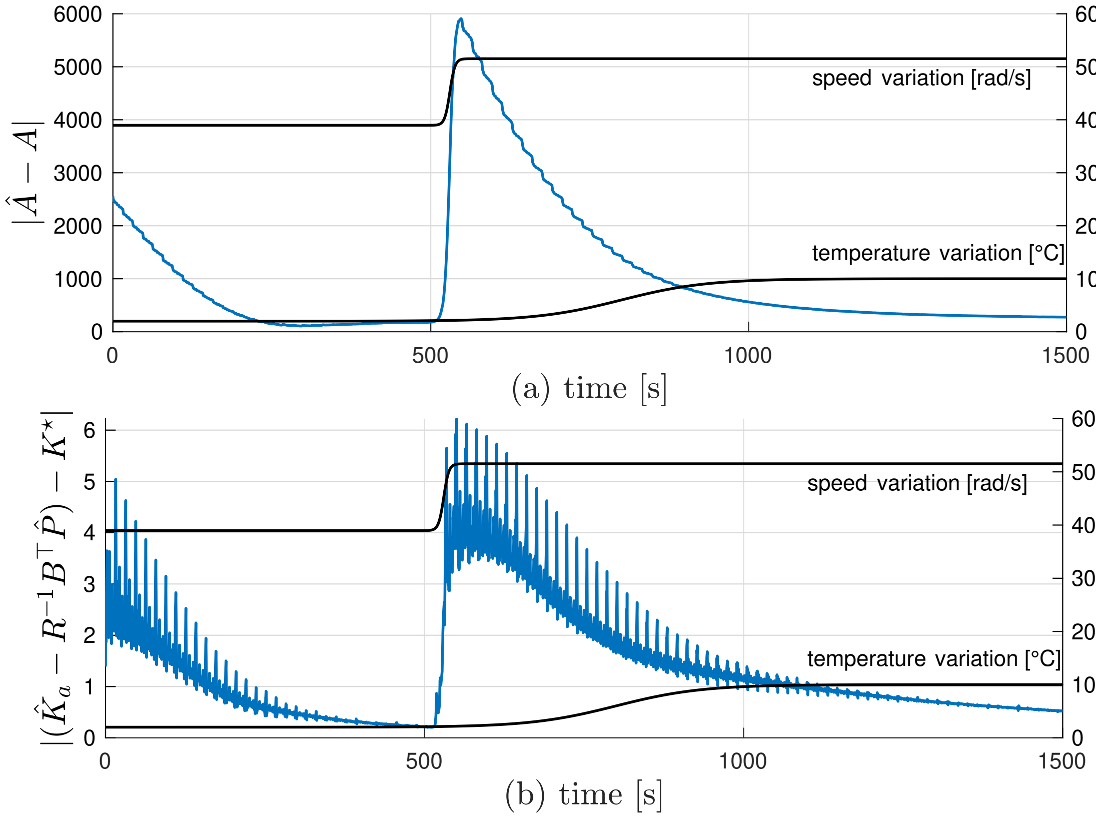

Figure IV.5: Convergence to true $A(t)$ and to optimal gain $K^{\star}(t)$.

$A(t)$. Notice that as soon as the speed disturbance ends, the gradient estimator is able to adapt and recover convergence of the estimation to a small ball about the true parameters. Next, in Fig. IV.5-(b), we show how the data-driven feedback gain approaches the optimal one. Since in this simulation we have a LTV plant, we calculate at each time instant the optimal gain $K^{\star}(t)$ by solving an LQR problem with constant $A(t)$. The importance of the adaptive controller action is particularly clear in presence of the speed disturbance, where the estimated matrix is far from the true one and thus the optimal action is likely to be destabilizing. Finally, we show in Fig. IV.6 how the error between the reference model and the real plant is kept bounded also in the presence of these disturbances.

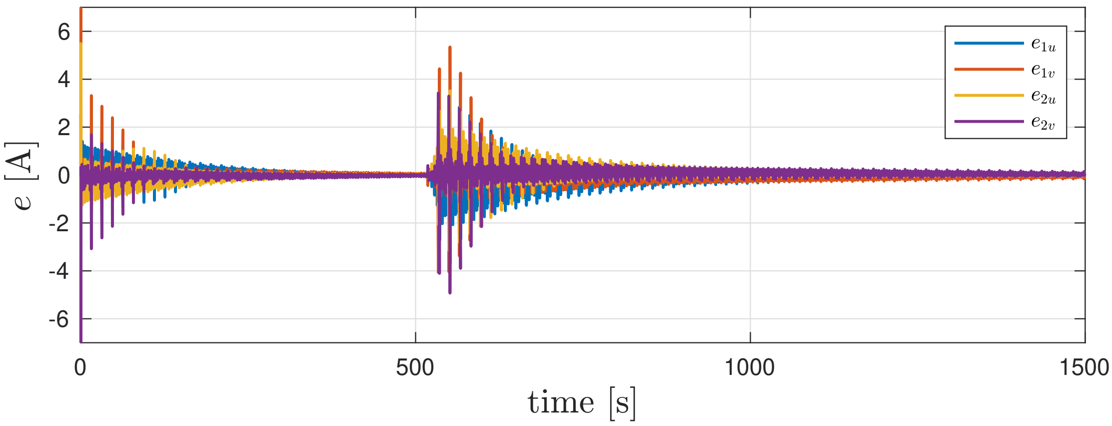

Figure IV.6: Tracking error between plant and reference model. Different colors stand for different components of e.

IV.3.4 Future work

In this chapter, we have addressed the problem of data-driven optimal control of partially unknown linear systems. First, we have proposed a framework that formalizes a robustly stable on-policy data-driven LQR problem in which optimality of the learned strategy is obtained while guaranteeing robust stability of the whole learning and control closed-loop system. Next, we have proposed a new solution to this problem consisting in the combination of model reference adaptive control and reinforcement learning. As main result, we showed that our design has a semiglobally uniformly asymptotically stable attractor where the plant follows the optimal reference model. To demonstrate the effectiveness of the solution, we tested it in the control of a doubly fed induction motor. The results show that our solution is also able to manage non-vanishing perturbations typical of real-world applications. Future work will be dedicated to consider the output-feedback framework and to relax the assumption of the knowledge of the input matrix B, as well as to extending the framework to the output-feedback scenario.

Chapter V

DATA-DRIVEN PRONTO: a model-free algorithm for numerical optimal control

I conclude this thesis with my first and one of the last works. The idea behind DATA-DRIVEN PRONTO is to design a data-driven solver for nonlinear optimal control building upon PRONTO [90], an iterative solver. Although the algorithm will be explained more in details in the next sections, we start by providing here some key concepts which are fundamental in understanding both PRONTO and DATA-DRIVEN PRONTO.

i) *PRojection Operator-based*: whilst "pure" SQP methods [66] satisfy the dynamic constraint only asymptotically, several algorithms [69, 118, 144] developed specifically for optimal control perform an additional "forward sweep" (on the whole time horizon) of the dynamics in order to satisfy at each algorithm iteration the dynamic constraints. Both PRONTO and DATA-DRIVEN PRONTO improve the numerical stability properties of this forward sweep by leveraging on an additional control law to be designed separately. This step is called "Projection" step, since it projects an infeasible curve to a trajectory of the system.

ii) *Newton method for Trajectory Optimization*: by including the projection operator in the cost function, PRONTO builds an unconstrained problem which can be solved through standard Newton method.

Whilst all the mentioned algorithms exploit nicely the structure of the optimal control problem, a major drawback is that they all require the knowledge of the dynamics and its derivatives. The main idea behind DATA-DRIVEN PRONTO is to leverage on the control law introduced by PRONTO (which "robustifies" the projection operator) to collect the data necessary to the estimation of derivatives.

This Chapter is organized as follows. In Section V.1, we give an overview of the scientific literature pertaining data-driven optimization and optimal control. Next, in Section V.2, we introduce the considered nonlinear optimal control problem with the necessary preliminaries on PRONTO, and we state the main contribution to the scientific community. Finally, in Section V.3 we present DATA-DRIVEN PRONTO. Theoretical guarantees on its convergence are given, and its hyper-parameters are discussed. The algorithm is analyzed via intermediate results, and a numerical example is given to show the effectiveness of the proposed algorithm. All the proofs of the intermediate results can be found in Appendix V.7. An article containing the results of this chapter is currently under preparation (to be submitted soon).

V.1 Literature review

Nonlinear optimal control problems are prevalent in various applications within the fields of Automation and Robotics, where the goal is to develop a control strategy that, when applied to a dynamical system, minimizes a specified performance index [28, 46, 66, 164, 189]. In general, the solution of Non-Linear Optimal Control Problems (NL-OCPs) strongly relies on a valid model of the system under control which, if inaccurate, can lead to the design of suboptimal trajectories for the true system. In this chapter, we propose a combination of identification techniques together with data-driven control which can solve NL-OCP in a model-free setting.
In particular, we focus on solving a finite-horizon NL-OCP where the dynamics is unknown but there is the possibility of performing multiple deterministic experiments (or simulations) from the same initial condition. Retrieving an optimal trajectory for this kind of setups is important, for example, in many industrial systems, where optimized execution of a repeated task results in significant savings. Pioneers in this field were the so called *Repetitive Control* (RC) and *Iterative Learning Control* (ILC) [2, 44, 94, 126, 153], which successfully approached the generic problem of tracking a given reference exploiting the possibility of learning a specific task by means of multiple repetition of that specific task. Besides RC and ILC, and taking more explicitly into account the problem of optimality, we provide an overview of two fields that deal with the problem of data-driven optimal control, distinguishing between the "Reinforcement Learning" (RL) and "Data-driven optimization" approaches (where, since the two fields have lots in common, we stress the fact that this distinction is done somehow arbitrarily, and several works may be placed in both categories).

Reinforcement Learning

Having its foundations in the idea of Dynamic Programming [20, 21], RL field considers the more generic problem of learning a control policy which minimizes (maximizes) a received reward through interaction with the environment [177, 202]. The learning happens during a so-called training phase, in which - while the agent explores the environment - *good* actions are rewarded and *bad* actions are penalized. At the end of the training, the result is in general a policy and not a sequence of open-loop inputs. In some cases, this feature makes RL more robust than optimal control with respect to unmodeled dynamics [199]. However, two problems arise in general with this approach: i) the number of required episodes is usually so huge that this learning is achievable only via simulations (and not in real setups); ii) (for continuous or highly-dimensional problems) a parametrization of the policy or the value function, e.g., a Neural Network (NN), has to be introduced, constraining the resulting solution to a particular form and introducing other complexity. In this field, [58, 59] propose to solve the optimal control problem by imposing a linear parametrization of the value function, and finding its parameters via a linear program constrained by collected data samples. In [1], it is proposed to use an approximate model to obtain local improvements in the parametrized policy, which is evaluated via real experiments. In [151], a NN is used to parametrize both policy and value function, and policy iteration is applied to solve the infinite-horizon NL-OCP. In [93], the authors consider a finite horizon NL-OCP, and they use a NN to approximate the costate and finding the optimal input. In [39], the authors consider a continuous-time NL-OCP and train a NN that parametrizes the input with gradient descent. In [141], the authors show by means of a practical application how the combination of a classic technique such as ILC with RL can reduce significantly the amount of required data for the training. Finally, in [70], the authors explore the idea of using RL to find open-loop optimal input instead of optimal policy, thus overcoming the parametrization problem.

Data-driven optimization

Parallel to the RL paradigm, optimization techniques capable of handling problems with parametric uncertainties where developed throught the years. Black-box optimization, derivative-free optimization and simulation optimization [3, 4, 53, 178] enclose all the techniques which are developed when the explicit cost function, its derivatives, or the constraints are not available to the optimization process (or they are too complicated). In general, in these cases the idea is to substitute the missing knowledge by *cleverly probing* the cost function, exploiting the knowledge of some known property (for example, in [75, 183, 184] the authors leverage on Lipschitz continuity). Given the particular structure of NL-OCPs, several iterative and efficient ways to solve them [60, 69, 90, 100, 118, 119, 158] have been proposed in the years. The structure of the NL-OCP for a data-driven resolution is exploited in [52], where the authors propose an algorithm to iteratively solve a NL-OCP with partial knowledge of the dynamics. We refer the reader to [64, 174, 177] for other ways of solving in a data-driven way NL-OCPs.

V.2 Problem setup

V.2.1 Problem statement: data-driven nonlinear optimal control

In this first section, we introduce the problem setup together with some preliminaries on model-based optimal control of nonlinear systems. In this chapter, we consider nonlinear systems described by the discrete-time dynamics

$$x_{t+1} = f(x_t, u_t), \qquad x_0 = x_{\text{init}} \tag{V.1}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is the dynamics and $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ are, respectively, the state of the system and the control input at time $t \in \mathbb{N}$. The initial condition is fixed to be $x_{\text{init}} \in \mathbb{R}^n$. Importantly, we assume dynamics (V.1) to be unknown, i.e., we do not have access to an explicit form of f. Nevertheless, we assume to be able to actuate the system with given input sequence $u_0, \ldots, u_{T-1}$ and to measure the noiseless states $x_0, \ldots, x_T$: consider, e.g., the possibility to retrieve these data from a realistic simulator, or from an experiment. To simplify the notation, we denote finite-dimensional stacks of vectors with bold letters, namely, we can denote the stacks of state and input sequence as

$$\begin{aligned} \boldsymbol{u} &:= (u_0, \ldots, u_{T-1}) \in \mathbb{R}^{nT} \\ \boldsymbol{x} &:= (x_0, x_1, \ldots, x_T) \in \mathbb{R}^{n(T+1)}. \end{aligned} \tag{V.2}$$

More formally, we assume we can measure trajectories of system (V.1), i.e., state-input sequences satisfying the following definition.

Definition V.1. *[System trajectory] The pair* $(\boldsymbol{x}, \boldsymbol{u}) \in R^s$ *is a trajectory of system* (V.1) *if it satisfies*

$$x_{t+1} = f(x_t, u_t) \tag{V.3}$$

for all $t = 0, \ldots, T-1$ *with* $x_0 = x_{init}$. ♔

Compactly, we denote a trajectory as $\eta := (\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^s$, where $s = s_x + s_u$ and $s_x := n(T+1)$, $s_u := mT$.

Definition V.2. *[Trajectory manifold] We denote as* $\mathcal{T} \subset \mathbb{R}^s$ *the manifold of all the trajectories of* (V.1) *as per Def. V.1 of fixed initial condition* x_{init}. *Notice that, by defining*

$$h(\boldsymbol{x}, \boldsymbol{u}) = \begin{bmatrix} x_0 - x_{init} \\ \ldots \\ x_T - f(x_{T-1}, u_{T-1}) \end{bmatrix}, \tag{V.4}$$

we may characterize systems trajectories as

$$(\boldsymbol{x}, \boldsymbol{u}) \in \mathcal{T} \iff h(\boldsymbol{x}, \boldsymbol{u}) = 0. \tag{V.5}$$

♔

Notice $\mathcal{T}$ is a manifold since the Jacobian of $h(\boldsymbol{x}, \boldsymbol{u})$ has constant rank independently on f (provided that f is differentiable). Conversely, a generic element of $\mathbb{R}^s$ not necessarily satisfying Definition V.1 is said to be a curve and we denoted it as $\xi := (\boldsymbol{\alpha}, \boldsymbol{\mu}) \in \mathbb{R}^s$, with

$$\begin{aligned} \boldsymbol{\alpha} &:= (\mu_0, \ldots, \mu_{T-1}) \in \mathbb{R}^{s_x}, \\ \boldsymbol{\mu} &:= (\alpha_0, \alpha_1, \ldots, \alpha_T) \in \mathbb{R}^{s_u}. \end{aligned} \tag{V.6}$$

Our objective is to design input sequences $\boldsymbol{u}$ for the unknown system (V.1) such that the resulting trajectory $(\boldsymbol{x}, \boldsymbol{u})$ minimizes a nonlinear performance index

$$\ell(\boldsymbol{x}, \boldsymbol{u}) := \sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T), \tag{V.7}$$

where $\ell_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ is the so-called stage cost and $\ell_T : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is the terminal cost. Compactly, we aim at solving the following optimal control problem

$$\begin{aligned} \min_{\boldsymbol{x}, \boldsymbol{u}} \quad & \sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T) \\ \text{subj. to} \quad & x_{t+1} = f(x_t, u_t), \quad t = 0, \ldots, T-1, \\ & x_0 = x_{\text{init}}. \end{aligned} \tag{V.8}$$

Remark V.1. The finite-time nature of the considered problem make its solution interesting especially for the generation of an optimal reference to be followed in repeated tasks. ◊

We introduce the following hypothesis on the smoothness of the dynamics and the cost function.

Assumption V.1 (Regularity). *The dynamics f in* (V.1) *and the cost function ℓ in* (V.7) *are twice continuously differentiable in their arguments, i.e., f, ℓ are C^2.*

As it will be clearer further in this chapter, we also introduce the following assumptions regarding the second-order derivatives of the cost function (V.7).

Assumption V.2. *For all state and input sequences $(\boldsymbol{x}, \boldsymbol{u}) \in \mathbb{R}^s$, it holds $\nabla^2 \ell(\boldsymbol{x}, \boldsymbol{u}) > 0$.*

Assumption V.2 guarantees the possibility of finding a valid descent direction; however, it can be relaxed via regularizations (see Remark V.2). Furthermore, if the cost ℓ is user-defined, it is possible to readily satisfy Assumption V.2. We highlight that the key challenge of solving problem (V.8) is that an expression for the dynamics (V.1) or its derivatives is not accessible in explicit form, i.e., it is not available to the solver.

V.2.2 Preliminaries: an introduction to PRONTO

In this preliminary section we present a discrete-time version of the optimal control algorithm PRONTO (on which we build DATA-DRIVEN PRONTO), which has been proposed in [90] for the continuous-time framework. The underlying idea is to leverage on a feedback policy to map generic elements $(\boldsymbol{\alpha}, \boldsymbol{\mu})$ of the space $\mathbb{R}^s$, the so-called curves, into the set of trajectories feasible for the dynamics (V.1). This projection is assumed to be implemented via the nonlinear tracking system

$$\begin{aligned} u_t &= \pi(\alpha_t, \mu_t, x_t, t) \\ x_{t+1} &= f(x_t, u_t), \qquad x_0 = x_{\text{init}}, \end{aligned} \tag{V.9}$$

where $\pi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$ is a generic tracking controller whose properties will be later clarified. The interconnection (V.9) between the control law π and the dynamics f can be seen as a *projection operator* which projects generic curves $(\boldsymbol{\alpha}, \boldsymbol{\mu})$ onto the trajectory manifold of system (V.1), i.e., it implements a map $(\boldsymbol{\alpha}, \boldsymbol{\mu}) \mapsto (\boldsymbol{x}, \boldsymbol{u})$ (we refer to [90] for a detailed discussion). More in detail, for all iterations indexed by k, the methodology proposed in [90] seeks for an update direction $(\boldsymbol{\Delta x}^k, \boldsymbol{\Delta u}^k)$ onto the tangent space of the trajectory manifold $\mathcal{T}$ at the current solution trajectory $(\boldsymbol{x}^k, \boldsymbol{u}^k)$ (step 1 of Algorithm 6). The descent direction is obtained by solving LQR problem (V.13), where A_t^k, B_t^k, q_t^k, r_t^k are defined as

$$\begin{aligned} A_t^k &:= \nabla_1 f(x_t^k, u_t^k)^\top, & B_t^k &:= \nabla_2 f(x_t^k, u_t^k)^\top, \\ q_t^k &:= \nabla_1 \ell_t(x_t^k, u_t^k)^\top, & r_t^k &:= \nabla_2 \ell_t(x_t^k, u_t^k)^\top, \end{aligned} \tag{V.10}$$

and $Q_t^k, Q_T^k \in \mathbb{R}^{n\times n}$, $S_t^k \in \mathbb{R}^{n\times m}$ and $R_t^k \in \mathbb{R}^{m\times m}$ are defined as

$$\begin{aligned} Q_t^k &:= \nabla_{11}^2 \ell_t(x_t^k, u_t^k), & S_t^k &:= \nabla_{12}^2 \ell_t(x_t^k, u_t^k), \\ R_t^k &:= \nabla_{22}^2 \ell_t(x_t^k, u_t^k), & Q_T^k &:= \nabla_{11}^2 \ell_t(x_T^k). \end{aligned} \tag{V.11}$$

After computing the descent direction, PRONTO updates the estimate of the solution $(\boldsymbol{\alpha}^k, \boldsymbol{\mu}^k)$ according to

$$\begin{bmatrix} \boldsymbol{\alpha}^k \\ \boldsymbol{\mu}^k \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}^k \\ \boldsymbol{\mu}^k \end{bmatrix} + \gamma^k \begin{bmatrix} \boldsymbol{\Delta x}^k \\ \boldsymbol{\Delta u}^k \end{bmatrix}, \tag{V.12}$$

where $\gamma^k \in (0, 1]$ is an appropriate stepsize (step 2 of Algorithm 6). Since the updated solution $(\boldsymbol{\alpha}^k, \boldsymbol{\mu}^k)$ does not satisfy, in general, the dynamics constraint, the updated trajectory $(\boldsymbol{x}^{k+1}, \boldsymbol{u}^{k+1})$ is then obtained from $(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\mu}^{k+1})$ via (V.9) (step 3 of Algorithm 6).

Remark V.2. [Approximations of descent step] In (V.11), we simplified the cost matrices by neglecting a term involving the Hessian of the dynamics. In order not to lose the convergence properties of the algorithm, it is sufficient to guarantee that the matrix in the quadratic part of the cost in (V.13) is positive definite for all t (e.g., using the identity matrix, regularizing the cost or using Assumption V.2) [66, Cor. 4.3]. ◇

Algorithm 6 recaps the procedure described so far.

Algorithm 6 PRONTO

Require: Initial trajectory $(\boldsymbol{x}^0, \boldsymbol{u}^0)$, controller π, dynamics f.

for $k = 0, 1, 2 \dots$ **do**

Step 1: find descent direction $(\boldsymbol{\Delta x}^k, \boldsymbol{\Delta u}^k)$ by solving

$$\min_{\boldsymbol{\Delta x}, \boldsymbol{\Delta u}} \sum_{t=0}^{T-1} \frac{1}{2} \begin{bmatrix} \Delta x_t \\ \Delta u_t \end{bmatrix}^\top \begin{bmatrix} Q_t^k & S_t^k \\ S_t^{\top,k} & R_t^k \end{bmatrix} \begin{bmatrix} \Delta x_t \\ \Delta u_t \end{bmatrix} + \begin{bmatrix} q_t^k \\ r_t^k \end{bmatrix}^\top \begin{bmatrix} \Delta x_t \\ \Delta u_t \end{bmatrix} + \Delta x_T^\top Q_T^k \Delta x_T + q_T^{k\top} \Delta x_T$$

$$\text{subj. to } \Delta x_{t+1} = A_t^k \Delta x_t + B_t^k \Delta u_t,$$

$$\Delta x_0 = 0, \quad t = 0, \dots, T-1.$$

Step 2: update curve $(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\mu}^{k+1})$:

$$\begin{aligned} \alpha_t^{k+1} &= x_t^k + \gamma^k \Delta x_t^k \\ \mu_t^{k+1} &= u_t^k + \gamma^k \Delta u_t^k, \end{aligned} \tag{V.13}$$

with $t = 0, \dots, T-1$.

Step 3: find new trajectory $(\boldsymbol{x}^{k+1}, \boldsymbol{u}^{k+1})$:

$$\begin{aligned} u_t^{k+1} &= \pi(\alpha_t^{k+1}, \mu_t^{k+1}, x_t^{k+1}, t) \\ x_{t+1}^{k+1} &= f(x_t^{k+1}, u_t^{k+1}), \quad x_0^{k+1} = x_{\text{init}}, \end{aligned} \tag{V.14}$$

with $t = 0, \dots, T-1$.

end for

To summarize, the steps $1-2$ of Algorithm 6 are the same of an SQP method. The main difference between PRONTO and other algorithms stands in the projection step (V.9), which is intended to numerically robustify its performance with repect to both SQP (where $(\boldsymbol{x}^{k+1}, \boldsymbol{u}^{k+1})$ do not in general satisfy the dynamic constraint, and this constraint violation needs to be taken into account) and shooting approaches, which are in general less stable (since only $\boldsymbol{u}^{k+1}$ is updated and $\boldsymbol{x}^{k+1}$ is found in open loop given dynamics (V.1)).

V.3 DATA-DRIVEN PRONTO

V.3.1 The algorithm

We are now ready to present DATA-DRIVEN PRONTO, our data-driven algorithm for optimal control. Moving from the foundings ideas of PRONTO, our approach extends the algorithm to the model-free framework. Indeed, notice that in order to implement Step 1 of Algorithm 6, perfect knowledge of the dynamics (V.1) is required. In the following, we introduce how DATA-DRIVEN PRONTO is capable of

Algorithm 7 DATA-DRIVEN PRONTO

Require: Initial trajectory $(\boldsymbol{x}^0, \boldsymbol{u}^0) \in \mathcal{T}$, controller π, exploration bounds $\delta_x, \delta_u > 0$, stepsize $\gamma > 0$.

for $k = 0, 1, 2 \dots$ **do**

Learning

Step $L1$: gather $i = 1, \dots, L$ trajectories perturbation $(\hat{\boldsymbol{x}}^{i,k}, \hat{\boldsymbol{u}}^{i,k})$ of $(\boldsymbol{x}^k, \boldsymbol{u}^k)$ via closed-loop experiment as in (V.21):

$$\begin{aligned} \hat{u}_t^{i,k} &= \pi(x_t^k, u_t^k, \hat{x}_t^{i,k}, t) + d_{u,t}^{i,k} \\ \hat{x}_{t+1}^{i,k} &= f(\hat{x}_t^{i,k}, \hat{u}_t^{i,k}), \qquad \hat{x}_0^{i,k} = x_{\text{init}} + d_x^{i,k}. \end{aligned} \tag{V.15}$$

Step $L2$: build $\Delta X_t^k, \Delta U_t^k, \Delta X_t^{+,k}$ for $t = 0, \dots T-1$ as in (V.24):

$$\begin{aligned} \Delta X_t^k &= [\hat{x}_t^{1,k} - x_t^k, \dots, \hat{x}_t^{L,k} - x_t^k] \\ \Delta U_t^k &= [\hat{u}_t^{1,k} - u_t^k, \dots, \hat{u}_t^{L,k} - u_t^k] \\ \Delta X_t^{+,k} &= [\hat{x}_{t+1}^{1,k} - x_{t+1}^k, \dots, \hat{x}_{t+1}^{L,k} - x_{t+1}^k]. \end{aligned} \tag{V.16}$$

Step $L3$: for all $t = 0, \dots, T-1$, estimate the linearizations of the dynamics:

$$\begin{bmatrix} \hat{A}_t^k & \hat{B}_t^k \end{bmatrix} = \Delta X_t^{+,k} \begin{bmatrix} \Delta X_t^k \\ \Delta U_t^k \end{bmatrix}^\dagger. \tag{V.17}$$

Optimization

Step $O1$: solve the approximate problem

$$\begin{aligned} \min_{\Delta\hat{\boldsymbol{x}}, \Delta\hat{\boldsymbol{u}}} \quad & \sum_{t=0}^{T-1} \frac{1}{2} \begin{bmatrix} \Delta\hat{x}_t \\ \Delta\hat{u}_t \end{bmatrix}^\top \begin{bmatrix} Q_t^k & S_t^k \\ S_t^{\top,k} & R_t^k \end{bmatrix} \begin{bmatrix} \Delta\hat{x}_t \\ \Delta\hat{u}_t \end{bmatrix} + \begin{bmatrix} q_t^k \\ r_t^k \end{bmatrix}^\top \begin{bmatrix} \Delta\hat{x}_t \\ \Delta\hat{u}_t \end{bmatrix} + \Delta\hat{x}_T^\top Q_T^k \Delta\hat{x}_T + q_T^{k\top} \Delta\hat{x}_T \\ \text{subj. to} \quad & \Delta\hat{x}_{t+1} = \hat{A}_t^k \Delta\hat{x}_t + \hat{B}_t^k \Delta\hat{u}_t, \\ & \Delta\hat{x}_0 = 0, \quad t = 0, \dots T-1. \end{aligned} \tag{V.18}$$

Step $O2$: update curve $(\boldsymbol{\alpha}^{k+1}, \boldsymbol{\mu}^{k+1})$

$$\begin{aligned} \alpha_t^{k+1} &= x_t^k + \gamma \Delta\hat{x}_t^k \\ \mu_t^{k+1} &= u_t^k + \gamma \Delta\hat{u}_t^k, \end{aligned} \tag{V.19}$$

with $t = 0, \dots, T-1$.

Step $O3$: obtain new trajectory $(\boldsymbol{x}^{k+1}, \boldsymbol{u}^{k+1})$

$$\begin{aligned} u_t^{k+1} &= \pi(\alpha_t^{k+1}, \mu_t^{k+1}, x_t^{k+1}, t) \\ x_{t+1}^{k+1} &= f(x_t^{k+1}, u_t^{k+1}), \quad x_0^{k+1} = x_{\text{init}}, \end{aligned} \tag{V.20}$$

with $t = 0, \dots, T-1$.

end for

generating a solution to problem (V.8) leveraging on successive *learning* and *optimization* steps, which are denoted in Algorithm 7 by prefixes $\boldsymbol{L}$ and $\boldsymbol{O}$, respectively.

Step $L1$: Closed-loop data collection At each iteration k, a set of state-input data are collected in the neighbourhood of the current trajectory $(\boldsymbol{x}^k, \boldsymbol{u}^k)$ by successive experimental sessions (or simulations) with additional exploration noise. The collected data are then used to estimate the Jacobians of the unknown dynamics (V.1). More in detail, the i^{th} perturbation of the nominal trajectory $(\boldsymbol{x}^k, \boldsymbol{u}^k)$, denoted as $(\hat{\boldsymbol{x}}^{i,k}, \hat{\boldsymbol{u}}^{i,k})$, is obtained from the real system by integrating the closed-loop dynamics

$$\begin{aligned} \hat{u}_t^{i,k} &= \pi(x_t^k, u_t^k, \hat{x}_t^{i,k}, t) + d_{u,t}^{i,k} \\ \hat{x}_{t+1}^{i,k} &= f(\hat{x}_t^{i,k}, \hat{u}_t^{i,k}), \qquad \hat{x}_0^{i,k} = x_{\text{init}} + d_x^{i,k}, \end{aligned} \tag{V.21}$$

where $d_x^{i,k} \in \mathbb{R}^n, d_{u,t}^{i,k} \in \mathbb{R}^m$ are appropriate exploration dithers injected to guarantee a successful identification. Notice the exploration dither may be added by the experimenter (especially in the case of simulation) or it may be spontaneous disturbances. In this article, for simplicity reasons, we suppose it is a degree of freedom introduced by the experimenter. We introduce now two hypotheses to characterize the control law π and the closed-loop experiments (V.21).

Assumption V.3 (Properties of π). *The state-feedback control law $\pi(\alpha, \mu, x, t)$ is twice continuously differentiable in its arguments, i.e., π is C^2 an designed such that $\pi(\alpha, \mu, \alpha, t) = \mu$ holds for all $\alpha \in \mathbb{R}^n, \mu \in \mathbb{R}^n, t \in \mathbb{N}$.*

This assumption implies that when the reference curve $(\boldsymbol{\alpha}, \boldsymbol{\mu})$ is a trajectory of the dynamics (V.1) as per Definition V.1, i.e., $(\boldsymbol{\alpha}, \boldsymbol{\mu}) \in \mathcal{T}$, it holds that the resulting trajectory $(\boldsymbol{x}, \boldsymbol{u})$ of the closed-loop system (V.9) is such that $(\boldsymbol{x}, \boldsymbol{u}) = (\boldsymbol{\alpha}, \boldsymbol{\mu})$. In fact, under Assumption V.3 and if $x_t = \alpha_t$, the input chosen by the closed-loop dynamics (V.9) is given by

$$u_t = \pi(\alpha_t, \mu_t, x_t, t) = \pi(\alpha_t, \mu_t, \alpha_t, t) = \mu_t, \tag{V.22}$$

resulting in $f(x_t, u_t) = f(\alpha_t, \mu_t) = \alpha_{t+1}$ if $(\boldsymbol{\alpha}, \boldsymbol{\mu}) \in \mathcal{T}$. As an example, a control law $\pi(\alpha, \mu, x, t) = \mu + K_t(\alpha - x)$ respects this hypothesis.

Remark V.3. Due to the genericity of the framework, we do not provide here a "standard" way to build the policy π, which should be designed for the specific application and leveraging the specific knowledge of the system (as an example, in robotics application one may use standard robust, adaptive, or sliding mode control techniques). ◇

Assumption V.4 (Dither boundedness). *The exploration dithers $d_{u,t}^{i,k}, d_x^{i,k}$ in (V.21) are known and uniformly bounded, i.e.,*

$$|d_x^{i,k}| \leq \delta_x \qquad\qquad |d_{u,t}^{i,k}| \leq \delta_u \tag{V.23}$$

for all $i = 1, \ldots, L$, $k \in \mathbb{N}$, $t = 0, \ldots, T-1$.

While the presence of exploration noise is necessary for performing a good identification, its integration in time may let nonlinearities show up, thus leaving the neighbourhood of the current trajectory $(\boldsymbol{x}^k, \boldsymbol{u}^k)$ and ruining the estimated linearization of the dynamics.

Remark V.4. In practical implementations, the experimenter may not be able to inject an arbitrarily small dither signal. The tracking capabilities of the closed-loop control law π become thus fundamental to allow the injection of stronger dithers without leaving the neighbourhood of $(\boldsymbol{x}^k, \boldsymbol{u}^k)$, thus, influencing the convergence properties of DATA-DRIVEN PRONTO. This aspect will be shown in the numerical example. ◇

Step $L2 - L3$: LTV dynamics identification For all $i = 1, \dots, L$, the perturbations $(\hat{\boldsymbol{x}}^{i,k}, \hat{\boldsymbol{u}}^{i,k})$ obtained via (V.21) are used to build matrices $\Delta X_t^k, \Delta U_t^k, \Delta X_t^{+,k}$, which are data batches stacking the differences between all the perturbations and the nominal trajectory, namely

$$\begin{aligned}
\Delta X_t^k &= [\hat{x}_t^{1,k} - x_t^k, \dots, \hat{x}_t^{L,k} - x_t^k] \in \mathbb{R}^{n \times L} \\
\Delta U_t^k &= [\hat{u}_t^{1,k} - u_t^k, \dots, \hat{u}_t^{L,k} - u_t^k] \in \mathbb{R}^{m \times L} \\
\Delta X_t^{+,k} &= [\hat{x}_{t+1}^{1,k} - x_{t+1}^k, \dots, \hat{x}_{t+1}^{L,k} - x_{t+1}^k] \in \mathbb{R}^{n \times L}.
\end{aligned} \tag{V.24}$$

Data $\Delta X_t^k, \Delta U_t^k, \Delta X_t^{+,k}$ are used to perform a least-squares identification (V.17) of the matrices A_t^k, B_t^k for all t. The following assumption ensures the identification step to be well posed.

Assumption V.5 (Well-posed identification)**.** *For all $k \in \mathbb{N}$ and for all $t = 0, \dots, T-1$, there exists $M > 0$ for which*

$$\begin{bmatrix} \Delta X_t^k \\ \Delta U_t^k \end{bmatrix} \in \mathcal{F}_M, \tag{V.25}$$

where

$$\mathcal{F}_M \coloneqq \left\{ F \in \mathbb{R}^{(n+m) \times L} : \kappa\left(FF^\top\right) \leq M \right\}. \tag{V.26}$$

Although it is challenging to establish sufficient conditions on the exploration noise to guarantee Assumption V.5 for generic nonlinear dynamics, the proposed strategy allows for the collection of an arbitrary number of trajectory perturbations and the verification of whether the conditions are met using the gathered data. Notice that DATA-DRIVEN PRONTO does not identify a specific model of the dynamics, but identifies linearizations of the dynamics about specific trajectories. This avoids the introduction of possible parametrization errors when performing the identification.

Step $O1$: Data-driven descent calculation Finally, in Algorithm 7 the descent direction is obtained by solving the data-based problem (V.18) where the linearized linear time-varying dynamics is replaced by its estimation based on data (V.24). The remaining steps remain identical to those in PRONTO.

Main result

In the following we present the main result of this chapter. For the sake of readability, for all iteration k we denote the solution estimate provided by Algorithm 7 as $\eta^k = (\boldsymbol{x}^k, \boldsymbol{u}^k)$ while each isolated local minimum solution of (V.8) is denoted as $\eta^\star = (\boldsymbol{x}^\star, \boldsymbol{u}^\star)$.

Theorem V.1. *Consider Algorithm 7 with unitary stepsize $\gamma = 1$. Let Assumptions V.1, V.2, V.3, V.4 and V.5 hold. Then, there exist dither bounds $\bar{\delta}_x, \bar{\delta}_u > 0$, an iteration $K \in \mathbb{N}$, and a radius $\rho > 0$ such that, if $\delta_x \in (0, \bar{\delta}_x)$, $\delta_u \in (0, \bar{\delta}_u)$, and $\eta^0 \in \mathbb{B}_\rho(\eta^\star) \cap \mathcal{T}$ then*

$$\begin{aligned} |\eta^k - \eta^\star| &\leq \phi(|\eta_0 - \eta^\star|, k) \quad \forall k < K \\ |\eta^k - \eta^\star| &\leq b(\delta_x, \delta_u) \quad \forall k \geq K, \end{aligned} \tag{V.27}$$

where ϕ is a class $\mathcal{KL}$ function and $b(\delta_x, \delta_u)$ is class $\mathcal{K}$ function of δ_x, δ_u. ♛

The proof of Theorem V.1 is given in Appendix V.7.6.

Remark V.5. The proof is given for unitary stepsize γ; however, there is a tradeoff between the basin of attraction of the algorithm and the amplitude of the chosen stepsize. It is also possible to choose γ with techniques such as Armijo's rule (depending on the type availability of simulation setup / experiments that one is able to implement / perform). ◇

V.3.2 Algorithm analysis

We approach the proof of the main theorem in two steps. First, by realizing that the solution update strategy implemented by Algorithm 7 can be viewed as a perturbed version of Algorithm 6 (where the perturbation is introduced by the Jacobian estimation), we prove that under the assumption that the difference between the descent direction obtained by solving problems (V.13) and (V.18) is sufficiently small, we can ensure convergence to a neighborhood of the solution $\eta^\star$. Second, we show how to pick the algorithm parameters δ_x and δ_u to ensure that the difference between the descent direction obtained by solving problems (V.13) and (V.18) is arbitrarily small. For the sake of clarity, we denote the current (unperturbed) trajectory at each iteration as $\eta^k = (\boldsymbol{x}^k, \boldsymbol{u}^k) \in \mathbb{R}^s$. To denote solutions of problems (V.13) and (V.18), i.e., the descent direction, we introduce the symbols $\zeta^k = (\boldsymbol{\Delta x}^k, \boldsymbol{\Delta u}^k) \in \mathbb{R}^s$ and $\hat{\zeta}^k = (\boldsymbol{\Delta \hat{x}}^k, \boldsymbol{\Delta \hat{u}}^k) \in \mathbb{R}^s$, respectively.

Practical stability of DD-PRONTO

We now study the convergence properties of PRONTO in case of errors in the descent direction calculation. At first, we briefly introduce a state-space reformulation of the exact version of PRONTO. Next, we show how to modify PRONTO to obtain Algorithm 6 (i.e., we introduce the cost regularization) and how its convergence properties are affected by the cost approximation. At last, we study Algorithm 7 as a perturbed version of Algorithm 6. We start by rewriting problem (V.8) as

$$\begin{aligned} \min_{\eta} \quad & \ell(\eta) \\ \text{subj. to} \quad & \eta \in \mathcal{T}, \end{aligned} \tag{V.28}$$

where $\ell(\cdot)$ and $\mathcal{T}$ are given in (V.7) and Def. V.2, respectively. We denote the projection operator given by the closed-loop system (V.9) as $\mathcal{P} : \mathbb{R}^s \to \mathcal{T}$. The idea is to project any curve ξ into a trajectory η by leveraging the tracking controller π and integrating the dynamics. Notice that Assumption V.3 ensures that i) $\mathcal{P}(\eta) = \eta$ if $\eta \in \mathcal{T}$ and ii) $\mathcal{P}(\xi) \in \mathcal{T}$ if $\xi \notin \mathcal{T}$, which are properties required to call $\mathcal{P}$ a projection. Further details on the projection operator, its derivatives and their properties can be found in [91]. We then embed the projection in the cost by defining $g(\xi) := \ell(\mathcal{P}(\xi))$ to obtain the *unconstrained* problem formulation

$$\min_{\eta} \quad g(\eta), \tag{V.29}$$

which is shown to have the same isolated minima of problem (V.28) [90]. Problem (V.29) is then solved via a quasi Newton's method, which can be rewritten as the autonomous dynamical system

$$\begin{aligned} \bar{\zeta}^k &= \operatorname*{argmin}_{\zeta \in T_{\eta^k}\mathcal{T}} \left(\frac{1}{2} \zeta^\top \nabla^2 g(\eta^k) \zeta + \nabla g(\eta^k)^\top \zeta \right) \\ \eta^{k+1} &= \mathcal{P}(\eta^k + \bar{\zeta}^k), \end{aligned} \tag{V.30}$$

where $T_{\eta^k}\mathcal{T}$ is the space tangent to the trajectory manifold $\mathcal{T}$ at η^k. Notice this is not a "pure" Newton method since it searches for updates in the tangent space $T_{\eta^k}\mathcal{T}$ and since it projects the updated curve onto the trajectory manifold. This allows to satisfy the dynamic constraints at each algorithm iteration, instead of satisfying it only asymptotically. Being PRONTO a projected SQP algorithm [18, Thm. 3], it is shown that the autonomous system (V.30), under Assumptions V.2, V.1 and V.3, has locally exponentially stable equilibria in isolated solutions of the optimal control problem $\eta^\star$ (or, more in general, in points which solve KKT conditions) [18, Thm. 2]. In the next lemma, we study Algorithm 6 which implements a regularized version of PRONTO (namely, Alg. 6) where the Hessians of the cost function g are approximated considering only the derivatives of the cost function ℓ (instead of their composition with $\mathcal{P}$). More in detail, Alg. 6 implements the iterative update

$$\begin{aligned} \zeta^k &= \operatorname*{argmin}_{\zeta \in T_{\eta^k}\mathcal{T}} \left(\frac{1}{2} \zeta^\top \nabla^2 \ell(\eta^k) \zeta + \nabla \ell^\top(\eta^k) \zeta \right) \\ \eta^{k+1} &= \mathcal{P}(\eta^k + \zeta^k). \end{aligned} \tag{V.31}$$

Notice that, in (V.31), only the second order derivatives of the cost function ℓ are considered, instead of the composition g (cf. dynamics (V.30)). The following lemma provides stability guarantees of the optimal solution for Algorithm 6.

Lemma V.1. *[Exponential stability of optimal solution] Consider the discrete-time autonomous dynamical system* (V.31). *Let Assumptions V.2, V.1 and V.3 hold. Then, the equilibrium $\eta^\star$ is Locally Exponentially Stable.* ♙

The proof is provided in Appendix V.7.1.
Given the stability properties of $\eta^\star$ for system (V.31), we are now able to provide a theoretical guarantee for a perturbed version of the algorithm.

Lemma V.2. *[Practical stability of optimal solution] Consider the non-autonomous dynamical system given by*

$$\begin{aligned} \zeta^k &= \underset{\zeta \in T_{\eta^k}\mathcal{T}}{\operatorname{argmin}} \left(\frac{1}{2} \zeta^\top \nabla^2 \ell(\eta^k) \zeta + \nabla \ell^\top(\eta^k) \zeta \right) \\ \eta^{k+1} &= \mathcal{P}\left(\eta^k + \zeta^k + \Delta\zeta^k \right), \end{aligned} \tag{V.32}$$

where $\Delta\zeta^k$ is a perturbation. Let Assumptions V.1, V.2 and V.3 hold. Let $\|\Delta\zeta^k\| \leq \delta_\zeta$ for all $k \in \mathbb{N}$ and for some $\delta_\zeta > 0$. Then, there exists $\bar{\delta}_\zeta > 0$ such that, if $\delta_\zeta \leq \bar{\delta}_\zeta$, the equilibrium $\eta^\star$ is Locally Uniformly Ultimately Bounded, i.e., there exists $K \in \mathbb{N}$, $p > 0$, class $\mathcal{KL}$ function ϕ and class $\mathcal{K}$ function $b(\delta_\zeta)$ such that, if $\eta^0 \in \mathbb{B}_p(\eta^\star)$, then it holds

$$\begin{aligned} |\eta^k - \eta^\star| &\leq \phi(|\eta_0 - \eta^\star|, k) && \forall k < K \\ |\eta^k - \eta^\star| &\leq b(\delta_\zeta) && \forall k \geq K. \end{aligned} \tag{V.33}$$

♙

The proof is provided in Appendix V.7.2.

Data-driven descent error

In the following, we present formal error bounds about the calculation of the descent direction based solely on input dithers. Since the results in this section are independent on the specific iteration k of the algorithm, we omit the superscript k for clarity. Additionally, to compact the notation, we define $\boldsymbol{d}_x$ and $\boldsymbol{d}_u$ as the stack of exploration dithers across all time steps and perturbations, i.e.,

$$\begin{aligned} \boldsymbol{d}_x &:= (d_x^1, \dots, d_x^L) \in \mathbb{R}^{nL} \\ \boldsymbol{d}_u &:= (d_{u,0}^1, \dots, d_{u,T-1}^1, \dots, d_{u,0}^L, \dots, d_{u,T-1}^L) \in \mathbb{R}^{s_u L}. \end{aligned} \tag{V.34}$$

The core idea is to show that the error between the data-driven descent direction $\hat{\zeta} = (\boldsymbol{\Delta\hat{x}}, \boldsymbol{\Delta\hat{u}})$ solution of (V.18) and the exact descent direction $\zeta = (\boldsymbol{\Delta x}, \boldsymbol{\Delta u})$ solution of (V.13) is a smooth function of the exploration dithers $\boldsymbol{d}_x, \boldsymbol{d}_u$. Let us denote the stack of data matrices (V.24) obtained by collecting L

perturbed trajectories from (V.21) as

$$\begin{aligned}\boldsymbol{\Delta X} &\coloneqq (\Delta X_0, \ldots, \Delta X_{T-1}) \in \mathbb{R}^{nTL} \\ \boldsymbol{\Delta U} &\coloneqq (\Delta U_0, \ldots, \Delta U_{T-1}) \in \mathbb{R}^{mTL}\end{aligned} \tag{V.35}$$

First, we show that, for a given trajectory η, the data matrices $(\boldsymbol{\Delta X}, \boldsymbol{\Delta U})$ in (V.35) are a smooth function of the exploration dithers $\boldsymbol{d}_x, \boldsymbol{d}_u$, namely we show the function

$$\begin{aligned}\Delta_{XU} :& \mathbb{R}^s \times \mathbb{R}^{nL} \times \mathbb{R}^{s_u L} \to \mathbb{R}^{nTL} \times \mathbb{R}^{mTL} \\ & \eta, \boldsymbol{d}_x, \boldsymbol{d}_u \longmapsto (\boldsymbol{\Delta X}, \boldsymbol{\Delta U}).\end{aligned} \tag{V.36}$$

is smooth in its arguments. Notice, $\Delta_{XU}(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u)$ depends both on the exploration dithers of all $i = 1, \ldots, L$ collected perturbations and on the current trajectory η, where the dependence on η accounts for the closed-loop dynamics (V.21). The next lemma formally proves the claimed property.

Lemma V.3. *Let Assumption V.1 and V.3 hold. Then,* Δ_{XU} *in* (V.36) *is a* C^1 *function of the current trajectory* η *and of all the exploration dithers* $\boldsymbol{d}_x, \boldsymbol{d}_u$*. Furthermore,*

$$\Delta_{XU}(\eta, 0, 0) = 0 \tag{V.37}$$

for all system trajectories $\eta \in \mathcal{T}$*.* ♙

The proof is provided in Appendix V.7.3.

Notice that a continuously differentiable function is also Lipschitz continuous in bounded sets; furthermore, being zero in zero, it is possible to obtain a bound for its norm which is linear and which is zero in zero. Denote now the error between the matrices estimated using data via (V.17) and the exact Jacobians of f about the current trajectory η at time instant t as

$$\Delta A_t \coloneqq \hat{A}_t - A_t, \qquad \Delta B_t \coloneqq \hat{B}_t - B_t. \tag{V.38}$$

and define their stack, for all t, as

$$\begin{aligned}\boldsymbol{\Delta A} &\coloneqq (\Delta A_0, \ldots, \Delta A_{T-1}) \in \mathbb{R}^{n^2 T} \\ \boldsymbol{\Delta B} &\coloneqq (\Delta B_0, \ldots, \Delta B_{T-1}) \in \mathbb{R}^{nmT}.\end{aligned} \tag{V.39}$$

The stack of estimation errors in (V.39) can be written as a function of both the current trajectory η and the data matrices $\boldsymbol{\Delta X}, \boldsymbol{\Delta U}$, which define the exact Jacobians and their data-based estimation, respectively. We denote this function as

$$\begin{aligned}\Delta_{AB} :& \mathbb{R}^s \times \mathcal{F}_M^T \to \mathbb{R}^{n^2 T} \times \mathbb{R}^{nmT} \\ & \eta, \boldsymbol{\Delta X}, \boldsymbol{\Delta U} \longmapsto (\boldsymbol{\Delta A}, \boldsymbol{\Delta B}),\end{aligned} \tag{V.40}$$

where $\mathcal{F}_M^T \coloneqq \mathcal{F}_M \times \mathcal{F}_M \times \ldots$ and $\mathcal{F}_M \subset \mathbb{R}^{(n+m)\times L}$ is given in (V.26). The next lemma provides formal guarantees for this relation.

Lemma V.4. *Let Assumption V.1 hold. Then,* Δ_{AB} *in* (V.40) *is a* C^1 *function of* η *and* $\mathbf{\Delta X}, \mathbf{\Delta U}$. *Furthermore, for each* $\eta \in \mathcal{T}$ *and bounded* $\mathcal{K} \subset \mathcal{F}_M^{TL}$, *there exists* $g(\eta, \mathcal{K}) > 0$ *such that, if* $(\mathbf{\Delta X}, \mathbf{\Delta U}) \in \mathcal{K}$, *then*

$$|\Delta_{AB}(\eta, \mathbf{\Delta X}, \mathbf{\Delta U})| \leq g(\eta, \mathcal{K})(|\mathbf{\Delta X}| + |\mathbf{\Delta U}|). \tag{V.41}$$

♙

The proof is provided in Appendix V.7.4.

Remark V.6. It is not possible to entirely eliminate the estimation error given by Δ_{AB}, as this would necessitate all trajectory perturbations $\hat{\eta}^i$ to coincide exactly with the current trajectory η, leading to singular matrices $(\Delta X_t, \Delta U_t) \notin \mathcal{F}_M$ for any $M > 0$. Nevertheless, it is possible for the matrices ΔX_t and ΔU_t to diminish in the directions that ensure a bounded condition number. ◇

We now consider the approximated problem (V.18) and the full knowledge problem (V.13). Notice that, for a given trajectory η, these two problems differ only in the constraint represented by the LTV dynamics. The descent direction error $\Delta_\zeta := \hat{\zeta} - \zeta$ can be expressed as function of both the current trajectory η and the estimation errors (V.39), so we define

$$\begin{aligned} \Delta_\zeta :& \mathbb{R}^s \times \mathbb{R}^{n^2 T} \times \mathbb{R}^{nmT} \to \mathbb{R}^s \\ & \eta, \mathbf{\Delta A}, \mathbf{\Delta B} \longmapsto \hat{\zeta} - \zeta, \end{aligned} \tag{V.42}$$

where, for a given trajectory η, $\hat{\zeta} = (\mathbf{\Delta \hat{x}}, \mathbf{\Delta \hat{u}})$ and $\zeta = (\mathbf{\Delta x}, \mathbf{\Delta u})$ are the solutions of (V.18) and (V.13), respectively. The dependence of Δ_ζ on η accounts for the cost matrices being Hessian and Jacobian of the cost function ℓ evaluated in the current trajectory η. The next lemma provides formal guarantees for this relation.

Lemma V.5. *Let Assumption V.1 and V.2 hold. For all system trajectories* $\eta \in \mathcal{T}$, *there exists a continuous function* $\delta_{AB} : \mathcal{T} \to \mathbb{R}_{>0}$, *such that, if* $|(\mathbf{\Delta A}, \mathbf{\Delta B})| \leq \delta_{AB}(\eta)$, *then* Δ_ζ *in* (V.42) *is a* C^1 *function of* $\eta, \mathbf{\Delta A}, \mathbf{\Delta B}$ *such that*

$$\Delta_\zeta(\eta, 0, 0) = 0. \tag{V.43}$$

♙

The proof is provided in Appendix V.7.5.

In other words, Lemma V.5 provides us with a bound $\delta_{AB}(\eta)$ on the identification error which, if respected, ensures that the error function $\Delta_\zeta(\eta, \mathbf{\Delta A}, \mathbf{\Delta B})$ is smooth in its arguments.

V.3.3 Numerical simulations

In this section, we demostrate the capabilities of DATA-DRIVEN PRONTO by solving an optimal control problem for a nonlinear underactuated robot with unknown dynamics. First, we present the setup and the nonlinear optimal control problem, then we show the performances of DATA-DRIVEN PRONTO under two different controllers π, the first more accurate and the second less reliable. The robot, cf. [195, 233], consists of two links and one actuator on the first joint (see Figure V.1).

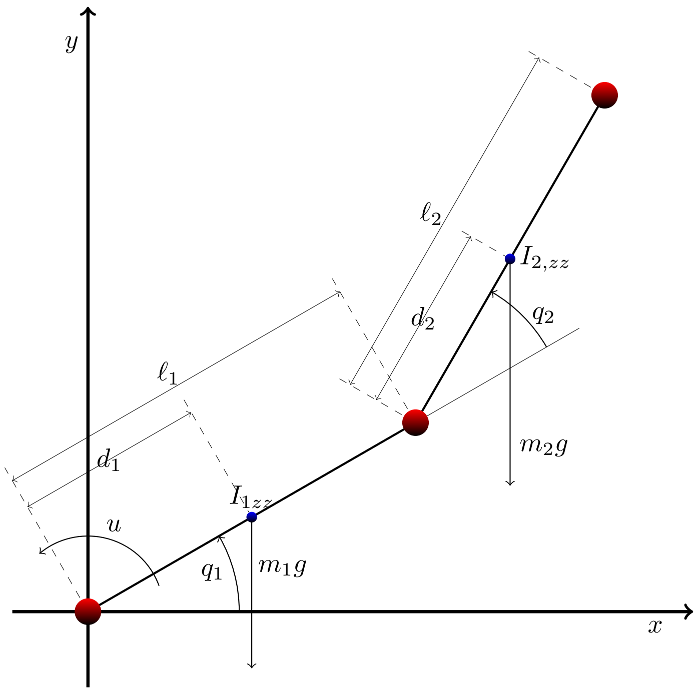

Figure V.1: The pendubot.

Its dynamics read as

$$M(q)\begin{bmatrix}\ddot{q}_1\\ \ddot{q}_2\end{bmatrix} + (C(q,\dot{q}) + F)\begin{bmatrix}\dot{q}_1\\ \dot{q}_2\end{bmatrix} + g(q) = \begin{bmatrix}u\\ 0\end{bmatrix}, \tag{V.44}$$

where $q = (q_1, q_2) \in [0, 2\pi]^2$ stacks the two joint angles $q_1, q_2 \in \mathbb{R}$, $u \in \mathbb{R}$ is the input torque on the first joint, $M(q) \in \mathbb{R}^{2\times 2}$ is the inertia matrix, $F \in \mathbb{R}^{2\times 2}$ accounts for friction, $C(q,\dot{q}) \in \mathbb{R}^2$ includes the Coriolis and centrifugal forces and $g(q) \in \mathbb{R}^2$ is the gravitational term. The matrices in (V.44) are defined as

$$\begin{aligned}
M(q) &:= \begin{bmatrix} a_1 + a_2 + 2a_3\cos(q_2) & a_2 + a_3\cos(q_2)\\ a_2 + a_3\cos(q_2) & a_2 \end{bmatrix}\\
C(q,\dot{q}) &:= \begin{bmatrix} -a_3\sin(q_2)\dot{q}_2 & -a_3\sin(q_2)(\dot{q}_1+\dot{q}_2)\\ a_3\sin(q_2)\dot{q}_1 & 0 \end{bmatrix}\\
g(q) &:= \begin{bmatrix} a_4\cos(q_1) + a_5\cos(q_1+q_2)\\ a_5\cos(q_1+q_2) \end{bmatrix}\\
F &:= \mathrm{diag}(f_1, f_2)
\end{aligned} \tag{V.45}$$

Table V.1: Physical parameters of the pendubot

Parameter	Value	Parameter	Value
m_1[Kg]	1.2	m_2[Kg]	1.2
ℓ_1[m]	1	ℓ_2[m]	1
d_1[m]	0.6	d_2[m]	0.6
$I_{1,zz}$[Kgm^2]	0.1	$I_{2,zz}$[Kgm^2]	0.1
f_1[Ns/rad]	1	f_2[Ns/rad]	1

where

$$\begin{aligned} &a_1 := I_{1,zz} + m_1 d_1^2 + m_2 \ell_1^2 \\ &a_2 := I_{2,zz} + m_2 d_2^2, \qquad a_3 := m_2 \ell_1 d_2 \\ &a_4 := g(m_1 d_1 + m_2 \ell_1), \qquad a_5 := g m_2 d_2. \end{aligned} \tag{V.46}$$

The physical parameters used to simulate the system are summarized in Table V.1. From (V.44), we obtain a state space model with state variable $x = (q_1, q_2, \dot{q}_1, \dot{q}_2)$. The dynamics is then discretized via forward Euler integration of step $dt = 0.01$s over a simulation time of $T = 10$s. The cost function is a quadratic cost function designed to follow a step reference $(\bar{x}^*, \bar{u}^*)$:

$$\begin{aligned} \ell(\bar{x}, \bar{u}) = \sum_{t=0}^{T-1} &(x_t - x_t^*)^\top Q (x_t - x_t^*) + \\ &(u_t - u_t^*)^\top R (u_t - u_t^*) + (x_T - x_T^*)^\top Q_T (x_T - x_T^*), \end{aligned}$$

with $Q = \operatorname{diag}(10^2, 10^2, 10, 10)$, $Q_T = Q$ and $R = 5 \cdot 10^3$. The reference state curve is a step from an initial (unstable) equilibrium condition $x_0 = (\frac{\pi}{8}, \frac{3}{8}\pi, 0, 0)$ to the final (unstable) equilibrium $x_T = (\frac{\pi}{4}, \frac{\pi}{4}, 0, 0)$. The reference input curve compensates for the gravity term at the two equilibrium position. The initial trajectory η^0 for the algorithm is chosen as the standstill robot in the starting position.

Exact-parameters projection operator

In this first example, the controller π in (V.9) and (V.21) is designed at each iteration k by solving a finite-horizon LQR problem over the linearized dynamics about the current nominal trajectory η^k.
In order to collect the perturbations of the current trajectory at each iteration, we add the exploration dither as in (V.21), with $d_{u,\cdot} \sim \mathcal{U}(0, \delta_u)$Nm and $d_x \sim \mathcal{U}(-\delta_x, \delta_x)$rad, with $\delta_u = 0.001, \delta_x = 0.01$. At each iteration, we collect $L = 6$ perturbed trajectories.
In Figure V.2, we plot the iterations of the DATA-DRIVEN PRONTO for all the states and the input. We use the Armijo rule to choose the stepsize, and we stop the algorithm when the update $\hat{\zeta}^k$ is not a descent direction, namely, when we have reached the ball about the optimal solution $\eta^\star$ in which the

cost does not improve significantly. In Figure V.3, we plot the norm of the descent direction $\hat{\zeta}^k$.

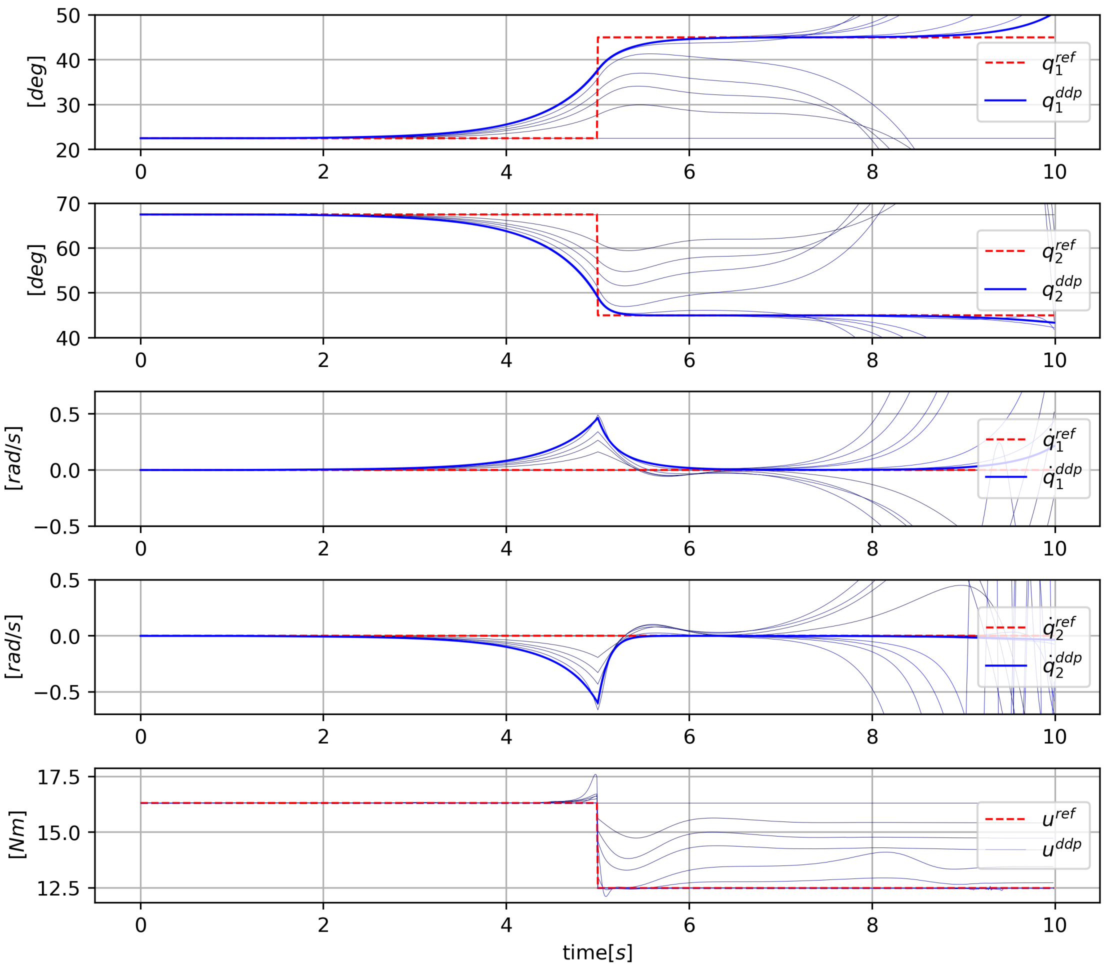

Figure V.2: In blue the reference curves for the states and the input. In red, the result of DATA-DRIVEN PRONTO.

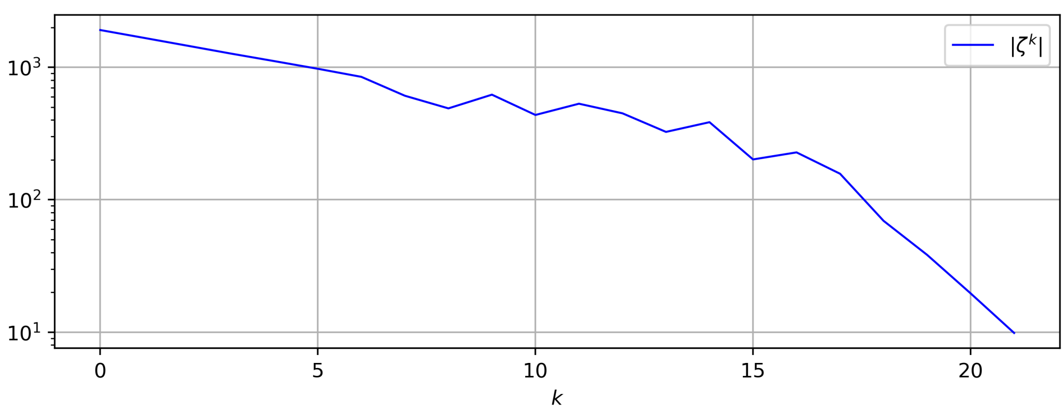

Figure V.3: Norm of descent direction.

In order to better see the convergence properties of DATA-DRIVEN PRONTO depending on the hyperparameters δ_x, δ_u, we test 15 times the algorithm convergence, solving the same optimal control problem. To do this, we halve at each instance the bound on the exploration noise, namely, $\delta_x^{i+1} = \delta_x^i/2$ and $\delta_u^{i+1} = \delta_u^i/2$. We choose at the first algorithm run $\delta_x^0 = 0.001$ and $\delta_u^0 = 0.01$, and we pick $d_x \sim \mathcal{U}(-\delta_x^i, \delta_x^i)$rad and $d_{u,t} \sim \mathcal{U}(0, \delta_u^i)$Nm, for all t. The results obtained, showed in Figure V.4, demonstrate the strictly increasing bound (in this case, exponential) between the amplitude of the dithers and the suboptimality of DD-PRONTO.

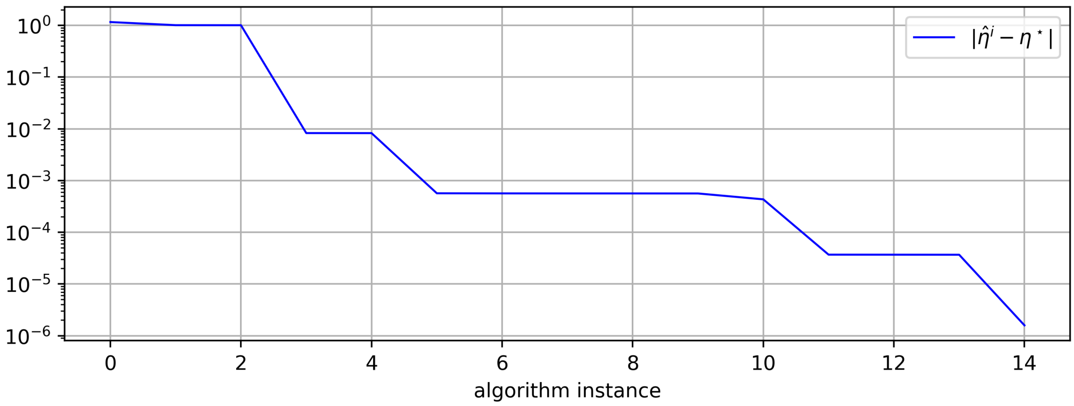

Figure V.4: Distance from the optimum depending on the dither amplitude.

Estimated-parameters projection operator

In this second example, the controller π in (V.9) and (V.21) is designed at each iteration k by solving a finite-horizon LQR problem over an inexact linearization of the current nominal trajectory η^k, relying on the estimated matrices A_t, B_t (which are used both to find the descent direction and to obtain an estimated controller).

In order to collect the perturbations of the current trajectory at each iteration, we add the exploration dither as in (V.21), with $d_{u,\cdot} \sim \mathcal{U}(0, \delta_u)$Nm and $d_x \sim \mathcal{U}(-\delta_x, \delta_x)$rad, with $\delta_u = 0.001, \delta_x = 0.0001$. At each iteration, we collect $L = 6$ perturbed trajectories. Notice that in this case, since the controller π is "less precise", dithers with the same amplitudes as in the previous example would lead to numerical instabilities, so we are forced to choose smaller amplitude dithers (which is not often possible in the practice). This strongly motivates the need of a good closed-loop policy π.

Figures V.5 and V.6 highlight how the importance of relying on a good control law π affects the quality of the obtained result.

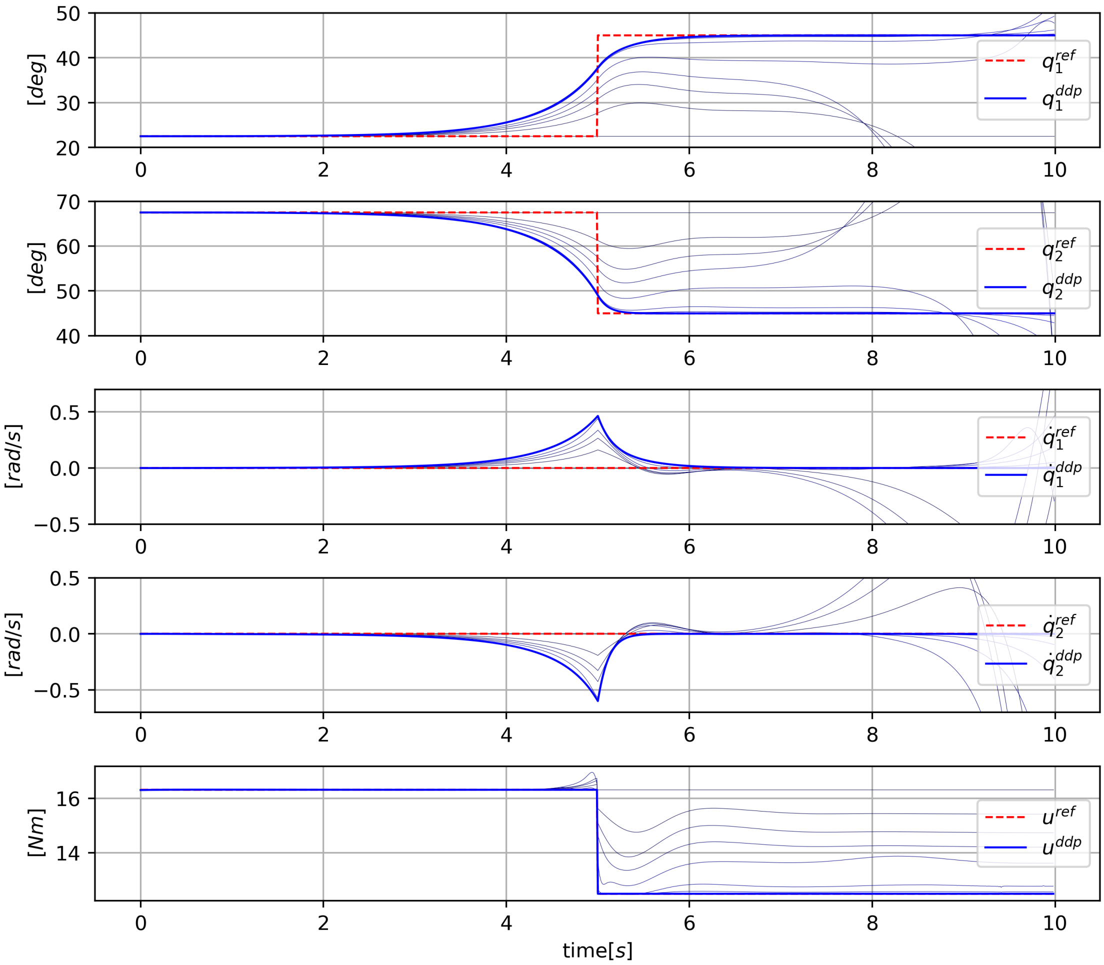

Figure V.5: In blue the reference curves for the states and the input. In red, the result of `DATA-DRIVEN PRONTO`.

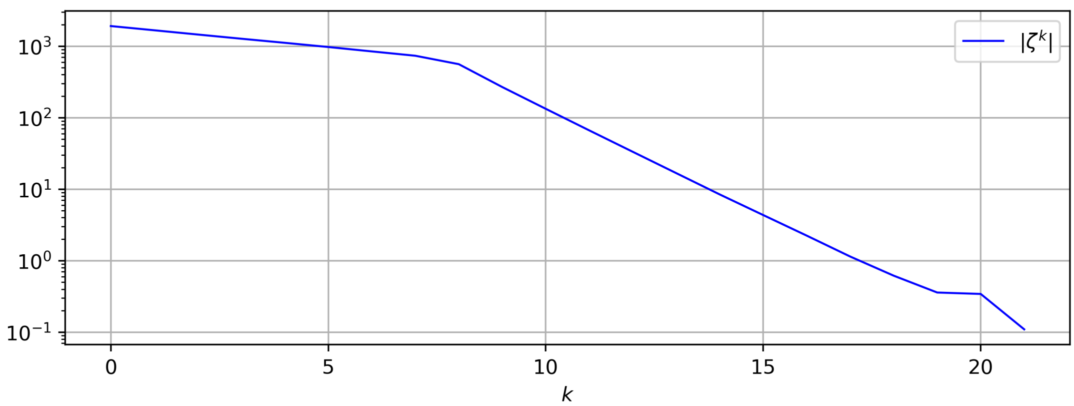

Figure V.6: Norm of descent direction.

V.3.4 Future work

In this chapter, we proposed a novel data-driven optimal control algorithm called DATA-DRIVEN PRONTO. This algorithm extends the applicability of PRONTO by removing the knowledge of the system model, relying on the ability to explore system trajectories. The main advantage of this algorithm is the possibility to overcome the suboptimality of model-based solution by estimating the linearizations about a trajectory instead of computing them from the model. Theoretical guarantees on the convergence of the algorithm have been given, together with insight on how to tune the design parameters of the algorithm. Future work will be directed to the real-world implementation of this algorithm.

Conclusions

Throghout this work, we showed how system theoretical tools can be used to solve modern problems regarding optimization, stabilization and learning guarantees in the context of uncertain systems. Motivated by the importance of enforcing the informativity of systems trajectories, we deeply studied the concepts of persistence of excitation and sufficient richness, retracing the origin of the definitions used today and proposing a clear separation between the two notions. Then, we developed new necessary and sufficient characterizations of sufficient richness for linear systems, presented in a notation unifying discrete- and continuous-time systems. The role of the system geometry in this context has been studied leveraging on the properties of sinusoidal inputs. We then faced the problem of information extraction from measured data. A new (to the authors' knowledge) type of observer, the "gazer", has been proposed and motivated for the model-free context in which this thesis is framed, and several possible designs (state feedback, single input single output, multi input multi output) have been given. We then showed a possible offline application of the gazer by studying how to design stabilizing gains from input-output data without the need of measuring the full state and its derivatives. Future work will be done in the direction of extending the obtained results to the nonlinear framework, as well as in obtaining a deeper understanding of the obtained equations. In the context of linear quadratic regulation, we developed an on-policy, online algorithm which deals with both the problems of optimization and system stabilization at the same time. Differently from other algorithms, to achieve both objectives we rely on a combination of model reference adaptive control and reinforcement learning, avoiding the requirement of an initial stabilizing policy. Furthermore, the obtained stability results are semi-global in the algorithms hyperparameters. Future work will be done in the direction of extending this approach to the output-feedback case. At last, we considered the case of model-free nonlinear optimal control, and we developed `DATA-DRIVEN PRONTO`, a numerical solver inspired to `PRONTO`. The underlying idea is to keep the data-efficient structure of solvers for optimal control problems whilst substituting the model knowledge with experiments on a simulator or a real setup. The presence of a stabilizing control law is included in the algorithm analysis, motivated by the need of guaranteeing a safe exploration of systems trajectories.

Appendix

V.4 Proofs for Chapter II

V.4.1 Proof of Lemma II.3

Since the arguments for discrete-time are analogous but more straightforward, we prove only the continuous-time part. Notice that for any $\mathrm{w} \in \Omega_d^{\mathrm{C}}$ and $\lambda > 0$, $\lambda\mathrm{w} \in \Omega_d^{\mathrm{C}}$, so Ω_d^{C} is a cone in $C_b^\infty(\mathbb{R}^d)$. Next, we show it is open. Let $\mathrm{w} \in \Omega_d^{\mathrm{C}}$. There exist $T, \alpha > 0$ such that

$$\int_t^{t+T} w(\tau)w(\tau)^\top d\tau \geq \alpha I, \quad \forall t \geq 0. \tag{V.47}$$

Choose any $\alpha' \in (0, \alpha)$ and any $\epsilon \in (0, \frac{\alpha'}{2TM})$, where $M := \|\mathrm{w}\|_\infty$. Choose any w' such that $\|\mathrm{w}' - \mathrm{w}\|_\infty = \|\Delta\mathrm{w}\|_\infty \leq \epsilon$. We have

$$\begin{aligned}
&\int_t^{t+T} w(\tau)'w(\tau)'^\top \mathrm{d}\tau = \\
&\quad = \int_t^{t+T} (w(\tau) + \Delta w(\tau))(w(\tau) + \Delta w(\tau))^\top \mathrm{d}\tau \\
&\quad \geq \alpha I + \int_t^{t+T} \left(w(\tau)\Delta w(\tau)^\top + \Delta w(\tau)w(\tau)^\top + \Delta w(\tau)\Delta w(\tau)^\top\right)\mathrm{d}\tau \\
&\quad \geq \alpha I + \int_t^{t+T} \left(w(\tau)\Delta w(\tau)^\top + \Delta w(\tau)w(\tau)^\top\right)\mathrm{d}\tau \\
&\quad \geq I\left(\alpha - \int_t^{t+T} 2M\epsilon \mathrm{d}\tau\right) \\
&\quad \geq I(\alpha - 2TM\epsilon) \geq (\alpha - \alpha')I > 0,
\end{aligned} \tag{V.48}$$

so, w' is PE. Therefore, for each point $\mathrm{w} \in \Omega_d^{\mathrm{C}}$, it is always possible to find an open ball about w which is still in Ω_d^{C}.

※

V.4.2 Proof of Lemma II.4

We divide the proof in three steps.

I) A useful characterization for PPE signals. Since $\mathrm{W}(\mathrm{w}) \coloneqq \mathrm{Q}^n(\mathrm{w})$ is PPE of degree at most $d' \leq d(n-1)$, there exist $i = 1, \dots, nd - d'$ orthonormal directions $z_i \in \mathbb{R}^{nd}$ such that for each $T \in \mathbb{N}, \epsilon > 0$ we can find $t \in \mathbb{N}$ such that

$$\sum_{\tau=t}^{t+T} |W_\tau^\top z_i| \leq \epsilon, \tag{V.49}$$

which means $|W_\tau^\top z_i| \leq \epsilon$, for all $\tau = t, \dots, t{+}T$. Consider an orthonormal basis for $\mathbb{R}^{nd}$, $\{z_1, \dots, z_{nd}\}$. Since $W_\tau = \sum_{i=1}^{nd} z_i W_\tau^\top z_i$, there exists $\lambda_\tau \in \mathbb{R}^{d'}$ such that

$$W_\tau = E\lambda_\tau + \tilde{W}_\tau \tag{V.50}$$

for all $\tau = t, \dots, t+T$, where $E = [z_{nd-d'+1}, \dots, z_{nd}] \in \mathbb{R}^{nd \times d'}$ stacks the directions in which w is PPE, the j-th component of λ_τ is given by $\lambda_\tau^j = W_\tau^\top z_j$, and $|\tilde{W}_\tau| \leq (nd - d')\epsilon$ by (V.49).

II) Signal sequences are constrained by PPE. Pick any $k \in \mathbb{N} : n \leq k \leq T$ and $\epsilon > 0$. Consider the signal $\mathrm{Q}^k(\mathrm{w})$ in the interval $\tau = t, \dots, t+T-k+1$, namely, $(w_\tau, \dots, w_{\tau+k-1}) \in \mathbb{R}^{kd}$. For each subsequence of k instants in the window $\tau = t, \dots, T-k+1$, we can write $(k-n+1)$ nd-dimensional equations of the type (V.50). Compactly, they read

$$M \underbrace{\begin{bmatrix} w_\tau \\ \vdots \\ w_{\tau+k-1} \end{bmatrix}}_{\coloneqq \hat{w}_\tau} = \tilde{E} \underbrace{\begin{bmatrix} \lambda_\tau \\ \vdots \\ \lambda_{\tau+k-n+1} \end{bmatrix}}_{\coloneqq \hat{\lambda}_\tau} + \underbrace{\begin{bmatrix} \tilde{W}_\tau \\ \vdots \\ \tilde{W}_{\tau+k-n+1} \end{bmatrix}}_{\coloneqq \tilde{w}_\tau}, \tag{V.51}$$

with $\hat{w}_\tau \in \mathbb{R}^{kd}$, $\hat{\lambda}_\tau \in \mathbb{R}^{d'(k-n+1)}$, $\tilde{w}_\tau \in \mathbb{R}^{nd(k-n+1)}$, and

$$M \coloneqq \begin{bmatrix} I_d & 0 & 0 & 0 & 0 & \dots \\ 0 & \ddots & 0 & 0 & 0 & \dots \\ 0 & 0 & \ddots & 0 & 0 & \dots \\ 0 & 0 & 0 & I_d & 0 & \dots \\ 0 & I_d & 0 & 0 & 0 & \dots \\ 0 & 0 & \ddots & 0 & 0 & \dots \\ 0 & 0 & 0 & \ddots & 0 & \dots \\ 0 & 0 & 0 & 0 & I_d & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \in \mathbb{R}^{nd(k-n+1)\times kd}$$
$$\tilde{E} \coloneqq I_{k-n+1} \otimes E \in \mathbb{R}^{nd(k-n+1)\times d'(k-n+1)}. \tag{V.52}$$

We are interested in characterizing the solutions $(\hat{w}_\tau, \hat{\lambda}_\tau)$ of (V.51), namely, the possible signals which fulfill the constraints which are imposed by PPE. In other words, we want to solve

$$\begin{bmatrix} M & -\tilde{E} \end{bmatrix} \begin{bmatrix} \hat{w}_\tau \\ \hat{\lambda}_\tau \end{bmatrix} = \tilde{w}_\tau. \tag{V.53}$$

III) A small enough degree of PPE overconstrains the signal sequence. Given the right pseudo-inverse $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger$ of $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}$, any solution $(\hat{w}_\tau, \hat{\lambda}_\tau)$ can be written as

$$\begin{bmatrix} \hat{w}_\tau \\ \hat{\lambda}_\tau \end{bmatrix} = \begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger \tilde{w}_\tau + v_\tau, \tag{V.54}$$

where $v_\tau \in \ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)$. Notice that, since $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}$ has full row rank, by the rank-nullity theorem it holds that

$$\begin{aligned} d'' &\coloneqq \dim\left(\ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) \\ &= \dim\left(\mathrm{dom}\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) - \dim\left(\mathrm{im}\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) \\ &= kd + d'(k-n+1) - nd(k-n+1) \\ &= (k-n)(d+d'-nd) + d', \end{aligned} \tag{V.55}$$

from which it is clear that, if $d' \leq d(n-1)$, then $d'' \leq d'$. Furthermore, notice that $|\tilde{w}_\tau| \leq \sqrt{nd(k-n+1)}\epsilon$ for all $\tau = t, \dots, t+T-k+1$. This means that for all $\tau = t, \dots, t+T-k+1$ we can write

$$\hat{w}_\tau = \begin{bmatrix} I_k & 0 \end{bmatrix} G v_\tau + \begin{bmatrix} I_k & 0 \end{bmatrix} \begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger \tilde{w}_\tau, \tag{V.56}$$

with G any surjective map

$$G : \mathbb{R}^{d''} \to \ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right), \tag{V.57}$$

and $\nu_\tau \in \mathbb{R}^{d''}$ such that $\nu_\tau = G\nu_\tau$. Since $\nu_\tau \in \mathbb{R}^{d''}$, it can span in time at most d'' directions, and the same holds for $\begin{bmatrix} I_k & 0 \end{bmatrix} G\nu_\tau$. Being $\tilde{w}_\tau$ arbitrarily small in the arbitrarily long interval $\tau = t, \dots, t+T-k+1$, we conclude that $\mathrm{Q}^k(\mathrm{w}) = (q^{k-1}\mathrm{w}, \dots, q^0\mathrm{w})$ is PPE of degree at most $d'' \le d'$.

V.4.3 Proof of Lemma II.5

Pick any time interval $[t, t+T]$ of arbitrary length $T \ge 0$ and an unitary direction $z \in \mathbb{R}^d$. Being w bounded, the quantity

$$\bar{w}(t,T,z) \coloneqq \int_t^{t+T} |z^\top w(\tau)| \mathrm{d}\tau, \tag{V.58}$$

is finite for all $t, T > 0$ and $z \in \mathbb{R}^d$. Choose an arbitrarily small $\alpha > 0$, and define the sets

$$\begin{aligned} T_{>\alpha}(t,T,z) &\coloneqq \{\tau \in [t, t+T] : |z^\top w(\tau)| > \alpha\} \\ T_{\le\alpha}(t,T,z) &\coloneqq \{\tau \in [t, t+T] : |z^\top w(\tau)| \le \alpha\}. \end{aligned} \tag{V.59}$$

Notice that, by continuity of w, $T_{>\alpha}(t,T,z)$ is a union of open sets for any t, T, z. Since $\|\mathrm{w}\|_\infty \le M$ for some $M > 0$, then we can find an upper bound for the measure $\mu(T_{\ge\alpha}(t,T,z))$:

$$\begin{aligned} \int_t^{t+T} |z^\top w(\tau)| \mathrm{d}\tau &= \bar{w}(t,T,z) \\ \int_{T_{>\alpha}} |z^\top w(\tau)| \mathrm{d}\tau + \int_{T_{\le\alpha}} |z^\top w(\tau)| \mathrm{d}\tau &= \bar{w}(t,T,z) \\ \int_{T_{>\alpha}} |z^\top w(\tau)| \mathrm{d}\tau &\le \bar{w}(t,T,z) \\ \int_{T_{>\alpha}} \alpha \mathrm{d}\tau &\le \bar{w}(t,T,z) \\ \alpha\mu(T_{\ge\alpha}) &\le \bar{w}(t,T,z) \\ \mu(T_{>\alpha}) &\le \alpha^{-1}\bar{w}(t,T,z) \end{aligned} \tag{V.60}$$

Overall, we obtained that for every $t, T, \alpha > 0$, $z \in \mathbb{R}^d$ may partition the interval $[t, t+T]$ such that

$$\begin{aligned} |w(\tau)^\top z| &\le \alpha, \quad \forall \tau \in T_{\le\alpha}(t,T,z), \\ |w(\tau)^\top z| &> \alpha, \quad \forall \tau \in T_{>\alpha}(t,T,z), \end{aligned} \tag{V.61}$$

with $\mu(T_{>\alpha}(t,T,z)) \le \alpha^{-1}\bar{w}(t,T,z)$. We want now to find an upper bound on $|w(\tau)^\top z|$ in the region $T_{>\alpha}(t,T,z)$.

Since $T_{>\alpha}(t,T,z)$ is a union of open sets of measure $\mu(T_{>\alpha}) \le \alpha^{-1}\bar{w}(t,T,z)$, we consider the case in which it is a unique interval (namely, the case in which $|w(\tau)|$ can grow more). Denote as $\bar{t} \in [t, t+T]$ the last time instant for which $|w(\bar{t})^\top z| \le \alpha$, and $M \coloneqq \|\mathrm{d}(\mathrm{w})\|_\infty$. Then, we express any $\tau \in T_{>\alpha}(t,T,z)$ as $\tau = \bar{t} + \delta$, with $\delta \leq \alpha^{-1}\bar{w}(t,T,z)$ (since we have shown that $\mu(T_{>\alpha}) \leq \alpha^{-1}\bar{w}(t,T,z)$). It holds that

$$\begin{aligned} \left|w(\bar{t}+\delta)^\top z\right| &= \left|w(\bar{t})^\top z + \int_t^{t+\delta} \dot{w}(\tau)^\top z \mathrm{d}\tau\right| \\ |w(\bar{t}+\delta)^\top z| &\leq \alpha + M\delta \\ |w(\bar{t}+\delta)^\top z| &\leq \alpha + M\alpha^{-1}\bar{w}(t,T,z). \end{aligned} \tag{V.62}$$

Choose any $\alpha > 0$. Pick $\epsilon(\alpha) := \frac{\alpha^2}{M}$. Then, for all $T > 0$, since $\mathrm{w} \notin \Omega_d^{\mathrm{C}}$, there exists $t > 0, z \in \mathbb{R}^d$ such that

$$\bar{w}(t,T,z) = \int_t^{t+T} |z^\top w(\tau)| \mathrm{d}\tau \leq \epsilon. \tag{V.63}$$

Substituting in (V.62), we have

$$\begin{aligned} |w(\tau)^\top z| &\leq \alpha + M\alpha^{-1}\epsilon \\ &\leq \alpha + \alpha, \end{aligned} \tag{V.64}$$

for all $\tau \in T_{>\alpha}(t,T,z)$, and thus for all $\tau \in [t, t+T]$.
To recap, we have proved that for any $\alpha, T > 0$, there exists $\epsilon > 0, z \in \mathbb{R}^d$ and thus $t > 0$ for which $|w(\tau)^\top z| \leq 2\alpha$ for all $\tau \in [t, t+T]$, and this concludes the proof.

※

V.4.4 Proof of Lemma II.6

We prove the theorem in three steps.

I) A useful characterization for PPE signals. Since $\mathrm{W}(\mathrm{w}) := \mathrm{D}^n(\mathrm{w})$ is PPE of degree at most $d' \leq d(n-1)$, by Lemma II.5, and following the same steps as in Lemma II.4 I), for all $T, \epsilon > 0$ we can find $t > 0$ such that

$$W(\tau) = E\lambda(\tau) + \tilde{W}(\tau), \tag{V.65}$$

where $E \in \mathbb{R}^{nd \times d'}$ stacks the directions in which $\mathrm{D}^n(\mathrm{w})$ is PPE, $\lambda(\tau) \in \mathbb{R}^{d'}$ stacks the projections of $\mathrm{D}^n(\mathrm{w})$ along these directions, and $|\tilde{W}(\tau)| \leq \epsilon$ for all $\tau \in [t, t+T]$.

II) Sampled signal sequencies are constrained by PPE. Pick an arbitrary $\bar{T} \leq T$ and $N \in \mathbb{N}$. We define $\delta := \bar{T}/N$. Notice that, being $\mathrm{d}(\mathrm{W})$ Lipschitz continuous, it holds that

$$W(\tau+\delta) = W(\tau) + \delta\dot{W}(\tau) + r(\delta), \tag{V.66}$$

where $|r(\delta)| \leq c\frac{\delta^2}{2}$, $c := \|\mathrm{d}^2(\mathrm{W})\|_\infty$.

By recalling that $W(\tau) = (w(\tau), \dot{w}(\tau), \ldots, \dot{w}^{(n-1)}(\tau))$, we can rewrite (V.66) as

$$W(\tau+\delta) = (I+\delta S)W(\tau) + \delta e_d \dot{w}^{(n)}(\tau) + r(\delta), \tag{V.67}$$

where $S \in \mathbb{R}^{nd\times nd}$, $e_d \in \mathbb{R}^{nd\times d}$ are given by

$$S = \begin{bmatrix} 0_{(n-1)d\times d} & I_{(n-1)d} \\ 0_{d\times d} & 0_{d\times(n-1)d} \end{bmatrix}, \quad e_d = \begin{bmatrix} 0_{d\times d} & 0_{d\times d} & \ldots & I_d \end{bmatrix}^\top. \tag{V.68}$$

By substituting (V.65) into (V.67), we obtain that in the interval $\tau \in [t, t+T-\delta]$ we can write

$$\begin{aligned} &E\lambda(\tau+\delta) + \tilde{W}(\tau+\delta) = \\ &= (I_{nd}+\delta S)(E\lambda(\tau) + \tilde{W}(\tau)) + \delta e_d \dot{w}^{(n)}(\tau) + r(\delta). \end{aligned} \tag{V.69}$$

Considering N consecutive instants $\tau, \tau+\delta, \ldots, \tau+\bar{T}-\delta$, similarly to (V.51) we obtain

$$M\hat{w}(\tau) = \tilde{E}\hat{\lambda}(\tau) + \tilde{w}(\tau), \tag{V.70}$$

where $\hat{w}(\tau) \in \mathbb{R}^{Nd}$, $\hat{\lambda}(\tau) \in \mathbb{R}^{d'(N+1)}$, $\tilde{w}(\tau) \in \mathbb{R}^{ndN}$ are given by

$$\begin{aligned} \hat{w}(\tau) &:= \begin{bmatrix} \dot{w}^{(n)}(\tau) \\ \dot{w}^{(n)}(\tau+\delta) \\ \vdots \\ \dot{w}^{(n)}(\tau+\bar{T}-\delta) \end{bmatrix}, \quad \hat{\lambda}(\tau) := \begin{bmatrix} \lambda(\tau) \\ \lambda(\tau+\delta) \\ \vdots \\ \lambda(\tau+\bar{T}) \end{bmatrix}, \\ \tilde{w}(\tau) &:= \begin{bmatrix} -(I_{nd}+\delta S)\tilde{W}(\tau) + \tilde{W}(\tau+\delta) + r(\delta) \\ -(I_{nd}+\delta S)\tilde{W}(\tau+\delta) + \tilde{W}(\tau+2\delta) + r(\delta) \\ \vdots \\ -(I_{nd}+\delta S)\tilde{W}(\tau+\bar{T}-\delta) + \tilde{W}(\tau+\bar{T}) + r(\delta) \end{bmatrix}, \end{aligned} \tag{V.71}$$

and

$$\begin{aligned} M &:= \delta I_N \otimes e_d \in \mathbb{R}^{ndN\times Nd} \\ \tilde{E} &:= \begin{bmatrix} [-(I_{nd}+\delta S)E \;\; E] & 0 & \ldots & 0 \\ 0 & [-(I_{nd}+\delta S)E \;\; E] & 0 & \ldots \\ & \vdots & & \\ 0 & \ldots & 0 & [-(I_{nd}+\delta S)E \;\; E] \end{bmatrix} \in \mathbb{R}^{ndN\times d'(N+1)}. \end{aligned} \tag{V.72}$$

We are interested in characterizing the solutions $(\hat{w}(\tau), \hat{\lambda}(\tau))$ of system (V.70), namely, the possible signals which fulfill the constraints which are imposed by PPE. In other words, we want to solve

$$\begin{bmatrix} M & -\tilde{E} \end{bmatrix} \begin{bmatrix} \hat{w}(\tau) \\ \hat{\lambda}(\tau) \end{bmatrix} = \tilde{w}(\tau) \tag{V.73}$$

in the interval $[t, t+T]$.

III) A small enough degree of PPE overconstrains the sampled signal sequence. Given the right pseudo-inverse $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger$ of $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}$, any solution $(\hat{w}(\tau), \hat{\lambda}(\tau))$ can be written as

$$\begin{bmatrix} \hat{w}(\tau) \\ \hat{\lambda}(\tau) \end{bmatrix} = \begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger \tilde{w}(\tau) + v(\tau), \tag{V.74}$$

where $v(\tau) \in \ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)$. Consider the term $\tilde{w}(\tau)$: we show at first it can be made arbitrarily small. By recalling the bounds on $r(\delta)$ (see (V.66)) and $\tilde{W}(\tau)$ (see (V.65)), it holds that

$$\begin{aligned} |\tilde{w}(\tau)| &\leq \left\| \begin{bmatrix} (I_{nd} + \delta S)\tilde{W}(\tau) \\ (I_{nd} + \delta S)\tilde{W}(\tau + \delta) \\ \vdots \\ (I_{nd} + \delta S)\tilde{W}(\tau + \bar{T} - \delta) \end{bmatrix} \right\| + \left\| \begin{bmatrix} \tilde{W}(\tau+\delta) \\ \tilde{W}(\tau+2\delta) \\ \vdots \\ \tilde{W}(\tau+\bar{T}) \end{bmatrix} \right\| + \left\| \begin{bmatrix} r(\delta) \\ r(\delta) \\ \vdots \\ r(\delta) \end{bmatrix} \right\| \\ &\leq (1+\delta)\sqrt{N}\epsilon + \sqrt{N}\epsilon + \sqrt{N}c\frac{\delta^2}{2} \\ &\leq \left(2\sqrt{N} + \frac{\bar{T}}{\sqrt{N}}\right)\epsilon + \sqrt{N}c\frac{\bar{T}^2}{2N^2}. \end{aligned} \tag{V.75}$$

We now show we can always find an arbitrarily long interval in which $\tilde{w}(\tau)$ is arbitrarily small. Notice that, given any $\epsilon' > 0$, by choosing

$$N \geq \bar{N} := \sqrt[3]{\frac{\bar{T}^4 c^2}{\epsilon'^2}} \implies \sqrt{N}c\frac{\bar{T}^2}{2N^2} \leq \frac{\epsilon'}{2}. \tag{V.76}$$

Exploting this, and choosing $\epsilon = \frac{\epsilon'}{2(2\sqrt{N} + \bar{T}/\sqrt{N})}$, we have from (V.75) that for any $\epsilon' > 0, \bar{T} > 0, T > \bar{T}$ there exists $\bar{N} \in \mathbb{N}, t > 0$ such that, for all $N \geq \bar{N}$,

$$|\tilde{w}(\tau)| \leq \epsilon' \quad \forall \tau \in [t, t+T]. \tag{V.77}$$

We move now to the term $v(\tau) \in \ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)$. Notice that, since $\begin{bmatrix} M & -\tilde{E} \end{bmatrix}$ has full row rank, by the rank-nullity theorem it holds

$$\begin{aligned} d'' &:= \dim\left(\ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) \\ &= \dim\left(\mathrm{dom}\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) - \dim\left(\mathrm{im}\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right)\right) \\ &= Nd + d'(N+1) - ndN \\ &= N(d + d' - nd) + d', \end{aligned} \tag{V.78}$$

from which it is clear that if $d' \leq d(n-1)$ then $d'' \leq d'$. This means that we can write the solution $\hat{\lambda}(\tau)$ of (V.74) as

$$\hat{\lambda}(\tau) = \begin{bmatrix} 0 & I_N \end{bmatrix} G v(\tau) + \begin{bmatrix} 0 & I_N \end{bmatrix} \begin{bmatrix} M & -\tilde{E} \end{bmatrix}^\dagger \tilde{w}(\tau), \tag{V.79}$$

with G any surjective map

$$G : \mathbb{R}^{d''} \to \ker\left(\begin{bmatrix} M & -\tilde{E} \end{bmatrix}\right), \tag{V.80}$$

and $v(\tau) \in \mathbb{R}^{d''}$ such that $v(\tau) = Gv(\tau)$. Using (V.65) and (V.79), and recalling the definition of $\hat{\lambda}(\tau)$ in (V.71), we can reconstruct the vector in which we are interested in:

$$\begin{bmatrix} W(\tau) \\ W(\tau + \frac{\bar{T}}{N}) \\ \vdots \\ W(\tau + \bar{T}) \end{bmatrix} = G' v(\tau) + F\tilde{w}(\tau) + \begin{bmatrix} \tilde{W}(\tau) \\ \tilde{W}(\tau + \frac{\bar{T}}{N}) \\ \vdots \\ \tilde{W}(\tau + \bar{T}) \end{bmatrix}, \tag{V.81}$$

where

$$\begin{aligned} G' &= (I_N \otimes E)\begin{bmatrix} 0 & I_N \end{bmatrix} G \\ F &= (I_N \otimes E)\begin{bmatrix} 0 & I_N \end{bmatrix}\begin{bmatrix} M & -E \end{bmatrix}^{\dagger}. \end{aligned} \tag{V.82}$$

The proof is complete by recalling that the choices of N, ϵ are such that $|\tilde{w}(\tau)| \leq \epsilon'$ (in (V.77)) and $|(\tilde{W}(\tau), \dots, \tilde{W}(\tau + \bar{T}))| \leq \epsilon'/2$ for all $\tau \in [t, t+T]$.

※

V.4.5 Proof of Lemma II.7

By Lemma II.5, if $\mathrm{w} \notin \Omega_d^{\mathrm{c}}$ then for all $T, \epsilon > 0$ we can find $t > 0, z \in \mathbb{R}^d$ such that

$$|w(\tau)^\top z| \leq \epsilon, \quad \forall \tau \in [t, t+T]. \tag{V.83}$$

Picking any $\tau, \delta > 0$ such that $\tau, \tau + \delta \in [t, t+T]$, we have that

$$|(w(\tau+\delta) - w(\tau))^\top z| \leq |w(\tau+\delta)^\top z| + |w(\tau)^\top z| \leq 2\epsilon \tag{V.84}$$

Expanding $w(\tau + \delta)$ in Taylor series, we obtain that

$$w(\tau+\delta) - w(\tau) = \dot{w}(\tau)\delta + o(\delta), \tag{V.85}$$

where $|o(\delta)| \leq M\frac{\delta^2}{2}$, with $M := \|\mathrm{d}^2(\mathrm{w})\|_\infty$. Using (V.85) in (V.84), we obtain

$$\begin{aligned} |(w(\tau+\delta) - w(\tau))^\top z| &\leq 2\epsilon \\ |(\dot{w}(\tau)\delta + o(\delta))^\top z| &\leq 2\epsilon. \end{aligned} \tag{V.86}$$

Since, by the triangle inequality, $|(\dot{w}(\tau)\delta + o(\delta))^\top z| \geq |\dot{w}(\tau)^\top z\delta| - |o(\delta)^\top z|$, we have that

$$\begin{aligned} |\delta\dot{w}(\tau)^\top z| - |o(\delta)^\top z| &\leq 2\epsilon \\ |\delta\dot{w}(\tau)^\top z| &\leq 2\epsilon + |o(\delta)| \\ |\dot{w}(\tau)^\top z| &\leq 2\frac{\epsilon}{\delta} + M\frac{\delta}{2}. \end{aligned} \tag{V.87}$$

Choosing a sufficiently small $\epsilon > 0$ and $\delta(\epsilon) := \sqrt{\epsilon}$, for all $\tau \in [t, t+T-\delta(\epsilon)]$ we have

$$|\dot{w}(\tau)^\top z| \leq \left(2 + \frac{M}{2}\right)\sqrt{\epsilon}. \tag{V.88}$$

By defining $\gamma(\epsilon) := \max(\sqrt{\epsilon}, (2+\frac{M}{2})\sqrt{\epsilon})$, we obtain

$$|w(\tau)^\top z| \leq \gamma(\epsilon) \qquad\qquad |\dot{w}(\tau)^\top z| \leq \gamma(\epsilon) \tag{V.89}$$

for all $\tau \in [t, t+T-\delta(\epsilon)]$.
Since $\gamma(\epsilon)$ is a strictly increasing function of ϵ such that $\gamma(0) = 0$, we may pick an arbitrarily small ϵ' and find $\epsilon : \epsilon' = \gamma(\epsilon)$; consequently, since t exists for any choice of $\epsilon, T > 0$, for any choice of ϵ' there exists an interval for which (V.89) holds.

※

V.4.6 Proof of Theorem II.6

We prove that $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{nm} \implies \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_n$ for $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{x}}$ and for all $x_0 \in \mathbb{R}^n$ by contraposition, i.e., we show that for all $x_0 \in \mathbb{R}^n$, $\sigma(\mathrm{u}, x_0) \notin \Omega^{\mathrm{D}}_n \implies \mathrm{Q}^n(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{nm}$. We prove this in four points.

I) The lack of PE of x **constrains the system input.** If $\mathrm{x} := \sigma(\mathrm{u}, x_0) \notin \Omega^{\mathrm{D}}_n$, applying Definition II.13 we have that for all $T, \epsilon > 0$ we can find a direction $z \in \mathbb{R}^n$ and $t \in \mathbb{N}$ such that

$$|z^\top x_\tau| \leq \epsilon \tag{V.90}$$

for all $\tau = t, \dots, t+T$. Along direction z, system dynamics read

$$z^\top x_{\tau+1} = z^\top A x_\tau + z^\top B u_\tau. \tag{V.91}$$

If $z^\top B \neq 0$, we can find the input u_τ as

$$u_\tau = \underbrace{-(z^\top B)^\dagger z^\top A}_{:=K} x_\tau + \underbrace{(z^\top B)^\dagger z^\top x_{\tau+1}}_{:=\tilde{u}_\tau} + v_\tau, \tag{V.92}$$

where $|\tilde{u}_\tau| \leq |(z^\top B)^\dagger|\epsilon$, and $v_\tau \in \ker(z^\top B)$.

II) The closed-loop dynamics depend only on x_τ, v_τ. The dynamics become

$$\begin{aligned} u_\tau &= Kx_\tau + \tilde{u}_\tau + v_\tau, \\ x_{\tau+1} &= \underbrace{(A - BK)}_{:=\tilde{A}} x_\tau + B\tilde{u}_\tau + Bv_\tau \end{aligned} \tag{V.93}$$

where, for all $\tau = t, \dots, t+T-1$, it holds that $|B\tilde{u}_\tau| \leq \epsilon$ and $v_\tau \in \ker(z^\top B)$. Consider the signal $\mathrm{U}(\mathrm{u}) := \mathrm{Q}^n(\mathrm{u})$. Given (V.92), we obtain

$$U_\tau = (I_n \otimes K) \begin{bmatrix} x_{\tau-n+1} \\ \vdots \\ x_\tau \end{bmatrix} + \underbrace{\begin{bmatrix} \tilde{u}_{\tau-n+1} \\ \vdots \\ \tilde{u}_\tau \end{bmatrix}}_{:=\tilde{U}_\tau} + \underbrace{\begin{bmatrix} v_{\tau-n+1} \\ \vdots \\ v_\tau \end{bmatrix}}_{:=V_\tau}. \tag{V.94}$$

Using (V.93), it holds that

$$\begin{bmatrix} x_{\tau-n+1} \\ \vdots \\ x_\tau \end{bmatrix} = Fx_{\tau-n+1} + G\left(\tilde{U}_\tau + V_\tau\right), \tag{V.95}$$

where

$$F = \begin{bmatrix} I \\ \vdots \\ \tilde{A}^{n-1} \end{bmatrix}, \quad G = \begin{bmatrix} 0 & \dots & \dots & 0 \\ B & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ A^{n-2}B & \dots & B & 0 \end{bmatrix}, \tag{V.96}$$

and substituting (V.95) in (V.94), we obtain

$$\begin{aligned} U_\tau &= \underbrace{(I_n \otimes K)F}_{:=F'} x_{\tau-n+1} + \underbrace{((I_n \otimes K)G + I)}_{G'} \left(\tilde{U}_\tau + V_\tau\right), \\ &= \begin{bmatrix} F' & G' \end{bmatrix} \begin{bmatrix} x_{\tau-n+1} \\ V_\tau \end{bmatrix} + G'\tilde{U}_\tau. \end{aligned} \tag{V.97}$$

III) Lack of PE in x implies $\mathrm{Q}^n(\mathrm{u})$ is not PE. Notice that in the period $\tau = t+n-1, \dots, t+T$, it holds that

i) $v_\tau \in \ker(z^\top B)$, with $\dim(\ker(z^\top B)) \leq m-1$, by construction of v_τ.

ii) $|z^\top x_\tau| \leq \epsilon$ by lack of PE of x.

iii) $|\tilde{U}_\tau| \leq \sqrt{n}|(z^\top B)^\dagger|\epsilon$ from (V.92).

Holding i) and ii), the space persistently spanned by (x_τ, V_τ) is at most $(n-1) + n(m-1) = nm - 1$ dimensional. In other words, by following the same procedure as in Lemma II.4, we can write for all

$\tau = t+n-1, t+T$

$$\begin{bmatrix} x_{\tau-n+1} \\ V_\tau \end{bmatrix} = E\lambda_\tau + \tilde{\lambda}_\tau, \tag{V.98}$$

where $E \in \mathbb{R}^{(nm)\times(nm-1)}$ stacks the directions which are persistently spanned by (x_τ, V_τ), $\lambda_\tau \in \mathbb{R}^{nm-1}$ stacks the projections of (x_τ, V_τ) along these directions, and $\tilde{\lambda}_\tau \in \mathbb{R}^{nm}$, $|\tilde{\lambda}_\tau| \le \epsilon$ is an arbitrarily small perturbation. Finally, we have

$$U_\tau = \underbrace{\begin{bmatrix} F' & G' \end{bmatrix} E}_{:=H} \lambda_\tau + \begin{bmatrix} F' & G' \end{bmatrix} \tilde{\lambda}_\tau + G'\tilde{U}_\tau, \tag{V.99}$$

and since $\begin{bmatrix} F' & G' \end{bmatrix} E \in \mathbb{R}^{nm\times(nm-1)}$, there exists $z \in \mathbb{R}^{nm}$ such that $z^\top \begin{bmatrix} F' & G' \end{bmatrix} E = 0$, which implies

$$z^\top U_\tau = z^\top \begin{bmatrix} F' & G' \end{bmatrix} \tilde{\lambda}_\tau + z^\top G'\tilde{U}_\tau. \tag{V.100}$$

Being both $\tilde{U}_\tau$ and $\tilde{\lambda}$ arbitrarily small in the arbitrarily long interval $\tau = t+n-1, \dots, t+T$, we conclude $\mathrm{Q}^n(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{nm}$.

IV) The case of $y^\top B = 0$. Consider the case where $z^\top B = 0$, namely, the columns $b_1, \dots, b_m$ of B satisfy $b_1, \dots, b_m \in \ker(z^\top)$. In that case, along direction z, the system dynamics read as

$$\begin{aligned} z^\top x_{\tau+1} &= z^\top A x_\tau + z^\top B u_\tau \\ &= z^\top A x_\tau. \end{aligned} \tag{V.101}$$

We can distinguish two cases: either $\ker(z^\top) = \ker(z^\top A)$ or $\ker(z^\top) \neq \ker(z^\top A)$. If $\ker(z^\top) = \ker(z^\top A)$, then $\ker(z^\top)$ must be an invariant subspace of A of dimension $n-1$. Since $\ker(z^\top)$ is A-invariant, and since $b_1, \dots, b_m \in \ker(z^\top)$, there are at most $n-1$ linearly independent vectors between the columns of $B, \dots, A^{n-1}B$. This is a contradiction, since we assumed (A, B) controllable, thus, it cannot hold $\ker(z^\top) = \ker(z^\top A)$.
At last, consider the case $\ker(z^\top) \neq \ker(z^\top A)$. Since (V.101) holds for all $\tau = t+1, \dots, t+T-1$, we can write

$$\begin{aligned} z^\top x_\tau &\le \epsilon \\ z^\top x_{\tau+1} = z^\top A x_\tau &\le \epsilon, \end{aligned} \tag{V.102}$$

from which we deduce that we can write $x_\tau = \bar{x}_\tau + \tilde{x}_\tau$, with $\bar{x}_\tau \in \ker(z^\top) \cap \ker(z^\top A)$ and $|\tilde{x}_\tau| \le \epsilon$. Since $\ker(z^\top) \neq \ker(z^\top A)$ and since the dimension of each kernel is at most $n-1$, then $\dim(\ker(z^\top) \cap \ker(z^\top A)) \le n-2$.
This means that $\bar{x}_\tau$ spans persistently at most $n-2$ directions, namely, there exists an unitary $z_2 \in$

$\mathbb{R}^n$, $z_2 \perp z \perp (\ker(z^\top) \cap \ker(z^\top A))$, such that

$$|y_2^\top x_\tau| = |y_2^\top \bar{x}_\tau + y_2^\top \tilde{x}_\tau| = |y_2^\top \tilde{x}_\tau| \le \epsilon. \tag{V.103}$$

We can thus repeat the same procedure as before, checking if $z_2^\top B \neq 0$. If $z_2^\top B \neq 0$, we can repeat the reasoning in points I), II), III); otherwise, we can repeat the above reasoning to find another direction z_3 such that $z_3^\top x_\tau \le \epsilon$ (and repeat the process until we find some $z_i^\top B \neq 0$, which must exist since $B \neq 0$. In each of these cases, $\mathrm{Q}^n(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{nm}$, which concludes the proof for the first statement of the theorem.

We move to the second statement of the theorem. Since it has an analogous proof, we sketch only the main differences with respect to the proof given before. We want to prove by contraposition that $\mathrm{Q}^n(\mathrm{u}) \in \Omega^{\mathrm{D}}_{(n+1)m} \implies \sigma(\mathrm{u}, x_0) \in \Omega^{\mathrm{D}}_{n+m}$ for $\sigma \in \mathbb{L}^{\mathrm{D}}_{\mathrm{xu}}$ and for all $x_0 \in \mathbb{R}^n$ by showing that for all $x_0 \in \mathbb{R}^n$, $\sigma(\mathrm{u}, x_0) \notin \Omega^{\mathrm{D}}_{n+m} \implies \mathrm{Q}^{n+1}(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{(n+1)m}$.

I) The lack of PE of (x, u) constrains the system input. If $(\mathrm{x}, \mathrm{u}) \notin \Omega^{\mathrm{D}}_{n+m}$ we have that for all $T, \epsilon > 0$ we can find $z = (z_x, z_u) \in \mathbb{R}^{n+m}$, $t \in \mathbb{N}$ such that

$$z_x^\top x_\tau = -z_u^\top u_\tau + \chi_\tau, \tag{V.104}$$

where $\chi_\tau \in \mathbb{R}$, $|\chi_\tau| \le \epsilon$ for all $\tau = t, \dots, t+T$. By pre-multiplying by $z_x^\top$ the systems dynamics, we obtain the update

$$u_{\tau+1} = -(z_u^\top)^\dagger z_x^\top (A x_\tau + B u_\tau) + (z_u^\top)^\dagger \chi_{\tau+1} + v_{\tau+1}, \tag{V.105}$$

where $v_{\tau+1} \in \ker(z_u^\top)$.

II) The closed-loop dynamics depends only on x_τ, v_τ, u_τ. Since we can write $x_{\tau+1} = A x_\tau + B u_\tau$, in the interval $\tau = t, \dots, t+T$ we can use (V.105) to express each $u_\tau, u_{\tau+1}, \dots, u_{\tau+n}$ as a linear function of $x_\tau, u_\tau, v_{\tau+1}, \dots, v_{\tau+n}$ (similarly as done in (V.94)) plus an arbitrarily small quantity $\tilde{U}_\tau$. In other words, we have

$$\begin{bmatrix} u_\tau \\ \vdots \\ u_{\tau+n} \end{bmatrix} = K \begin{bmatrix} u_\tau \\ x_\tau \\ v_{\tau+1} \\ \vdots \\ v_{\tau+n} \end{bmatrix} + \tilde{U}_\tau. \tag{V.106}$$

III) A lack of PE in (x, u) implies $\mathrm{Q}^{n+1}(\mathrm{u})$ is not PE. Since (x, u) is not PE, it spans at most $n+m-1$ directions. Since $v_\tau \in \ker(z_u^\top)$, the vector $(v_{\tau+1}, \dots, v_{\tau+n})$ spans at most $(m-1)n$ directions. Overall, we have that the right-hand side of (V.106) spans persistently only $n+m-1+(m-1)n = (n+1)m-1$ directions, which means that the signal $(u_\tau, \dots, u_{\tau+n}) \in \mathbb{R}^{(n+1)m}$ on the left-hand side of (V.106) spans persistently only $(n+1)m-1$ directions, namely, $\mathrm{Q}^{n+1}(\mathrm{u}) \notin \Omega^{\mathrm{D}}_{(n+1)m}$.

※

V.4.7 Proof of Theorem II.4

We prove this result by contraposition, i.e., we show that given $\sigma \in \mathbb{L}_{\mathrm{x}}^{\mathrm{D}}$, if $\mathrm{Q}^n(\mathrm{u}) \in \ell_\infty(\mathbb{R}^{nm})$ is PPE of degree at most $n' \leq n-1$, then for all $x_0 \in \mathbb{R}^n$, $\mathrm{x} = \sigma(\mathrm{u}, x_0) \notin \Omega_n^{\mathrm{D}}$ regardless of the initial condition.

I) x_t can be approximated arbitrarily well by a linear function of older inputs. Let $x_0 = 0$. We can write system dynamics as

$$x_{\tau+1} = A^n x_{\tau-n+1} + RU_\tau, \tag{V.107}$$

where R is the reachability matrix and $U_\tau = (u_\tau, \ldots, u_{\tau-n+1}) \in \mathbb{R}^{nm}$. By writing $x_{\tau-n+1}$ as a function of the previous inputs, and repeating this recursion for arbitrary K steps, we obtain

$$\begin{aligned} x_{\tau+1} =& A^{Kn} x_{\tau-Kn+1} \\ &+ \underbrace{\begin{bmatrix} I & A^n & \ldots & A^{Kn} \end{bmatrix}}_{=:\bar{A}} \underbrace{(I_K \otimes R)}_{=:\bar{R}} \begin{bmatrix} U_\tau \\ U_{\tau-n} \\ \vdots \\ U_{\tau-(K-1)n} \end{bmatrix}. \end{aligned} \tag{V.108}$$

Notice that for any $\epsilon > 0$, we can choose $K > 0$ such that $|A^{Kn} x_{\tau-Kn+1}| \leq \epsilon$ for all $\tau \in \mathbb{N}$, being the signal x bounded. Notice the vector $(U_\tau, \ldots, U_{\tau-(K-1)n})$ is given by the signal $\mathrm{Q}^{(K-1)n}(\mathrm{u})$ evaluated at time τ.

II) The degree of PPE of $\mathrm{Q}^{(K-1)n}(\mathrm{u})$ is limited by the degree of PPE of $\mathrm{Q}^n(\mathrm{u})$. If $\mathrm{Q}^n(\mathrm{u})$ is PPE of degree at most $n' \leq n-1$, then $n' \leq m(n-1)$ and we can apply Lemma II.4 to ensure that for any $K \geq 1$, $\mathrm{Q}^{(K-1)n}(\mathrm{u})$ is PPE of degree at most $n' \leq n-1$.

III) If the spanned directions are $n-1$, $\mathrm{x} \notin \Omega_n^{\mathrm{C}}$ Leveraging the lack of PPE demonstrated in the above point, for all $T, K, \epsilon > 0$ there exists $t > 0$ for which we can rewrite (V.108) as (see the derivation in Lemma II.4)

$$x_\tau = A^{Kn} x_{\tau-Kn} + \bar{A}\bar{R}E\lambda_t + \bar{A}\bar{R}\tilde{\lambda}_t, \tag{V.109}$$

for all $\tau = t, \ldots, t+T$, for some $E \in \mathbb{R}^{(K-1)nm \times (n-1)}$ stacking the directions in which there is PPE, $\lambda_t \in \mathbb{R}^{n-1}$ stacking the projections of $(U_t, \ldots, U_{t-(K-1)n})$ along these directions, $\tilde{\lambda}_t \in \mathbb{R}^{(K-1)nm}$ such that $|\tilde{\lambda}_{t_{\epsilon,K}}| \leq \epsilon$.

Since $\bar{A}\bar{R}E \in \mathbb{R}^{n \times (n-1)}$, there exists $z \in \mathbb{R}^n$ such that $z^\top \bar{A}\bar{R}E = 0$, which implies that

$$z^\top x_\tau = z^\top A^{Kn} x_{\tau-Kn} + z^\top \bar{A}\bar{R}\tilde{\lambda}_t. \tag{V.110}$$

Being both $|A^{Kn}x_{\tau-Kn}|$ and $\tilde{\lambda}$ arbitrarily small in the arbitrarily long interval $\tau = t+n-1, t+T$, we conclude $\mathrm{x} = \sigma(\mathrm{u}, 0) \notin \Omega_n^{\mathrm{D}}$ and thus $\mathrm{u} \notin C_{\mathrm{SR}}(\sigma, 0)$. To conclude the proof, since σ is a stable linear system, we have that all solutions $\mathrm{x} = \sigma(\mathrm{u}, x_0)$ converge exponentially to those initialized in $x_0 = 0$. Since a vanishing term cannot guarantee PE of a signal, we can conclude that for all $x_0 \in \mathbb{R}^n$, if $\mathrm{Q}^n(\mathrm{u})$ is PPE of degree at most $n' \le n-1$, then $\mathrm{x} = \sigma(\mathrm{u}, x_0) \notin \Omega_n^{\mathrm{D}}$.

We pass now to the second statement of the theorem. Since it has an analogous proof, we sketch only the main differences with respect to the proof given before. We want to prove that, given $\sigma \in \mathbb{L}_{\mathrm{xu}}^{\mathrm{D}}$, if $\mathrm{Q}^{n+1}(\mathrm{u}) \in \ell_\infty(\mathbb{R}^{nm})$ is PPE of degree $n' \le n+m-1$, then for all $x_0 \in \mathbb{R}^n$, $\mathrm{x} = \sigma(\mathrm{u}, x_0) \notin \Omega_n^{\mathrm{D}}$ regardless of the intitial condition.

I)(x_t, u_t) can be approximated arbitrarily well by a linear function of the previous inputs. Similarly as done in (V.108), we obtain

$$\begin{bmatrix} u_{\tau+1} \\ x_{\tau+1} \end{bmatrix} = \begin{bmatrix} 0 \\ A^{Kn}x_{\tau-Kn+1} \end{bmatrix} + \begin{bmatrix} I_m & 0 \\ 0 & \bar{A}\bar{R} \end{bmatrix} \begin{bmatrix} u_{\tau+1} \\ U_\tau \\ \vdots \\ U_{\tau-(K-1)n} \end{bmatrix}, \tag{V.111}$$

where $\bar{A}, \bar{R}$ are the same as in (V.111), and by choosing an appropriate $K \in \mathbb{N}$, for any $\epsilon > 0$ we achieve $|A^{Kn}x_{\tau-Kn+1}| \le \epsilon$, since A is Shur and x is bounded. Notice that the vector $(u_{\tau+1}, U_\tau, \ldots, U_{\tau-(K-1)n})$ is given by the signal $\mathrm{Q}^{(K-1)n+1}(\mathrm{u})$ evaluated at time $\tau+1$.

II) The degree of PPE of $\mathrm{Q}^{(K-1)n+1}(\mathrm{u})$ is limited by the degree of PPE of $\mathrm{Q}^{n+1}(\mathrm{u})$. Applying Lemma II.4, if $\mathrm{Q}^{n+1}(\mathrm{u})$ is PPE of degree at most $n' \le n+m-1$, then, since $n+m-1 \le m(n+1-1) = nm$ for all $n, m \in \mathbb{N}$, $\mathrm{Q}^{(K-1)n+1}(\mathrm{u})$ is PPE of degree at most n'.

III) If the spanned directions are $n+m-1$, then $(\mathrm{x}, \mathrm{u}) \notin \Omega_{n+m}^{\mathrm{D}}$ This point proceeds exactly as for the previous case, so we omit it.

※

V.4.8 Proof of Theorem II.7

We prove that $\mathrm{D}^n(\mathrm{u}) \in \Omega_{nm}^{\mathrm{C}} \implies \sigma(\mathrm{u}, x(0)) \in \Omega_n^{\mathrm{C}}$ for $\sigma \in \mathbb{L}_{\mathrm{x}}^{\mathrm{C}}$ and for all $x(0) \in \mathbb{R}^n$ by contraposition, i.e., we show that for all $x(0) \in \mathbb{R}^n$, $\sigma(\mathrm{u}, x(0)) \notin \Omega_n^{\mathrm{C}} \implies \mathrm{D}^n(\mathrm{u}) \notin \Omega_{nm}^{\mathrm{C}}$. We prove this in four points.

I) The lack of PE of x constrains the system input. If $\mathrm{x} := \sigma(\mathrm{u}, x(0)) \notin \Omega_n^{\mathrm{C}}$, by applying Lemma II.7 we have that for all $T, \epsilon > 0$ we can find a unitary direction $z \in \mathbb{R}^n$ and $t > 0$ such that

$$|z^\top \dot{x}(\tau)| \le \epsilon \qquad |z^\top x(\tau)| \le \epsilon \tag{V.112}$$

for all $\tau \in [t, t+T]$. Along direction z, the system dynamics read

$$z^\top \dot{x}(\tau) = z^\top A x(\tau) + z^\top B u(\tau). \tag{V.113}$$

If $z^\top B \neq 0$, we can find the input $u(\tau)$ as

$$u(\tau) = \underbrace{-(z^\top B)^\dagger z^\top A}_{:=K} x(\tau) + \underbrace{(z^\top B)^\dagger z^\top \dot{x}(\tau)}_{:=\tilde{u}(\tau)} + v(\tau), \tag{V.114}$$

where $|\tilde{u}(\tau)| \leq |(z^\top B)^\dagger|\epsilon$, and $v(\tau) \in \ker(z^\top B)$.

II) The closed-loop dynamics depends only on $x(\tau), v(\tau)$. The dynamics become

$$\begin{aligned} u(\tau) &= Kx(\tau) + \tilde{u}(\tau) + v(\tau), \\ \dot{x}(\tau) &= \underbrace{(A - BK)}_{:=\tilde{A}} x(\tau) + B\tilde{u}(\tau) + Bv(\tau) \end{aligned} \tag{V.115}$$

where, for all $\tau \in [t, t+T]$, it holds $|B\tilde{u}(\tau)| \leq \epsilon$ and $v(\tau) \in \ker(z^\top B)$. Consider the signal $\mathrm{U}(\mathrm{u}) := \mathrm{D}^n(\mathrm{u})$. Given (V.114), we obtain

$$U(\tau) = (I_n \otimes K) \begin{bmatrix} \dot{x}^{(n-1)}(\tau) \\ \vdots \\ x(\tau) \end{bmatrix} + \underbrace{\begin{bmatrix} \dot{\tilde{u}}^{(n-1)}(\tau) \\ \vdots \\ \tilde{u}(\tau) \end{bmatrix}}_{:=\tilde{U}(\tau)} + \underbrace{\begin{bmatrix} \dot{v}^{(n-1)}(\tau) \\ \vdots \\ v(\tau) \end{bmatrix}}_{:=V(\tau)}. \tag{V.116}$$

Using (V.115), it holds

$$\begin{bmatrix} \dot{x}^{(n-1)}(\tau) \\ \vdots \\ x(\tau) \end{bmatrix} = F\dot{x}^{(n-1)}(\tau) + G\left(\tilde{U}(\tau) + V(\tau)\right), \tag{V.117}$$

where

$$F = \begin{bmatrix} I \\ \vdots \\ \tilde{A}^{n-1} \end{bmatrix}, \quad G = \begin{bmatrix} 0 & \dots & \dots & 0 \\ B & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ A^{n-2}B & \dots & B & 0 \end{bmatrix}, \tag{V.118}$$

and substituting in (V.116), we obtain

$$\begin{aligned} U(\tau) &= \underbrace{(I_n \otimes K)F}_{:=F'} \dot{x}^{(n-1)}(\tau) + \underbrace{((I_n \otimes K)G + I)}_{G'} (\tilde{U}(\tau) + V(\tau)), \\ &= \begin{bmatrix} F' & G' \end{bmatrix} \begin{bmatrix} \dot{x}^{(n-1)}(\tau) \\ V(\tau) \end{bmatrix} + G'\tilde{U}(\tau). \end{aligned} \tag{V.119}$$

III) A lack of PE in x **implies** $D^n(u)$ **is not PE.** Notice that in the period $\tau \in [t, t+T]$, it holds

i) $v(\tau) \in \ker(z^\top B)$, with $\dim(\ker(z^\top B)) \leq m-1$, as per (V.114).

ii) $|z^\top x(\tau)| \leq \epsilon$ by assumption.

iii) $|\tilde{U}(\tau)| \leq \sqrt{n}|(z^\top B)^\dagger|\epsilon$, from (V.114).

Holding i) and ii), the space persistently spanned by $(x(\tau), V(\tau))$ is at most $(n-1)+n(m-1) = nm-1$ dimensional. In other words, by following the same procedure as in Lemma II.6, we can write for all $\tau \in [t, t+T]$

$$\begin{bmatrix} \dot{x}^{(n-1)}(\tau) \\ V(\tau) \end{bmatrix} = E\lambda(\tau) + \tilde{\lambda}(\tau), \tag{V.120}$$

where $E \in \mathbb{R}^{(nm)\times(nm-1)}$ stacks the directions which are persistently spanned by $(x(\tau), V(\tau))$, $\lambda(\tau) \in \mathbb{R}^{nm-1}$ stacks the projections of $(x(\tau), V(\tau))$ along these directions, and $\tilde{\lambda}(\tau) \in \mathbb{R}^{nm}$, $|\tilde{\lambda}(\tau)| \leq \epsilon$ is an arbitrarily small perturbation. We have

$$U(\tau) = \underbrace{\begin{bmatrix} F' & G' \end{bmatrix} E}_{:=H} \lambda(\tau) + \begin{bmatrix} F' & G' \end{bmatrix} \tilde{\lambda}(\tau) + G'\tilde{U}(\tau), \tag{V.121}$$

and since $\begin{bmatrix} F' & G' \end{bmatrix} E \in \mathbb{R}^{nm\times(nm-1)}$, there exists $z \in \mathbb{R}^{nm}$ such that $z^\top \begin{bmatrix} F' & G' \end{bmatrix} E = 0$, which reads

$$z^\top U(\tau) = z^\top \begin{bmatrix} F' & G' \end{bmatrix} \tilde{\lambda}(\tau) + z^\top G'\tilde{U}(\tau). \tag{V.122}$$

Being both $\tilde{U}(\tau)$ and $\tilde{\lambda}$ arbitrarily small in the arbitrarily long interval $\tau \in [t, t+T]$, we conclude $D^n(u) \notin \Omega^C_{nm}$.

IV) The case of $y^\top B = 0$. Consider the case in which $z^\top B = 0$, namely, the columns $b_1, \ldots, b_m$ of B satisfy $b_1, \ldots, b_m \in \ker(z^\top)$. In that case, along direction z, the system dynamics reads as

$$\begin{aligned} z^\top \dot{x}(\tau) &= z^\top A x(\tau) + z^\top B u(\tau) \\ &= z^\top A x(\tau), \end{aligned} \tag{V.123}$$

We can distinguish two cases: either $\ker(z^\top) = \ker(z^\top A)$ or $\ker(z^\top) \neq \ker(z^\top A)$. If $\ker(z^\top) = \ker(z^\top A)$, then $\ker(z^\top)$ must be an invariant subspace of A of dimension $n-1$. Since $\ker(y^\top)$ is

A-invariant, and since $b_1, \dots, b_m \in \ker(y^\top)$, there are at most $n-1$ linearly independent vectors between the columns of $B, \dots, A^{n-1}B$. This is a contradiction, since we assumed (A, B) controllable, thus, it cannot hold $\ker(z^\top) = \ker(z^\top A)$.
At last, consider the case $\ker(z^\top) \neq \ker(z^\top A)$. Since (V.101) holds for all $\tau \in [t, t+T]$, we can write

$$\begin{aligned} z^\top x(\tau) &\leq \epsilon \\ z^\top \dot{x}(\tau) = z^\top A x(\tau) &\leq \epsilon, \end{aligned} \tag{V.124}$$

from which we deduce that we can write $x(\tau) = \bar{x}(\tau) + \tilde{x}(\tau)$, with $\bar{x}(\tau) \in \ker(z^\top) \cap \ker(y^\top A)$ and $|\tilde{x}(\tau)| \leq \epsilon$. Since $\ker(z^\top) \neq \ker(z^\top A)$ and since the dimension of each kernel is at most $n-1$, then $\dim(\ker(z^\top) \cap \ker(z^\top A)) \leq n-2$.
This means that $\bar{x}(\tau)$ spans persistently at most $n-2$ directions, namely, there exists $z_2 \in \mathbb{R}^n$, $z_2 \perp z \perp (\ker(z^\top) \cap \ker(z^\top A))$, such that

$$|z_2^\top x(\tau)| = |z_2^\top \bar{x}(\tau) + z_2^\top \tilde{x}(\tau)| = |z_2^\top \tilde{x}(\tau)| \leq \epsilon. \tag{V.125}$$

We can thus repeat the same procedure as before, checking if $z_2^\top B \neq 0$. If $z_2^\top B \neq 0$, we can repeat the reasoning in points I), II), III); otherwise, we can repeat the above reasoning to find another direction z_3 such that $z_3^\top x_\tau \leq \epsilon$ (and repeat the process until we find some $z_i^\top B \neq 0$, which must exist since $B \neq 0$. In each of these cases, $\mathrm{D}^n(\mathrm{u}) \notin \Omega^{\mathrm{C}}_{nm}$, which concludes the proof for the first statement of the theorem.

We move now to the second statement of the theorem. Since it has an analogous proof, we sketch only the main differences with respect to the proof given before. We want to prove by contraposition that $\mathrm{D}^{n+1}(\mathrm{u}) \in \Omega^{\mathrm{C}}_{(n+1)m} \implies \sigma(\mathrm{u}, x(0) \in \Omega^{\mathrm{C}}_{n+m}$ for $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{xu}}$ and for all $x(0) \in \mathbb{R}^n$ by showing that for all $x(0) \in \mathbb{R}^n$, $\sigma(\mathrm{u}, x(0)) \notin \Omega^{\mathrm{C}}_{n+m} \implies \mathrm{D}^{n+1}(\mathrm{u}) \notin \Omega^{\mathrm{C}}_{(n+1)m}$.

I) The lack of PE of (x, u) constrains the system input. If $(\mathrm{x}, \mathrm{u}) \notin \Omega^{\mathrm{C}}_{n+m}$ we have that for all $T, \epsilon > 0$ we can find $z = (z_x, z_u) \in \mathbb{R}^{n+m}$, $t > 0$ such that

$$z_x^\top x(\tau) = -z_u^\top u(\tau) + \chi(\tau), \tag{V.126}$$

where $\chi(\tau) \in \mathbb{R}$, $|\chi(\tau)| \leq \epsilon$ for all $\tau \in [t, t+T]$. By isolating $u(\tau)$ and deriving it, we obtain

$$\dot{u}(\tau) = -(z_u^\top)^\dagger z_x^\top \left(Ax(\tau) + Bu(\tau)\right) + (z_u^\top)^\dagger \dot{\chi}(\tau) + \dot{v}(\tau), \tag{V.127}$$

where $\dot{v}(\tau) \in \ker(z_u^\top)$.

II) The closed-loop dynamics depends only on $x(\tau), v(\tau), u(\tau)$. Since we can write $\dot{x}(\tau) = Ax(\tau) + Bu(\tau)$, in the interval $\tau \in [t, t+T]$ we can use (V.127) to express each $u(\tau), \dot{u}(\tau), \dots, \dot{u}^{(n)}(\tau)$ as a linear function of $x(\tau), u(\tau), \dot{v}(\tau), \dots, \dot{v}^{(n)}(\tau)$ (similarly to (V.116)) plus an arbitrarily small quantity $\tilde{U}(\tau)$. In other words, we have

$$\begin{bmatrix} u(\tau) \\ \vdots \\ \dot{u}^{(n)}(\tau) \end{bmatrix} = K \begin{bmatrix} u(\tau) \\ x(\tau) \\ \dot{v}(\tau) \\ \vdots \\ \dot{v}^{(n)}(\tau) \end{bmatrix} + \tilde{U}(\tau) \tag{V.128}$$

III) A lack of PE in (x, u) **implies** $\mathrm{D}^{n+1}(\mathrm{u})$ **is not PE.** Since (x, u) is not PE, it spans at most $n+m-1$ directions. Since $v(\tau) \in \ker(z_u^\top)$, the vector $(\dot{v}(\tau), \dots, \dot{v}^{(n)}(\tau))$ spans at most $(m-1)n$ directions. Overall, we have that the right-hand side of (V.128) spans persistently only $n+m-1+(m-1)n = (n+1)m-1$ directions, which means that the signal $(u(\tau), \dots, \dot{u}^{(n)}(\tau)) \in \mathbb{R}^{(n+1)m}$ on the left-hand side of (V.128) spans persistently only $(n+1)m-1$ directions, namely, $\mathrm{D}^{n+1}(\mathrm{u}) \notin \Omega^{\mathrm{C}}_{(n+1)m}$.
※

V.4.9 Proof of Theorem II.5

Since the proof is analogous to the one for discrete-time systems, we only highlight where they differ. We show that given $\sigma \in \mathbb{L}^{\mathrm{C}}_{\mathrm{x}}$, if $\mathrm{D}^n(\mathrm{u}) \in C^\infty_b(\mathbb{R}^{nm})$ is PPE of degree at most $n' \leq n-1$, then for all $x(0) \in \mathbb{R}^n$, $\mathrm{x} = \sigma(\mathrm{u}, x_0) \notin \Omega^{\mathrm{C}}_n$ regardless of the initial condition.

I) $x(t)$ **as a linear function of older inputs.** Consider $x_0 = 0$. For all t, we can write

$$x(t+\delta) = e^{A\delta}x(t) + \int_t^{t+\delta} e^{A(t+\delta-\tau)}Bu(\tau)\mathrm{d}\tau. \tag{V.129}$$

By writing the Taylor expansion for $x(t+\delta)$, it holds also

$$\begin{aligned} x(t+\delta) =& x(t) + \dot{x}(t)\delta + \frac{\ddot{x}}{2!}(t)\delta^2 + \dots \\ =& e^{A\delta}x(t) + RU(t) + r(t), \end{aligned} \tag{V.130}$$

where

$$\begin{aligned} R &:= \begin{bmatrix} B\delta & \frac{AB}{2}\delta^2 & \dots & \frac{A^{n-1}B}{n!}\delta^n \end{bmatrix}, \\ U(t) &:= (u(t), \dots, \dot{u}^{(n-1)}(t)),\ r(t) := \sum_{i=n}^{\infty} o(\delta^i)\dot{u}^{(i)}(t). \end{aligned} \tag{V.131}$$

Notice that, given $M := \|\mathrm{d}^n(\mathrm{x})\|_\infty$, the remainder $r(t)$ can be bounded by $|r(t)| \leq M\delta^n/(n!)$. Confronting (V.129) and (V.130), we obtain

$$\int_t^{t+\delta} e^{A(t+\delta-\tau)} Bu(\tau)\mathrm{d}\tau = RU(t) + r(t) \tag{V.132}$$

Given any $\bar{T} > 0$ (to be chosen later), we can write $x(t+\bar{T})$ as

$$x(t+\bar{T}) = e^{A\bar{T}}x(t) + \int_t^{t+\bar{T}} e^{A(t+\bar{T}-\tau)} Bu(\tau)\mathrm{d}\tau. \tag{V.133}$$

Now, let $N \in \mathbb{N}$ (to be chosen later) and $\delta = \bar{T}/N$. Recalling (V.132) to approximate the integral in (V.133), we obtain

$$\begin{aligned} x(t+N\delta) &= e^{AN\delta}x(t) + \int_t^{t+N\delta} e^{A(t+N\delta-\tau)} Bu(\tau)\mathrm{d}\tau \\ &= e^{AN\delta}x(t) + \sum_{i=0}^{N-1} \int_{t+i\delta}^{t+(i+1)\delta} e^{A(t+N\delta-\tau)} Bu(\tau)\mathrm{d}\tau \\ &= e^{AN\delta}x(t) + \sum_{i=0}^{N-1} r(t+i\delta) + \\ & \underbrace{\begin{bmatrix} e^0 & e^{A\delta} & \dots & e^{A(N-1)\delta} \end{bmatrix}}_{:=\bar{A}} \underbrace{(I_N \otimes R)}_{:=\bar{R}} \begin{bmatrix} U(t+(N-1)\delta) \\ U(t+(N-2)\delta) \\ \vdots \\ U(t) \end{bmatrix}. \end{aligned} \tag{V.134}$$

II) $x(t)$ can be approximated arbitrarily well by a linear function of the previous inputs. We show now the terms $e^{AN\delta}x(t)$ and $\sum_{i=0}^{N-1} r(t+i\delta)$ can be made arbitrarily small. Pick any $\epsilon > 0$. Since e^A is Schur and x is bounded, there exists a sufficiently large $\bar{T} = N\delta > 0$ such that $|e^{A\bar{T}}x(t)| \leq \epsilon$ for all $t \in \mathbb{R}_{\geq 0}$. Next, we choose N. It holds

$$\left|\sum_{i=0}^{N-1} r(t+i\delta)\right| \leq N\frac{M\delta^n}{n!} = \frac{M\bar{T}^n}{n!N^{n-1}}, \tag{V.135}$$

so, given any ϵ and picking

$$N \geq \sqrt[n-1]{\frac{M\bar{T}^n}{n!\epsilon}}, \tag{V.136}$$

we obtain $|\sum_{i=0}^{N-1} r(t+i\delta)| \leq \epsilon$.

III) If the remaining term is not PPE of degree n, $\mathrm{x} \notin \Omega_n^{\mathrm{c}}$. At last, applying Lemma II.6, if $\mathrm{D}^n(\mathrm{u}) \notin \Omega_{nm,n}^{\mathrm{c}}$, we have that for any $T, \bar{T}, \epsilon > 0$ there exists $\bar{N}, \bar{t} > 0$ such that, if $N \geq$

$\max(\bar{N}, \sqrt[n-1]{\frac{M\bar{T}^n}{n!\epsilon}})$, we can rewrite (V.134) as

$$\begin{aligned} x(\tau + N\delta) =& e^{AN\delta}x(t) + \sum_{i=0}^{N-1} r(t+i\delta) \\ &+ \bar{A}\bar{R}G\lambda(\tau) + \bar{A}\bar{R}\tilde{X}(\tau), \end{aligned} \tag{V.137}$$

where $\lambda(\tau) \in \mathbb{R}^{n-1}$, $|\tilde{X}(\tau)| \leq \epsilon$ for all $\tau \in [\bar{t}, \bar{t}+T]$. We can thus conclude that $\mathrm{x} = \sigma(\mathrm{u}, 0) \notin \Omega_n^{\mathrm{C}}$ similarly as done in Theorem II.4 (since x is a sum of a signal spanning only $n-1$ directions plus arbitrarily small perturbations). To conclude the proof, since σ is a stable linear system, we have that all solutions $\mathrm{x} = \sigma(\mathrm{u}, x(0))$ converge exponentially to those initialized in $x(0) = 0$. Since a vanishing term cannot guarantee PE of a signal, we can conclude that for all $x(0) \in \mathbb{R}^n$, if $\mathrm{D}^n(\mathrm{u}) \notin \Omega_{nm,n}^{\mathrm{C}}$, then $\mathrm{x} = \sigma(\mathrm{u}, x(0)) \notin \Omega_n^{\mathrm{C}}$. We omit the second statement since it combines arguments from previous proofs.

※

V.4.10 Proof of Theorem II.9

Looking at decomposition (II.70), the dynamics of the last subsystem can be written as a single-input system in the form

$$\dot{x}_m = A_m x_m + b_{mm} u_m, \tag{V.138}$$

with $x_m \in \mathbb{R}^{n_m}, u_m \in \mathbb{R}$ and A_m, b_m of appropriate dimensions. Since $u_m(t)$ contains at least n_m spectral lines, by [41, Lemma 3.3] also $x_m(t)$ contains the same number of spectral lines at the same frequencies. Furthermore, by [41, Prop. 5.1], $x_m(t)$ is PE, since these spectral lines are linearly independent. We denote these spectral lines as $\hat{x}_m(\omega_1^m), \dots, \hat{x}_m(\omega_{n_m}^m) \in \mathbb{C}^{n_m}$.
Consider the $(m-1)$-th subsystem. Its dynamics is given by

$$\dot{x}_{m-1} = A_{m-1}x_{m-1} + b_{(m-1)(m-1)}u_{m-1} + b_{(m-1)m}u_m, \tag{V.139}$$

with $x_{m-1} \in \mathbb{R}^{n_{m-1}}$, $u_{m-1} \in \mathbb{R}$ and the other matrices of appropriate dimensions. Being the system linear, we can apply the superposition principle and write the solution $x_{m-1}(t)$ as

$$x_{m-1}(t) = \bar{x}_{m-1}(t) + \tilde{x}_{m-1}(t) \tag{V.140}$$

where $\bar{x}_{m-1}(t)$ is the system response to input $u_{m-1}(t)$ and $\tilde{x}(t)_{m-1}$ is the system response to input $u_m(t)$. Repeating the above reasoning, we know $\bar{x}_{m-1}(t)$ contains at least n_{m-1} linearly independent spectral lines $\hat{x}_{m-1}(\omega_1^{m-1}), \dots, \hat{x}_{m-1}(\omega_{n_{m-1}}^{m-1}) \in \mathbb{C}^{n_{m-1}}$, and so does $x_{m-1}(t)$, since $\bar{x}_{m-1}(t)$ and $\tilde{x}_{m-1}(t)$ have different spectral content by assumption.
Notice that the signal $(x_{m-1}(t), x_m(t))$ contains at least $n_{m-1} + n_m$ spectral lines which are linearly independent, since

i) n_{m-1} of those in $x_{m-1}(t)$ that are not present in $x_m(t)$ are linearly independent eachother

ii) n_m of those in $x_m(t)$ are linearly independent eachother,

iii) the spectral lines in these two sets are linearly independent, since they are in the forms

$$\begin{bmatrix} \hat{x}_{m-1}(\omega_1^{m-1}) \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} \hat{x}_{m-1}(\omega_{n_{m-1}}^{m-1}) \\ 0 \end{bmatrix}, \begin{bmatrix} \hat{x}_{m-1}(\omega_1^{m}) \\ \hat{x}_m(\omega_1^m) \end{bmatrix}, \dots, \begin{bmatrix} \hat{x}_{m-1}(\omega_{n_m}^{m}) \\ \hat{x}_m(\omega_{n_m}^m) \end{bmatrix} \in \mathbb{C}^{n_{m-1}+n_m}. \tag{V.141}$$

so the signal $(x_{m-1}(t), x_m(t))$ is PE by [41, Lemma 3.4]. This reasoning can be done for all subsystems, and since the overall system solution is given by

$$x(t) = (x_1(t), \dots, x_m(t)), \tag{V.142}$$

then $x(t)$ contains $\sum_{i=1}^m n_i = n$ linearly independent spectral lines and thus it is PE by [41, Lemma 3.4].

※

V.5 Proofs for Chapter III

V.5.1 Proof of Lemma III.1

Without loss of generality, we consider $(c^\top, A)$ in observability canonical form. We prove now the first statement. Notice that, for a generic $M \in \mathbb{R}^{n\times n+1}$, it holds

$$\operatorname{rank}(M) = n - \dim\left(\operatorname{leftker}(M)\right). \tag{V.143}$$

For this reason, we want to show that, for all $\lambda \in \sigma(A - \psi c^\top) \cap \sigma(A)$,

$$\dim\left(\operatorname{leftker}\begin{bmatrix} A - \psi c^\top - \lambda I & b \end{bmatrix}\right) = \dim\left(\operatorname{leftker}\begin{bmatrix} A - \lambda I & b \end{bmatrix}\right). \tag{V.144}$$

Being A and $A - \psi c^\top$ cyclic ($A - \psi c^\top$ must be cyclic since $(c^\top, A - \psi c^\top)$ is observable), the geometric multiplicity of each eigenvalue is one. This means that the matrices $A - \lambda I$ and $A - \psi c^\top - \lambda I$ will have a left kernel of dimension 1 for each shared eigenvalue. We want, however, to prove that the associated left eigenvectors are the the same for both matrices (in order to conclude the proof for the matrices $[A - \lambda I\ b]$ and $[A - \psi c^\top - \lambda I\ b]$). For all $\lambda \in \sigma(A - \psi c^\top) \cap \sigma(A)$, the associated left eigenvector $w_\lambda \in \mathbb{R}^n$ is such that $w_\lambda^\top(A - \psi c^\top - \lambda I) = 0$. Given the observability canonical form of

$(c^\top, A - \psi c^\top)$, it holds that

$$w_\lambda^\top (A - \lambda I) = \begin{bmatrix} 0 & \dots & w^\top \psi \end{bmatrix}$$

$$w_\lambda^\top \begin{bmatrix} -\lambda & 0 & \dots & 0 & a_1 \\ 1 & -\lambda & 0 & \dots & a_2 \\ 0 & 1 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 1 & a_n - \lambda \end{bmatrix} = \begin{bmatrix} 0 & \dots & w_\lambda^\top \psi \end{bmatrix}, \tag{V.145}$$

where $a_1, \dots, a_n$ are the opposite of the coefficients of the characteristic polynomial of A. From the first $n-1$ equation, we obtain that $w_\lambda = (w_1, \dots, w_n)$ must be of the form

$$w_\lambda^\top = w_1 \begin{bmatrix} 1 & \lambda & \dots & \lambda^{n-1} \end{bmatrix}. \tag{V.146}$$

However, it can be checked that this vector is also the unique left eigenvector of A associated to the eigenvalue λ, and so $w_\lambda^\top (A - \lambda I) = 0$. To conclude the proof, notice that any $w_\lambda \in \mathtt{leftker}(A - \psi c^\top - \lambda I)$ is in the left kernel of $[A - \psi c^\top \; b]$ only if $w_\lambda^\top b = 0$, and the same holds for the matrix $[A - \lambda I \; b]$ We have thus shown that the two matrices $[A - \psi c^\top \; b]$ and $[A - \lambda I \; b]$ have the same left eigenvectors, which implies they have the same rank.
We prove now the second statement. We show the result by contraposition, namely, we suppose that the second statement does not hold and then we derive that $\lambda \in \sigma(A)$. If

$$\operatorname{rank} \begin{bmatrix} A - \psi c^\top - \lambda I & \psi \end{bmatrix} < n \tag{V.147}$$

for some $\lambda \in \sigma(A - \psi c^\top)$, then there exists $w \in \mathbb{R}^n$ such that i) $w^\top \psi = 0$ and ii) $w^\top (A - \psi c^\top - \lambda I) = 0$. Condition ii) becomes

$$w^\top (A - \lambda I) = w^\top \psi c, \tag{V.148}$$

and using condition i), we have $w^\top (A - \lambda I) = 0$, namely, $\lambda \in \sigma(A)$. This proves, by contraposition, that

$$\lambda \in \sigma(A) \implies \operatorname{rank} \begin{bmatrix} A - \psi c^\top - \lambda I & \psi \end{bmatrix} \geq n \tag{V.149}$$

and concludes the proof.

※

V.5.2 Proof of Proposition III.3

Recall that (following Remark III.4), the existence of a full rank matrix $\Pi \in \mathbb{R}^{n\times z}$ satisfying (III.14) defines a coordinate change $\zeta \mapsto (\xi, \eta)$ such that the gazer is in Kalman observability form:

$$\begin{bmatrix}\dot{\xi}\\ \dot{\eta}\end{bmatrix} = \begin{bmatrix}A & 0\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}\xi\\ \eta\end{bmatrix} + \begin{bmatrix}B\\ B_2\end{bmatrix}u \qquad \text{(V.150)}$$
$$y = C\xi.$$

Notice that any uncontrollable subspace $\mathscr{Z} \subset \mathbb{R}^z$ must be such that $\mathscr{Z} \subset \ker(\Pi)$. In fact, given the change of coordinates in (V.150), $\xi = \Pi\zeta$ follows a dynamics which is stabilizable by assumption and that does not depend on η (so, any uncontrollable subspace must lie in $\ker(\Pi)$). Consider the gazer in filter form, whose dynamics read

$$\dot{\zeta} = \mathscr{A}\zeta + \mathscr{L}y + \mathscr{B}u. \qquad \text{(V.151)}$$

Notice that, for any $\zeta \in \ker(\Pi)$, $\mathscr{L}y = \mathscr{L}\mathscr{C}\zeta = \mathscr{L}C\Pi\zeta = 0$. So, for all $\zeta \in \ker\Pi$, it holds that $\dot{\zeta} = \mathscr{A}\zeta + \mathscr{B}u$. Since $(\mathscr{A}, \mathscr{B})$ is stabilizable by assumption, the system must be at least stabilizable.

※

V.5.3 Proof of Lemma III.5

Without loss of generality, consider Λ, A in Jordan form with the same ordering of eigenvalues. In particular, since the pair (Λ, ℓ) is controllable by assumption, we use the transform [95, Example 3.4], which brings the pair (Λ, ℓ) in the form described in Section III.2.1.

We start by proving that if the two matrix are similar, a solution to (III.28) can always be found. Since A, Λ are in Jordan form, we consider solutions Π to the first equation ($\Pi\Lambda = A\Pi$) of (III.28) of the form explained in Subsection (III.2.1). Expliciting the matrix Π and the vector ℓ using (III.3) and (III.6), we find

$$\Pi\ell = b$$
$$\begin{bmatrix}\Pi_{11} & \dots & \Pi_{1u}\\ \dots & \dots & \dots\\ \Pi_{u1} & \dots & \Pi_{uu}\end{bmatrix}\begin{bmatrix}e_{n_1^\Lambda}\\ \dots\\ e_{n_u^\Lambda}\end{bmatrix} = b, \qquad \text{(V.152)}$$

where u is the number of Jordan blocks in A and Λ. Being the matrices A and Λ similar and cyclic, for each corresponding eigenvalue λ_i, we have that $\Pi_{ii} \in \mathbb{R}^{n_i\times n_i}$ satisfying the first equation in (III.28) is square and upper triangular (of the form (III.4)), n_i being the algebraic multiplicity of λ_i (which is the same in both matrices). All the other Π_{ij} are zero. For this reason, equation (V.152) reads as u equations of the type

$$\Pi_{ii}e_{n_i^\Lambda} = b_i. \qquad \text{(V.153)}$$

To make an example, if $n_i = 3$, we have

$$\begin{gathered}\Pi_{ii} e_{n_i^\Lambda} = b_i, \\ \begin{bmatrix} a_i & b_i & c_i \\ 0 & a_i & b_i \\ 0 & 0 & a_i \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_i^1 \\ b_i^2 \\ b_i^3 \end{bmatrix}.\end{gathered} \tag{V.154}$$

Notice that, for all $i = 1, \ldots, u$ these systems have a solution for any b_i in the unknowns describing Π_{ii}, which can be easily found as (for the case of the example (V.154))

$$\Pi_{ii} = \begin{bmatrix} b_i^3 & b_i^2 & b_i^1 \\ 0 & b_i^3 & b_i^2 \\ 0 & 0 & b_i^3 \end{bmatrix}. \tag{V.155}$$

We prove now the second part, namely, that if A and Λ are not similar, system (III.28) may not be solvable for certain b. If A, Λ are not similar, being the matrices of the same dimension, either there exists an eigenvalue $\lambda_i \in \sigma(A)$, $\lambda_i \notin \sigma(\Lambda)$ or there exists an eigenvalue λ_i in A whose algebraic multiplicity in A is greater than its algebraic multiplicity in Λ. In the first case, Π has a row of zeros, thus it is not full rank and for all $b \in \operatorname{im}(\Pi)^\perp$, $b \neq \Pi\ell$. In the second case, Π_{ii} is rectangular, tall, and system (V.152) is of the form

$$\begin{gathered}\Pi_{ii} e_{n_i^\Lambda} = b_i \\ \begin{bmatrix} a_i & b_i & c_i \\ 0 & a_i & b_i \\ 0 & 0 & a_i \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_i^1 \\ b_i^2 \\ b_i^3 \\ b_i^4 \end{bmatrix}.\end{gathered} \tag{V.156}$$

Notice however that if $b_i = (0, 0, 0, 1)$, there are no solutions to (V.156) in the unknowns a_i, b_i, c_i.

※

V.5.4 Proof of Lemma III.6

Assuming a solution to equations (III.28) exists, being the system linear its solutions are in the form

$$\Pi = \Pi_P + \Pi_H, \tag{V.157}$$

where Π_P is the particular solution and Π_H are the homogeneus solutions, namely any Π_H which solves

$$\begin{aligned} \Pi_H \Lambda &= A \Pi_H \\ \Pi_H \ell &= 0. \end{aligned} \tag{V.158}$$

We start by finding the particular solution. Premultiplying the second equation in (III.28) by A, and using the first equation we obtain

$$\Pi_P \Lambda \ell = Ab. \tag{V.159}$$

By repeating this procedure, the particular solution Π_P must satisfy

$$\begin{aligned} \Pi_P \ell &= b \\ \Pi_P \Lambda \ell &= Ab \\ \ldots &= \ldots \\ \Pi_P \Lambda^{n-1} \ell &= A^{n-1} b, \end{aligned} \tag{V.160}$$

which means that it holds

$$\Pi_P R_{\Lambda,\ell} = R_{A,b} \tag{V.161}$$

from which, being (Λ, ℓ) controllable, $R_{\Lambda,\ell}$ is invertible, and we obtain

$$\Pi_P = R_{A,b} R_{\Lambda,\ell}^{-1}. \tag{V.162}$$

By repeating the same procedure, it is easy to see that the homogeneous solution is such that

$$\Pi_H R_{\Lambda,\ell} = 0, \tag{V.163}$$

from which we derive that $\Pi_H = 0$ is unique (being $R_{\Lambda,\ell}$ full rank by controllability of (Λ, ℓ)), and thus the complete solution is unique and it is given by

$$\Pi = \Pi_P + \Pi_H = R_{A,b} R_{\Lambda,\ell}^{-1} + 0. \tag{V.164}$$

※

V.5.5 Proof of Lemma III.7

In order to prove this result, it is sufficient to repeat the same proof of Lemma III.5 by applying the Jordan controllability decomposition [95, Example 3.4] to both pairs (A, b) and (Λ, ℓ). So, without loss of generality, consider A, Λ in Jordan form with the same ordering of the eigenvalues. Under the same arguments as per Lemma III.5, if A, Λ are similar we obtain that the second equation in (III.28) reads as

$$\Pi_{ii} e_{n_i} = e_{n_i}, \tag{V.165}$$

for all $i = 1, \ldots, u$, u being the number of Jordan blocks in A, Λ. Given the square form of each Π_{ii} under the similarity assumption (see (III.4)), the unique solution of this system is $\Pi_{ii} = I_{n_i}$, where n_i is the algebraic multiplicity of the i-th eigenvalue in A and Λ.

We can prove that similarity of A, Λ is also necessary by following the same arguments of Lemma III.5 but considering both pairs (A, b) and (Λ, ℓ) in Jordan controllability form. Full rankness of Π is then derived since the expression of the solution is given by (Lemma III.6)

$$\Pi = R_{A,b} R_{\Lambda,\ell}^{-1}. \tag{V.166}$$

※

V.5.6 Proof of Theorem III.1

We prove this theorem by considering $r = n$. Leveraging Remark III.6, it is easy to obtain the final statement of Theorem III.1. Define $\Pi := [\Pi_y \ \Pi_u]$, with $\Pi_y, \Pi_u \in \mathbb{R}^{n\times n}$. Then, (III.36) can be written as

$$\begin{cases} \Pi_y \Lambda = (A - \Pi_y \ell c^\top)\Pi_y \\ \theta_y^\top = c^\top \Pi_y \end{cases} \tag{V.167}$$

and

$$\begin{cases} \Pi_u \Lambda = (A - \Pi_y \ell c^\top)\Pi_u \\ \Pi_u \ell = b \\ \theta_u^\top = c^\top \Pi_u, \end{cases} \tag{V.168}$$

where we notice that $A - \Pi_y \ell c^\top$ appears in both equations. Consider the second system (V.168), which is in the same form as (III.28) (notice that, since θ_u is a degree of freedom, it does not introduce constraints).

Due to the observability of $(c^\top, A)$, we can choose $l \in \mathbb{R}^n$ as the unique vector such that $\sigma(A - LC) = \sigma(\Lambda)$ (furthermore, being $A - lc^\top, \Lambda$ cyclic, the geometric multiplicity of all the eigenvalues is 1 and thus Λ and $A - lc^\top$ are also similar). We thus impose the new equation $l = \Pi_y \ell$, and by denoting $A' := A - lc^\top$, we obtain from (V.167), (V.168) the following systems:

$$\begin{cases} \Pi_y \Lambda = A'\Pi_y \\ \Pi_y \ell = \psi, \end{cases} \qquad \begin{cases} \Pi_u \Lambda = A'\Pi_u \\ \Pi_u \ell = b, \end{cases} \tag{V.169}$$

where we have dropped the equations $\theta_y^\top = c^\top \Pi_y$ and $\theta_u^\top = c^\top \Pi_u$ since, being θ_y, θ_u degrees of freedom, they do not impose constraints on Π_y, Π_u. We then apply Proposition III.5, and we know that, since the pair (Λ, ℓ) is controllable and since Λ, A' are similar by construction, both systems can be solved and their solutions Π_y, Π_u are given (Lemma III.6) by

$$\Pi = \begin{bmatrix} \Pi_y & \Pi_u \end{bmatrix} = \begin{bmatrix} R_{A',l} R_{\Lambda,\ell}^{-1} & R_{A',b} R_{\Lambda,\ell}^{-1} \end{bmatrix} \tag{V.170}$$

To conclude full rankness of Π, it is sufficient to apply Lemma III.1 and notice that, either if (A, b) is controllable or $\sigma(A) \cap \sigma(\Lambda) = \emptyset$, then the pair $(A - lc^\top, [b \ l])$ is controllable, and thus $[R_{A',l} \ R_{A',b}]$

is of full row rank.

※

V.5.7 Proof of Theorem III.2

This proof follows the same steps as the one of Theorem III.1 and considering $r = n$. Leveraging Remark III.6, it is easy to obtain the final statement of Theorem III.2 for $r > n$.
Defining $\Pi = [\Pi_y^1, \dots, \Pi_y^p, \Pi_u^1, \dots, \Pi_u^m]$ and considering the block structure in (III.42), system (III.14) can be rewritten as

$$\Pi_y^i(\Lambda + \ell\theta_y^{ii}) + \sum_{j\neq i} \Pi_y^j \ell\theta_y^{ji} = A\Pi_y^i, \quad i = 1, \dots, p, \tag{V.171}$$

together with

$$\sum_{i=1}^{p} \Pi_y^i \ell\theta_u^{ij} + \Pi_u^j \Lambda = A\Pi_u^j, \quad j = 1, \dots, m. \tag{V.172}$$

Imposing $\mathscr{C} = C\Pi$,

$$\theta_y^{ij} = c_i \Pi_y^j, \qquad \theta_u^{ij} = c_i \Pi_u^j, \tag{V.173}$$

where c_i, $i = 1, \dots, p$ are the rows of C. Substituting (V.173) in (V.171), we obtain

$$\begin{aligned} \Pi_y^i \Lambda &= A\Pi_y^i - \sum_{j=1}^{p} \Pi_y^j \ell\theta_y^{ji}, \quad i = 1, \dots, p, \\ \Pi_y^i \Lambda &= A\Pi_y^i - \sum_{j=1}^{p} \Pi_y^j \ell c_j \Pi_y^i, \quad i = 1, \dots, p, \\ \Pi_y^i \Lambda &= (A - \sum_{j=1}^{p} \Pi_y^j \ell c_j)\Pi_y^i, \quad i = 1, \dots, p, \end{aligned} \tag{V.174}$$

and substituting (V.173) in (V.172),

$$\begin{aligned} \Pi_u^j \Lambda &= A\Pi_u^j - \sum_{i=1}^{p} \Pi_y^i \ell\theta_u^{ij}, \quad j = 1, \dots, m, \\ \Pi_u^j \Lambda &= A\Pi_u^j - \sum_{i=1}^{p} \Pi_y^i \ell c_i \Pi_u^j, \quad j = 1, \dots, m, \\ \Pi_u^j \Lambda &= (A - \sum_{i=1}^{p} \Pi_y^i \ell c_i)\Pi_u^j, \quad j = 1, \dots, m. \end{aligned} \tag{V.175}$$

Notice that we may rewrite $A - \sum_{i=1}^{p} \Pi_y^i \ell c_i$ as

$$A - \sum_{i=1}^{p} \Pi_y^i \ell c_i = A - \begin{bmatrix} \Pi_y^1 \ell & \dots & \Pi_y^p \ell \end{bmatrix} C. \tag{V.176}$$

By observability of (C, A), there exists $L \in \mathbb{R}^{n \times p}$ such that $A - LC$ is cyclic and $\sigma(A - LC) = \sigma(\Lambda)$ [228, Lemma 2.2, Thm. 2.1]. We thus impose the new equation

$$\begin{bmatrix} \Pi_y^1 \ell & \dots & \Pi_y^p \ell \end{bmatrix} = L, \tag{V.177}$$

which, writing $L = [l_1, \dots, l_p]$, results in p equations of the type

$$\Pi_y^i \ell = l_i, \quad i = 1, \dots, p. \tag{V.178}$$

Denoting $A' \coloneqq A - LC$, we obtain that equations (V.174) are equivalent to

$$\begin{aligned} \Pi_y^i \Lambda &= A' \Pi_y^i, \\ \Pi_y^i \ell &= l_i, \quad i = 1, \dots, p. \end{aligned} \tag{V.179}$$

Next, considering also the equation $\Pi \mathscr{B} = B$, and denoting $B = [b_1, \dots, b_m]$, we obtain that equations (V.175) are equivalent to

$$\begin{aligned} \Pi_u^j \Lambda &= A' \Pi_u^j, \\ \Pi_u^j \ell &= b_j, \quad j = 1, \dots, m. \end{aligned} \tag{V.180}$$

By applying Lemma III.6, all of these $p + m$ systems of equations have a solution Π_y^i, Π_u^j which can be found with Lemma III.6. Full rankness of Π follows as per Theorem III.1.

※

V.5.8 Proof of Lemma III.8

It can be checked by simple calculations that $\Pi \coloneqq \gamma^{-1} [A + \lambda I_n \; B]$ solves the equations

$$\Pi \begin{bmatrix} A & B \\ 0 & -\lambda I_m \end{bmatrix} = A\Pi, \quad \Pi \begin{bmatrix} 0 \\ \gamma I_m \end{bmatrix} = B, \quad \Pi I = \Pi \tag{V.181}$$

arising by substituting the matrix $\mathscr{A}, \mathscr{B}, \mathscr{C}$ in (III.14) with the matrices in (III.53). The fact that Π is full rank follows directly from the controllability Assumption III.1 on (A, B), which ensures that

$$\operatorname{rank} \begin{bmatrix} A + \lambda I & B \end{bmatrix} = n \tag{V.182}$$

for all $\lambda \in \sigma(A)$.

※

V.5.9 Proof of Lemma III.9

We use the PBH test, which requires that

$$\operatorname{rank}\begin{bmatrix} sI_n - A & -B & 0 \\ 0 & (s+\lambda)I_m & \gamma I_m \end{bmatrix} = n + m \tag{V.183}$$

for all $s \in \sigma(A) \cup \{-\lambda\}$. From Assumption III.1, the first n rows are linearly independent for all $s \in \sigma(A) \cup \{-\lambda\}$. Therefore, the result follows by noticing that the remaining m rows are linearly independent from the previous ones. ※

V.5.10 Proof of Lemma III.10

Recall the notation F and g from (III.92). By [11, Thm. 2.17], (F, g) is controllable if and only if the $2n$ rows of

$$\mathcal{H}(s) := (sI - F)^{-1}g \tag{V.184}$$

are linearly independent over the field of complex numbers, i.e., if and only if there exists no $w \in \mathbb{C}^{2n}$, $w \neq 0$, such that

$$w^\top \mathcal{H}(s) = 0, \quad \forall s \in \mathbb{C}. \tag{V.185}$$

Notice that $\mathcal{H}(s)$ can be equivalently seen as the transfer function of the cascade between plant (III.78) and the filters (III.80):

$$\mathcal{H}(s) = \begin{bmatrix} (sI - \Lambda)^{-1}\ell c^\top (sI - A)^{-1} b \\ (sI - \Lambda)^{-1}\ell. \end{bmatrix} \tag{V.186}$$

We follow similar steps to [191, Thm. 2.7.3]. Suppose that there exists $w \neq 0$ such that $w^\top \mathcal{H}(s) = 0$ for all s. Split $w = \mathrm{col}(w_1, w_2)$, with $w_1, w_2 \in \mathbb{R}^n$. Then, it holds that

$$\frac{\mathcal{N}_1(s)}{\mathcal{D}_\Lambda(s)}\frac{\mathcal{N}(s)}{\mathcal{D}(s)} + \frac{\mathcal{N}_2(s)}{\mathcal{D}_\Lambda(s)} = 0, \quad \forall s \in \mathbb{C}, \tag{V.187}$$

where $\mathcal{N}(s)$, $\mathcal{D}(s)$, $\mathcal{N}_1(s)$, $\mathcal{N}_2(s)$, and $\mathcal{D}_\Lambda(s)$ are polynomials of s such that $\frac{\mathcal{N}(s)}{\mathcal{D}(s)} = c^\top (sI - A)^{-1} b$, $\frac{\mathcal{N}_1(s)}{\mathcal{D}_\Lambda(s)} = w_1^\top (sI - \Lambda)^{-1}\ell$, and $\frac{\mathcal{N}_2(s)}{\mathcal{D}_\Lambda(s)} = w_2^\top (sI - \Lambda)^{-1}\ell$. For the above expression to be identically zero, it must hold that

$$\mathcal{N}_2(s) = -\mathcal{N}_1(s)\frac{\mathcal{N}(s)}{\mathcal{D}(s)}, \quad \forall s \in \mathbb{C}. \tag{V.188}$$

To ensure that the left- and the right-hand sides are equal, since $\mathcal{N}_2(s)$ has no poles, there must be n pole-zero cancellations in $\mathcal{N}_1(s)\frac{\mathcal{N}(s)}{\mathcal{D}(s)}$. Notice that $\mathcal{N}_1(s)$ and $\mathcal{N}_2(s)$ are at most of degree $n - 1$. Therefore, there must be at least one pole-zero cancellation in $\frac{\mathcal{N}(s)}{\mathcal{D}(s)}$. However, $\mathcal{N}(s)$ and $\mathcal{D}(s)$ are coprime by Assumption III.2, hence we have a contradiction. ※

V.6 Proofs for Chapter IV

V.6.1 Proof of Lemma IV.1

At first, to simplify the expressions for the Lyapunov function, we introduce the following vectorized coordinates:

$$\hat{\theta}_{\mathrm{A}} := \mathrm{vec}(\hat{A}) \in \mathbb{R}^{n^2}, \qquad \tilde{\theta}_{\mathrm{A}} := \mathrm{vec}(\tilde{A}) \in \mathbb{R}^{n^2}. \tag{V.189}$$

It can be verified the following relation holds:

$$\mathrm{vec}(BB^{\dagger}\epsilon\xi^{\top}) = \bar{B}(\xi \otimes I_n)\epsilon, \quad \bar{B} := (I_n \otimes BB^{\dagger}), \tag{V.190}$$

where $\bar{B}$ defines a projection onto $\mathrm{Im}(I_n \otimes B) \subset \mathbb{R}^{n^2}$. Notice that, for any $\hat{\theta}_{\mathrm{A}} \in \mathrm{Im}(I_n \otimes B)$ and $\tau \in \mathbb{R}^{n^2}$, then $\tilde{\theta}_{\mathrm{A}} \in \mathrm{Im}(I_n \otimes B)$ and, since the scalar product of orthogonal vectors is zero and by idempotence of the projection,

$$\begin{aligned}\tilde{\theta}_{\mathrm{A}}^{\top}\tau &= \tilde{\theta}_{\mathrm{A}}^{\top}(\tau_{\parallel} + \tau_{\perp}) = \tilde{\theta}_{\mathrm{A}}^{\top}\tau_{\parallel} + 0 \\ &= \tilde{\theta}_{\mathrm{A}}^{\top}\bar{B}\tau_{\parallel} = \tilde{\theta}_{\mathrm{A}}^{\top}\bar{B}(\tau_{\parallel} + \tau_{\perp}) = \tilde{\theta}_{\mathrm{A}}^{\top}\bar{B}\tau,\end{aligned} \tag{V.191}$$

where $\tau_{\parallel} \in \mathrm{Im}(I_n \otimes B)$ and $\tau_{\perp} \in (I_n \otimes B)^{\perp}$. We rewrite (IV.22) by using the vectorized coordinates defined above:

$$\dot{\hat{\theta}}_{\mathrm{A}} = \mathcal{P}_{\mathrm{vec}^{-1}\,\hat{\theta}_{\mathrm{A}} \in C}\left\{-\gamma\bar{B}\frac{(\xi \otimes I_n)\epsilon}{1 + \nu|\xi||\epsilon|}\right\}. \tag{V.192}$$

The computations from here are similar to [112, Lemma 6.1] but we report them for the reader's convenience. Define

$$V_A(\tilde{\epsilon}, \tilde{\theta}_{\mathrm{A}}) := \frac{1}{\lambda}|\tilde{\epsilon}|^2 + \frac{1}{2\gamma}|\tilde{\theta}_{\mathrm{A}}|^2, \tag{V.193}$$

which is positive definite with respect to $(0, 0)$ and radially unbounded. Note that $\epsilon = (\xi \otimes I_n)^{\top}\tilde{\theta}_{\mathrm{A}} + \tilde{\epsilon}$. Then, using (V.191) and [112, Lemma E.1] to treat the projection operator $\mathcal{P}_{\hat{A} \in C}\{\cdot\}$, the derivative of

V_A along the solutions of (IV.46) is

$$\begin{aligned}
\dot{V}_A &= -2|\tilde{\epsilon}|^2 + \tilde{\theta}_{\mathrm{A}}^\top \mathcal{P}_{\mathrm{vec}^{-1}\ \hat{\theta}_{\mathrm{A}} \in C}\left\{-\bar{B}\frac{(\xi \otimes I_n)\epsilon}{1+\nu|\xi||\epsilon|}\right\} \\
&\leq -2|\tilde{\epsilon}|^2 - \frac{\tilde{\theta}_{\mathrm{A}}^\top \bar{B}(\xi \otimes I_n)\epsilon}{1+\nu|\xi||\epsilon|} = -2|\tilde{\epsilon}|^2 - \frac{\tilde{\theta}_{\mathrm{A}}^\top (\xi \otimes I_n)\epsilon}{1+\nu|\xi||\epsilon|} \\
&\leq -2|\tilde{\epsilon}|^2 - \frac{|\epsilon|^2 - \tilde{\epsilon}^\top \epsilon}{1+\nu|\xi|^2} \\
&\leq -2|\tilde{\epsilon}|^2 - \frac{1}{4}\frac{|\epsilon|^2}{(1+\nu|\xi|^2)^2} + \frac{\tilde{\epsilon}^\top \epsilon}{1+\nu|\xi|^2} - \frac{3}{4}\frac{|\epsilon|^2}{1+\nu|\xi|^2} \\
&= -|\tilde{\epsilon}|^2 - \left(\frac{1}{2}\frac{|\epsilon|}{1+\nu|\xi|^2} - \tilde{\epsilon}\right)^2 - \frac{3}{4}\frac{|\epsilon|^2}{1+\nu|\xi|^2} \leq 0
\end{aligned} \tag{V.194}$$

implying that $(\tilde{\epsilon}(t), \tilde{\theta}_A(t))$ is contained for all $t \in [0, t_f)$ in a compact sublevel set of V_A. We conclude the proof by recalling [112, Lemma E.1] to ensure $\hat{A}(t) \in \Theta$, for all its domain of existence.
We prove now the last point. Using [112, Lemma E.1] to treat the projection operator $\mathcal{P}_{\hat{A} \in C}\{\cdot\}$ and the fact that $|\bar{B}| = 1$ due to (V.190), we can bound $|\dot{\hat{A}}|$ as follows:

$$|\dot{\hat{A}}| \leq |\dot{\hat{A}}|_F = |\dot{\hat{\theta}}_{\mathrm{A}}| \leq \gamma|\bar{B}|\frac{|(\xi \otimes I_n)||\epsilon|}{1+\nu|\xi||\epsilon|} \leq \frac{\gamma|\xi||\epsilon|}{1+\nu|\xi||\epsilon|} \leq \gamma. \tag{V.195}$$

※

V.6.2 Proof of Lemma IV.2

Function $\mathcal{P}(\hat{A})$ being continuous and Θ a compact set, there exist scalars $p_{\min}, p_{\max} > 0$ such that

$$p_{\min} I_n \leq \mathcal{P}(\hat{A}) \leq p_{\max} I_n, \quad \forall \hat{A} \in \Theta. \tag{V.196}$$

Then, define the Lyapunov function

$$V_{\mathrm{m}}(x_{\mathrm{m}}, t) := {x_{\mathrm{m}}}^\top \mathcal{P}(\hat{A}) x_{\mathrm{m}} \tag{V.197}$$

which is positive definite and radially unbounded, and whose derivative along the solutions of (IV.24) is given by:

$$\begin{aligned}
\dot{V}_{\mathrm{m}} &= {x_{\mathrm{m}}}^\top \left(\mathcal{P}(\hat{A})A_{\mathrm{cl}}(\hat{A}) + A_{\mathrm{cl}}(\hat{A})^\top \mathcal{P}(\hat{A})\right) x_{\mathrm{m}} \\
&\quad + {x_{\mathrm{m}}}^\top \left(\frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}\right) x_{\mathrm{m}} + 2{x_{\mathrm{m}}}^\top \mathcal{P}(\hat{A}) B d,
\end{aligned} \tag{V.198}$$

where

$$A_{\mathrm{cl}}(\hat{A}) := \hat{A} - BR^{-1}B^\top \mathcal{P}(\hat{A}) \tag{V.199}$$

and the product $\odot$ is defined as

$$\frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial \mathcal{P}(\hat{A})}{\partial [\hat{A}]_{ij}} [\dot{\hat{A}}]_{ij}, \tag{V.200}$$

with $[\hat{A}]_{ij}$ the i-th row and j-th column entry of matrix $\hat{A}$. Since $\hat{P} = \mathcal{P}(\hat{A})$ solves at each time instant ARE (IV.29), it holds that:

$$A_{\text{cl}}(\hat{A})^\top \mathcal{P}(\hat{A}) + \mathcal{P}(\hat{A}) A_{\text{cl}}(\hat{A}) = \underbrace{-Q - K(\hat{A})^\top R K(\hat{A})}_{=:-\bar{Q}(\hat{A})}, \tag{V.201}$$

where from Assumption IV.1, $Q > 0$ and Θ being compact, $\bar{Q}(\hat{A}) \geq q > 0$ for all $\hat{A} \in \Theta$, with q defined as

$$q \coloneqq \min_{\hat{A}\in\Theta} \lambda_{\min}\left(-Q - \mathcal{P}(\hat{A}) B R B^\top \mathcal{P}(\hat{A})\right), \tag{V.202}$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a matrix. Define $c \coloneqq \max_{i,j\in\{1,\ldots,n\}} \left\{\max_{\hat{A}\in\Theta} \left|\frac{\partial \mathcal{P}(\hat{A})}{\partial [\hat{A}]_{ij}}\right|\right\}$, then we obtain

$$\begin{aligned} \left|\frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}\right| &= \left|\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial \mathcal{P}(\hat{A})}{\partial [\hat{A}]_{ij}} [\dot{\hat{A}}]_{ij}\right| \\ &\leq c \sum_{i=1}^{n}\left(\sum_{j=1}^{n} |[\dot{\hat{A}}]_{ij}|\right) \leq cn \max_{1\leq i\leq n} \sum_{j=1}^{n} |[\dot{\hat{A}}]_{ij}| \leq cn^{\frac{3}{2}} |\dot{\hat{A}}|. \end{aligned} \tag{V.203}$$

By letting $|\dot{\hat{A}}| \leq \gamma_b^\star \coloneqq q/(2cn^{\frac{3}{2}})$, (V.198) becomes

$$\begin{aligned} \dot{V}_{\text{m}} =& -x_{\text{m}}{}^\top \left(\bar{Q}(\hat{A}) - \frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}\right) x_{\text{m}} + 2x_{\text{m}}{}^\top \mathcal{P}(\hat{A}) B d \\ \leq& -(q - c\rho|\dot{\hat{A}}|)|x_{\text{m}}|^2 + 2p_{\max}|B||x_{\text{m}}||d| \\ \leq& -\frac{q}{2}|x_{\text{m}}|\left(|x_{\text{m}}| - \frac{4p_{\max}|B||d|}{q}\right). \end{aligned} \tag{V.204}$$

Therefore,

$$|x_{\text{m}}| \geq \frac{8p_{\max}|B||d|}{q} \implies \dot{V}_{\text{m}} \leq -\frac{q}{4}|x_{\text{m}}|^2, \tag{V.205}$$

which concludes the statement. ※

V.6.3 Proof of Lemma IV.3

At first, to simplify the expressions for the Lyapunov function, we introduce the following vectorized coordinates:

$$\hat{\theta}_{\text{a}} \coloneqq \text{vec}(\hat{K}_{\text{a}}) \in \mathbb{R}^{mn}, \qquad \tilde{\theta}_{\text{a}} \coloneqq \text{vec}(\tilde{K}_{\text{a}}) \in \mathbb{R}^{mn}. \tag{V.206}$$

We rewrite dynamics (IV.49) and (IV.51) by using the vectorized coordinates above defined:

$$\begin{aligned}\dot{e} &= A_{\mathrm{cl}}(\hat{A})e + B\hat{K}_{\mathrm{a}}x - BK_{\mathrm{a}}(\hat{A})x \\ &= A_{\mathrm{cl}}(\hat{A})e + B(x \otimes I_m)^\top \tilde{\theta}_{\mathrm{a}}, \\ \dot{\tilde{\theta}}_{\mathrm{a}} &= -\mu(x \otimes I_m)B^\top \mathcal{P}(\hat{A})e.\end{aligned} \tag{V.207}$$

Consider the Lyapunov function

$$V_e(e, \tilde{\theta}_{\mathrm{a}}, t) := e^\top \mathcal{P}(\hat{A})e + \frac{1}{\mu}|\tilde{\theta}_{\mathrm{a}}|^2, \tag{V.208}$$

which is positive definite and radially unbounded. The time derivative of V_e along the trajectories of (V.207) is given by

$$\begin{aligned}\dot{V}_e = e^\top &\left(\mathcal{P}(\hat{A})A_{\mathrm{cl}}(\hat{A}) + A_{\mathrm{cl}}(\hat{A})^\top \mathcal{P}(\hat{A}) + \frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}\right) e \\ &+ 2e^\top \mathcal{P}(\hat{A})B(x \otimes I_m)^\top \tilde{\theta}_{\mathrm{a}} - \frac{2}{\mu}\tilde{\theta}_{\mathrm{a}}^\top(\mu(x \otimes I_m)B^\top \mathcal{P}(\hat{A})e) \\ = - e^\top &\left(\bar{Q}(\hat{A}) - \frac{\partial \mathcal{P}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}\right) e \leq -\frac{q}{2}|e|^2 \leq 0,\end{aligned} \tag{V.209}$$

where $A_{\mathrm{cl}}(\hat{A})$ is defined in (V.199), $\bar{Q}(\hat{A})$ is given in (V.201), and q is found in (V.202). We have ensured that $(e(t), \tilde{\theta}_{\mathrm{a}}(t))$ is contained for all $t \in [0, t_f)$ in a compact sublevel set of V_e, thus concluding the proof. ※

V.6.4 Proof of Lemma IV.4

For all $\hat{A} \in \Theta$, pair $(A_{\mathrm{cl}}(\hat{A}), B)$ in (IV.48), with $A_{\mathrm{cl}}(\hat{A})$ given in (V.199), is controllable because $(\hat{A}, B)$ is controllable from Assumption (IV.1). Additionally, the origin of system $\dot{x}_{\mathrm{m}} = A_{\mathrm{cl}}(\hat{A})x_{\mathrm{m}}$ is UGES from Lemma IV.2. If $d(t) \in \mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1})$, then by definition (II.55) of $\mathcal{C}_{\mathrm{SR}}(\mathbb{L}^{\mathrm{C}}_{\mathrm{x},>1})$, if $\dot{\hat{A}} = 0$ then $x_{\mathrm{m}}(t)$ is PE, i.e, there exist $T > 0$, $\alpha > 0$ such that

$$\int_t^{t+T} x_{\mathrm{m}}(s)x_{\mathrm{m}}(s)^\top \mathrm{d}s \geq \alpha I_n, \quad \forall t \geq 0. \tag{V.210}$$

By [142, Thm. 6.1], there exist a constant scalar $\eta > 0$, such that, if

$$|A_{\mathrm{cl}}(\hat{A}(s)) - A_{\mathrm{cl}}(\hat{A}(\tau))| \leq \eta, \quad \forall s, \tau \in [t, t+T], \tag{V.211}$$

for all $t \geq 0$, then $x_{\mathrm{m}}(t)$ is PE also when $\dot{\hat{A}} \neq 0$. Recall that $A_{\mathrm{cl}}(\hat{A}) := \hat{A} - BR^{-1}B^\top \mathcal{P}(\hat{A})$ is an analytic function of $\hat{A}$ [175, Thm. 4.1]. Thus, from the mean-value theorem and similar computations to (V.203), we obtain:

$$|A_{\text{cl}}(\hat{A}(s)) - A_{\text{cl}}(\hat{A}(\tau))| = \left|(s-\tau)\frac{\partial A_{\text{cl}}(\hat{A})}{\partial \hat{A}} \odot \dot{\hat{A}}(\varsigma)\right| \leq |s-\tau| c n^{\frac{3}{2}} |\dot{\hat{A}}(\varsigma)|, \tag{V.212}$$

where $\varsigma \in [s,\tau]$, $c := \max_{i,j\in\{1,\dots,n\}}\left\{\max_{\hat{A}\in\Theta}\left|\frac{\partial A_{\text{cl}}(\hat{A})}{\partial[\hat{A}]_{ij}}\right|\right\}$. From the fact that $|\dot{\hat{A}}(\cdot)| \leq \gamma$, we conclude that for $\gamma^{\star}_{PE} := \eta/(Tcn^{\frac{3}{2}})$, if $\gamma \in (0, \gamma^{\star}_{PE}]$, then bound η in (V.211) is enforced and thus $x_{\text{m}}(t)$ is PE. ※

V.6.5 Proof of Lemma IV.5

From Lemma IV.3 and the solutions being forward complete, it holds that the origin $(e, \tilde{\theta}_{\text{a}}) = 0$ of system (V.207) is UGS. Note that the regressor in (IV.51) is given by $x(t)$. Therefore, if $x(t)$ is uniformly PE (u-PE) as in [171, Def. 5], then UGAS and ULES of $(e, \tilde{\theta}_{\text{a}}) = 0$ follows from [171, Thm. 1 and 2]. To prove u-PE of $x(t)$, note that $x(t) = x_{\text{m}}(t) + e(t)$, where $x_{\text{m}}(t)$ is PE from Lemma IV.4. Therefore, we conclude u-PE of $x(t)$ from [171, Prop. 2]. ※

V.6.6 Proof of Lemma IV.6

From Lemma IV.1 and UGES of the $\tilde{\epsilon}$ subsystem, we only need to prove UGES of system (IV.46) with $\tilde{\epsilon} = 0$, which we write here in vectorized coordinates:

$$\dot{\tilde{\theta}}_A = \mathcal{P}_{\text{vec}^{-1}\,\hat{\theta}_{\text{A}}\in C}\left\{-\gamma\bar{B}\frac{(\xi\otimes I_n)(\xi\otimes I_n)^{\top}\tilde{\theta}_A}{1+\nu|\xi||\tilde{A}\xi|}\right\}. \tag{V.213}$$

Since the directions where learning happens are unchanged by the projection operator and by $\bar{B}$, we are interested in studying regressor $\bar{\xi}(t) := \frac{\xi(t)\otimes I_n}{\sqrt{1+\nu|\xi(t)||\tilde{A}(t)\xi(t)|}}$ in order to prove our result. Given a small enough gain γ, it holds from Lemma IV.4 that $x_m(t)$ is PE, while $e(t) \to 0$ exponentially fast from Lemma IV.5. From (IV.47), $\xi(t)$ is a filtered version of the PE signal $x_{\text{m}}(t) + e(t)$, thus $\xi(t)$ is PE [98, Lemma. 4.8.3]. Since all signals are bounded and $(\xi\otimes I_n)(\xi\otimes I_n)^{\top} = (\xi\xi^{\top})\otimes I_n$, PE of $\xi(t)$ implies that

$$\int_t^{t+T}\bar{\xi}(s)\bar{\xi}(s)^{\top}\mathrm{d}s \geq \int_t^{t+T}\frac{(\xi(s)\xi(s)^{\top})\otimes I_n}{1+\xi_M^2\tilde{A}_M}\mathrm{d}s \geq \alpha I_{n^2}, \tag{V.214}$$

for some $T, \alpha > 0$ and all $t \in \mathbb{R}_{\geq 0}$, with $\xi_M := \sup_{t\in\mathbb{R}}|\xi(t)|$, $\tilde{A}_M := \sup_{t\in\mathbb{R}}|\tilde{A}(t)|$, thus $\bar{\xi}(t)$ is PE. From [98, Thm. 8.5.6], we conclude that $\tilde{A} = 0$ is UGES. ※

V.6.7 An expression for the projection operator

For the reader's convenience, we report the expression of the projector $\mathcal{P}_{\hat{A}\in C}\{\tau\}$ as presented in [112, Appendix E] for the case where the set C is a closed ball having center in the nominal parameter $\bar{A}$

and radius $\rho > 0$. In this scenario, C is given by

$$C := \{\hat{A} \in \mathbb{R}^{n\times n} : |\hat{A} - \bar{A}|_F \leq \rho\}, \tag{V.215}$$

where $|\cdot|_F$ is the Frobenius norm. Denoting $\tilde{\theta}_{\mathrm{A}} = \mathrm{vec}(\hat{A} - A)$ and following the same procedure as in [112, Pag. 512], we introduce an arbitrarily small parameter $\sigma > 0$ which is used to define a boundary region around the given ball C. The expression for $\mathcal{P}_{\hat{A}\in C}\{\tau\}$ is then given by

$$\mathcal{P}_{\hat{A}\in C}\{\tau\} = \begin{cases} \tau & \text{(A)} \\ \mathrm{vec}^{-1}\left(\left(I - c(\tilde{\theta}_{\mathrm{A}})\Gamma\frac{\tilde{\theta}_{\mathrm{A}}\tilde{\theta}_{\mathrm{A}}^\top}{|\tilde{\theta}_{\mathrm{A}}|^2}\right)\mathrm{vec}(\tau)\right) & \text{(B)} \end{cases} \tag{V.216}$$

where $\Gamma \in \mathcal{S}_+^{n^2}$ is a tuning gain, $c(\tilde{\theta}_{\mathrm{A}}) = \min(1, \frac{|\tilde{\theta}_{\mathrm{A}}|^2 - \rho^2}{\sigma})$, and the conditions (A), (B) are:

(A) $|\tilde{\theta}_{\mathrm{A}}| < \rho$ or $\tilde{\theta}_{\mathrm{A}}^\top \tau \leq 0$.

(B) $|\tilde{\theta}_{\mathrm{A}}| \in [\rho, \rho + \sigma]$ and $\tilde{\theta}_{\mathrm{A}}^\top \tau > 0$.

※

V.6.8 Proof of Theorem IV.1

Pick $\gamma^\star := \min\{\gamma_b^\star, \gamma_{PE}^\star\} = \gamma_{PE}^\star$, where $\gamma_b^\star$ is from Proposition IV.1 and $\gamma_{PE}^\star$ is the one of Lemma IV.4. Then, if $\gamma \leq \gamma^\star$, the closed-loop solutions are bounded and forward complete. Moreover, x_{m} is PE. The remainder of the proof involves showing the existence of a UGAS attractor using the concept of ω-limit set of a set, see [79, Def. 6.23]. By Lemmas IV.5 and IV.6, from any compact set of initial conditions, it holds that $\hat{A} \to A$, $\tilde{\epsilon} \to 0$, $e \to 0$, $\tilde{K}_{\mathrm{a}} \to 0$ exponentially. Moreover, by Lemma IV.2, the model reference subsystem (IV.48) is ISS with uniformly bounded input $d(t)$, in particular we have that:

$$|x_{\mathrm{m}}| \geq X_{\mathrm{m}} \implies \dot{V}_{\mathrm{m}} \leq -\frac{q}{4}|x_{\mathrm{m}}|^2, \tag{V.217}$$

where X_{m} can be found in (V.205) and depends on $\|d(\cdot)\|_\infty$, and V_{m} in (V.197) is an ISS Lyapunov function for the reference model. Consider the ξ subsystem in (IV.21). It holds that

$$|\xi| \geq \frac{2|x|}{\lambda} \implies \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}|\xi|^2\right) \leq -\frac{\lambda}{2}|\xi|^2. \tag{V.218}$$

Denote $\Xi := \frac{2}{\lambda}X_{\mathrm{m}}$ and define the compact set

$$\begin{aligned} \mathcal{K}_{\mathrm{s}}^\star := \{(w, x, z_s) \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}_{\mathrm{s}} : \hat{A} = A, \tilde{\epsilon} = 0, \\ e = 0, |x_{\mathrm{m}}| \leq X_{\mathrm{m}}, |\xi| \leq \Xi\} \subset \mathcal{W} \times \mathbb{R}^n \times \mathcal{L}_{\mathrm{s}}, \end{aligned} \tag{V.219}$$

where $\mathcal{L}_{\mathrm{s}}$ is the learning set given in (IV.38). Consider a set of initial conditions $\mathcal{K}_s := \mathcal{K}_s^\star + c\mathbb{B}$, with $c > 0$ arbitrary, and note that the solutions are empty if they start outside $\mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}_{\mathrm{s}}$. We now prove that $\mathcal{K}_s^\star$ is uniformly attractive from $\mathcal{K}_s$. By the above-mentioned properties for the subsystems $\tilde{A}, \tilde{\epsilon}, e, \tilde{K}_\mathrm{a}, x_\mathrm{m}$, there exists $T' > 0$ such that, for any $\varepsilon > 0$, it holds that

$$\begin{aligned}
&|\tilde{A}(t)| \le \varepsilon, \quad |\tilde{\epsilon}(t)| \le \varepsilon, \quad |e(t)| \le \min\left(\varepsilon, \frac{\lambda}{2}\frac{\varepsilon}{3}\right) \\
&|x_\mathrm{m}(t)| \le X_\mathrm{m} + \min\left(\varepsilon, \frac{\lambda}{2}\frac{\varepsilon}{3}\right).
\end{aligned} \tag{V.220}$$

for all $t \ge T'$, from which it holds also that

$$\begin{aligned}
\frac{2|x(t)|}{\lambda} &\le \frac{2}{\lambda}(|x_\mathrm{m}(t)| + |e(t)|) \\
&\le \Xi + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \le \Xi + \frac{2}{3}\varepsilon.
\end{aligned} \tag{V.221}$$

Thus, from (V.218), there exists $T \ge T'$ such that $|\xi(t)| \le \Xi + \varepsilon$ for all $t \ge T$. For compactness of notation, denote $\boldsymbol{x}_\mathrm{s} := (w, x, z_\mathrm{s})$. The arguments above have proved that $\mathcal{K}_s^\star$ is uniformly attractive from $\mathcal{K}_s$. Namely, for any $\rho > 0$, there exists $T_\rho \ge 0$ such that $|\phi(t, \boldsymbol{x}_\mathrm{s})|_{\mathcal{K}_\mathrm{s}^\star} \le \rho$, for all $t \ge T_\rho$ and $\boldsymbol{x}_\mathrm{s} \in \mathcal{K}_\mathrm{s}$, where $\phi(t, \boldsymbol{x}_\mathrm{s})$ is the solution at time t of the closed-loop system having initial condition $\boldsymbol{x}_\mathrm{s}$. Denote with $\mathcal{A}_s := \Omega(\mathcal{K}_\mathrm{s})$ the ω-limit set of $\mathcal{K}_\mathrm{s}$. We want to prove that $\mathcal{A}_s \subset \mathcal{K}_\mathrm{s}^\star$. We do it by contradiction, i.e., we suppose that $\mathcal{A}_s \subset \mathcal{K}_\mathrm{s}^\star$ is false. Under this hypothesis, there exists $\bar{\boldsymbol{x}}_\mathrm{s} \in \mathcal{A}_s$ and $\rho > 0$ such that $|\bar{\boldsymbol{x}}_\mathrm{s}|_{\mathcal{K}_\mathrm{s}^\star} \ge 3\rho$. By definition [79, Def. 6.23], the ω-limit set of $\mathcal{K}_\mathrm{s}$ is the set of all points $\boldsymbol{x}_\mathrm{s}$ such that there exist sequences $\boldsymbol{x}_{\mathrm{s},n} \in \mathcal{K}_\mathrm{s}$, $t_n \ge 0$ such that $\lim_{n\to\infty} t_n = \infty$ and $\lim_{n\to\infty} \phi(t_n, \boldsymbol{x}_{\mathrm{s},n}) = \boldsymbol{x}_\mathrm{s}$. Therefore, by definition of limit, there exists $\bar{n} \in \mathbb{N}$ such that

$$|\phi(t_n, \boldsymbol{x}_{\mathrm{s},n}) - \bar{\boldsymbol{x}}_\mathrm{s}| \le \rho, \quad \forall n \ge \bar{n}. \tag{V.222}$$

Pick any subsequence $\boldsymbol{x}_{\mathrm{s},n_i}, t_{n_i}$ such that, for $n_i \ge \bar{n}$, then $t_{n_i} \ge T_\rho$, where T_ρ derives from the uniform attractivity of $\mathcal{K}_\mathrm{s}^\star$ (see above). We have thus proved that, for $n_i \ge \bar{n}$, $|\phi(t_{n_i}, \boldsymbol{x}_{\mathrm{s},n_i}) - \bar{\boldsymbol{x}}_\mathrm{s}| \le \rho$, thus $|\phi(t_{n_i}, \boldsymbol{x}_{\mathrm{s},n_i})|_{\mathcal{K}_\mathrm{s}^\star} \ge 2\rho$, and at the same time $|\phi(t_{n_i}, \boldsymbol{x}_{\mathrm{s},n_i})|_{\mathcal{K}_\mathrm{s}^\star} \le \rho$ by uniform attractivity of $\mathcal{K}_\mathrm{s}^\star$. This is a contradiction, hence necessarily $\mathcal{A}_s \subset \mathcal{K}_\mathrm{s}^\star$. To summarize the previous results, we have thus proved that the solutions are globally bounded and forward complete and

$$\mathcal{A}_s := \Omega(\mathcal{K}_\mathrm{s}) \subset \mathcal{K}_\mathrm{s}^\star \subset \mathrm{Int}(\mathcal{K}_\mathrm{s}) \subset \mathcal{K}_\mathrm{s}. \tag{V.223}$$

By [79, Corollary 7.7], $\mathcal{A}_s = \Omega(\mathcal{K}_\mathrm{s})$ is asymptotically stable, with domain of attraction containing $\mathcal{K}_\mathrm{s}$. Since $\mathcal{K}_\mathrm{s}$ can be chosen arbitrarily large due to Proposition IV.1, we conclude UGAS of $\mathcal{A}_s$. ※

V.6.9 Proof of Theorem IV.2

Consider the reduced-order system with state $\boldsymbol{x}_s := (w, x, z_s)$ and the boundary layer system with state $\hat{P}$. Define the indicator functions

$$\begin{aligned} \omega_s(\boldsymbol{x}_s) &:= \begin{cases} |\boldsymbol{x}_s|_{\mathcal{A}_s} & \boldsymbol{x}_s \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}_s \\ \infty & \text{elsewhere} \end{cases} \\ \omega_f(\boldsymbol{x}_s, \hat{P}) &:= \begin{cases} |\hat{P} - \mathcal{P}(\hat{A})| & (\boldsymbol{x}_s, \hat{P}) \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z} \\ \infty & \text{elsewhere.} \end{cases} \end{aligned} \tag{V.224}$$

By Theorem IV.1, the reduced-order system satisfies

$$\omega_s(\boldsymbol{x}_s(t)) \le \beta_s(\omega_s(\boldsymbol{x}_s(0)), t), \tag{V.225}$$

where β_s is a class $\mathcal{KL}$ function. Moreover, by the DRE properties [47, Thm. 4], the boundary-layer system $d\hat{P}/d\tau(\tau) = \mathcal{R}(\hat{P}(\tau), \hat{A})$, with $\hat{A} \in \Theta$ constant and $\tau := gt$, satisfies

$$\omega_f(\boldsymbol{x}_s, \hat{P}(\tau)) \le \beta_f(\omega_f(\boldsymbol{x}_s, \hat{P}(0)), \tau), \tag{V.226}$$

where β_f is a class $\mathcal{KL}$ function. From [206, Thm. 1] (Assumptions 1, 3, 4, 7, 8 can be verified), from any compact set of initial conditions $\mathcal{K} \subset \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z}$ and for any $\delta > 0$, there exists $g^\star > 0$ such that, for all $g \ge g^\star$, the solutions are forward complete and satisfy:

$$\begin{aligned} \omega_s(\boldsymbol{x}_s(t)) &\le \beta_s(\omega_s(\boldsymbol{x}_s(0)), t) + \delta \\ \omega_f(\boldsymbol{x}_s(t), \hat{P}(t)) &\le \beta_f(\omega_f(\boldsymbol{x}_s(0), \hat{P}(0)), gt) + \delta. \end{aligned} \tag{V.227}$$

In particular, choose $\mathcal{K} := \mathcal{A}_s \times P^\star + c\mathbb{B}$, with $c > 0$ arbitrary. Reference model dynamics (IV.24) can be rewritten as:

$$\begin{aligned} \dot{x}_m =& (\hat{A} - BR^{-1}B^\top \hat{P})x_m + Bd \\ =& (A - BR^{-1}B^\top P^\star)x_m + Bd + \\ & (\tilde{A} - BR^{-1}B^\top(\hat{P} - P^\star))x_m \\ =& (A - BR^{-1}B^\top P^\star)x_m + Bd + \\ & (\tilde{A} - BR^{-1}B^\top(\hat{P} - \mathcal{P}(\hat{A}) + \mathcal{P}(\hat{A}) - \mathcal{P}(A)))x_m. \end{aligned} \tag{V.228}$$

Notice that x_m is PE if $\hat{P} = \mathcal{P}(\hat{A})$ and $\tilde{A} = 0$. Furthermore, since $\mathcal{P}(\hat{A})$ is an analytic function of $\hat{A}$, x_m is PE by [191, Lemma 6.1.2] if $|\hat{P} - \mathcal{P}(\hat{A})|$ and $|\tilde{A}|$ are sufficiently small, because the solutions of (V.228) are sufficiently close to those with $\hat{P} = \mathcal{P}(\hat{A})$ and $\tilde{A} = 0$. Moreover, also x and ξ are PE if x_m is PE and $|e|$ is sufficiently small. Choose $\delta > 0$ such that the conditions $\omega_s(\boldsymbol{x}_s) \le 2\delta$ and $\omega_f(\boldsymbol{x}_s, \hat{P}) \le 2\delta$ imply that x_m, x, and ξ are PE. Then, pick $g \ge g^\star$, where $g^\star$ is obtained from the considered $\mathcal{K}$ and δ. From (V.227), the closed-loop solutions converge in finite time T to a compact set satisfying $\omega_s(\boldsymbol{x}_s) \leq 2\delta$ and $\omega_f(\boldsymbol{x}_s, \hat{P}) \leq 2\delta$. Then, for $t \geq T$, $\tilde{\epsilon} \to 0$, $\hat{A} \to A$ exponentially from Lemma IV.6 since ξ is PE. From the local exponential stability of the DRE [47, Thm. 4], it follows that $\hat{P} \to P^\star$ exponentially. By Lemma IV.5 and $\hat{P} \to \mathcal{P}(\hat{A})$, we conclude that $e \to 0$, $\hat{K}_a \to 0$ exponentially. As a consequence, the same arguments of Theorem IV.1 (omitted here to avoid repetition) can be used to show that the compact set

$$\begin{aligned}\mathcal{K}^\star &:= \{(\boldsymbol{x}_s, \hat{P}) \in \mathcal{W} \times \mathbb{R}^n \times \mathcal{Z} : \hat{A} = A, \tilde{\epsilon} = 0, \\ &\qquad\qquad e = 0, |x_m| \leq X_m, |\xi| \leq \Xi, \hat{P} = P^\star\} \\ &= \mathcal{K}_s^\star \times P^\star,\end{aligned} \tag{V.229}$$

is uniformly attractive from $\mathcal{K}$, with $\mathcal{K}_s^\star$ given in (V.219) . The same steps as in Theorem IV.1 allow to prove that $\mathcal{A} := \Omega(\mathcal{K}) \subset \mathcal{K}^\star \subset \mathrm{Int}(\mathcal{K}) \subset \mathcal{K}$, thus $\mathcal{A}$ is uniformly asymptotically stable with domain of attraction containing $\mathcal{K}$, and since $\mathcal{K}$ can be chosen arbitrarily large we can conclude semiglobal uniform asymptotic stability of $\mathcal{A}$. Finally, we want to prove that $\mathcal{A} = \mathcal{A}_s \times P^\star$, where $\mathcal{A}_s = \Omega(\mathcal{K}_s^\star)$. In $\mathcal{A}$, it holds that $\hat{P} = P^\star$, $\hat{A} = A$ and $\tilde{\epsilon} = 0$, from which it holds that $\hat{P} = \mathcal{P}(\hat{A}) = \mathcal{P}(A) = P^\star$ for all points in this set. For this reason, in $\mathcal{A}$, the vector field of Algorithm 5 coincides with the vector field of the reduced-order system with $\hat{A} = A$ and $\tilde{\epsilon} = 0$. Since the vector fields coincide, we have that in this set solutions $\boldsymbol{x}(t)$ of Algorithm 5 can be written as $\boldsymbol{x}(t) = \boldsymbol{x}_s(t) \times P^\star$, where $\boldsymbol{x}_s(t)$ is the solution of the reduced-order system having the same initial conditions. From (V.229) and $\mathcal{A} \subset \mathcal{K}^\star$, it follows that $\mathcal{A} = \Omega(\mathcal{K}) = \Omega(\mathcal{K}^\star) = \Omega(\mathcal{K}_s^\star \times P^\star) = \mathcal{A}' \times P^\star$. Since for the slow states of Algorithm 1 the solutions coincide with those of the reduced-order system, it follows that $\mathcal{A}' = \mathcal{A}_s$. As a consequence, $\mathcal{A} = \mathcal{A}_s \times P^\star$. ※

V.7 Proofs for Chapter V

V.7.1 Proof of Lemma V.1

Consider problem (V.28). The (approximate) Newton method for constrained optimization applied on (V.28) updates the tentative solution ξ^k by applying, for all iterations k, the iterative rule, cf. [66, Algorithm 4.1],

$$\begin{bmatrix}\xi^{k+1} \\ \lambda^{k+1}\end{bmatrix} = \begin{bmatrix}\xi^{k} \\ 0\end{bmatrix} - \underbrace{\begin{bmatrix}\nabla^2 \ell(\xi^k) & \nabla h(\xi^k) \\ \nabla h(\xi^k)^\top & 0\end{bmatrix}^{-1}}_{Z(\xi^k)} \begin{bmatrix}\nabla \ell(\xi^k) \\ h(\xi^k),\end{bmatrix} \tag{V.230}$$

where $\lambda^k \in \mathbb{R}^{s_x}$ are the lagrangian multipliers associated to the equality constraint. This means the update law for ξ^k is given by

$$\begin{aligned} \zeta^k &= \underbrace{Z_{11}(\xi^k)\nabla\ell(\eta^k)}_{:=\zeta_1^k} + \underbrace{Z_{12}(\xi^k)h(\xi^k)}_{:=\zeta_2^k} \\ \xi^{k+1} &= \xi^k - \zeta^k. \end{aligned} \tag{V.231}$$

where Z_{ij} denotes the ij-th block of matrix Z in (V.230). It can be proved [66, Thm. 4.2] that, under Assumptions, V.2 and V.1, a local solution $\eta^\star$ to (V.28) is LES for dynamics (V.231). Notice that in general the tentative solution ξ^k is not a trajectory of the system. It can be shown that the term ζ_1^k in (V.231) is the result of the following optimization problem

$$\begin{aligned} \zeta_1^k = \operatorname*{argmin}_{\zeta} \quad & \zeta^\top \nabla^2\ell(\xi^k)\zeta + \nabla\ell(\xi^k)\zeta \\ \text{s.t.} \quad & \nabla h(\xi^k)^\top \zeta = 0, \end{aligned} \tag{V.232}$$

namely, it satisfies by construction $\zeta_1^k \in T_{\eta^k}\mathcal{T}$ at all iterations. The second term, ζ_2^k, takes into account the constraint violation, and it is zero when $\xi^k \in \mathcal{T}$.
We show now that, introducing a projection, the update law (V.231) does not lose the convergence properties of the algorithm. Consider the update law

$$\begin{aligned} \zeta^k &= Z_{11}(\xi^k)\nabla\ell(\xi^k) + Z_{12}(\xi^k)h(\xi^k) \\ \xi^{k+1} &= \mathcal{P}(\xi^k - \zeta^k). \end{aligned} \tag{V.233}$$

Thanks to the projection $\mathcal{P}$ in the update, we know it holds $\xi^k \in \mathcal{T}$ for all $k \in \mathbb{N}$. For this reason, we re-write (V.233) as

$$\begin{aligned} \zeta^k &= Z_{11}(\eta^k)\nabla\ell(\eta^k) + Z_{12}(\eta^k)h(\eta^k) \\ \eta^{k+1} &= \mathcal{P}(\eta^k - \zeta^k), \end{aligned} \tag{V.234}$$

where we only highlithted by using η^k the fact that iteration (V.234), unlike (V.231), produces only system trajectories. Under Assuption V.3, we can expand the update in Taylor series and we obtain

$$\begin{aligned} \zeta^k &= Z_{11}(\eta^k)\nabla\ell(\eta^k) + Z_{12}(\eta^k)h(\eta^k) \\ \eta^{k+1} &= \mathcal{P}(\eta^k) - \nabla\mathcal{P}(\eta^k)^\top\zeta^k + o(\zeta^k) \\ &= \eta^k - \nabla\mathcal{P}(\eta^k)^\top\zeta^k + o(\zeta^k), \\ &= \eta^k - \nabla\mathcal{P}(\eta^k)^\top(\zeta_1^k + \zeta_2^k) + o(\zeta^k), \end{aligned} \tag{V.235}$$

where we exploited the fact that $\mathcal{P}(\eta) = \eta$ for all $\eta \in \mathcal{T}$. It can be shown (see [90] for more insight) that, if $\eta \in \mathcal{T}$ and $\zeta \in T_\eta \mathcal{T}$, it holds

$$\nabla \mathcal{P}(\eta)^\top \zeta = \zeta, \tag{V.236}$$

namely, $\nabla \mathcal{P}^\top : \mathbb{R}^s \to T_\eta \mathcal{T}$ is itself a projection into the space tangent to the trajectory manifold at η. Since $\zeta_1^k \in T_{\eta^k}\mathcal{T}$ by construction (being the solution of (V.232)), and since $\zeta_2^k = Z_{12}(\eta^k)h(\eta^k) = 0$ (since $\eta \in \mathcal{T} \implies h(\eta) = 0$), we re-write (V.235) as

$$\begin{aligned} \zeta^k &= Z_{11}(\eta^k)\nabla\ell(\eta^k) + Z_{12}(\eta^k)h(\eta^k) \\ \eta^{k+1} &= \eta^k - \zeta^k + o(\zeta^k), \end{aligned} \tag{V.237}$$

Notice that dynamics (V.231) and (V.237) differ only in the little-o term. By definition of $o(\zeta)$, for all $\epsilon > 0$ we can find $\delta_\epsilon > 0$ such that

$$|\zeta| \le \delta_\epsilon \implies |o(\zeta)| \le \epsilon|\zeta|. \tag{V.238}$$

Furthermore, holding SOSC for $\eta^\star$ and being f, ℓ smooth, it can be shown that for any ball $\mathbb{B}_r(\eta^\star)$, there exists $k(r) > 0$ such that $|\zeta| \le k(r)|\eta - \eta^\star|$ for all $\eta \in \mathbb{B}_r(\eta^\star)$. Overall, we have that for any $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that

$$|\eta - \eta^\star| \le \delta_\epsilon \implies |o(\zeta(\eta))| \le \epsilon|\eta - \eta^\star|. \tag{V.239}$$

In [66, Thm. 4.2], it is shown that by choosing the Lyapunov function $V(\xi) = |\xi - \eta^\star|$, the optimal solution $\eta^\star$ is Locally Exponentially Stable for dynamics (V.231), and for some $s < 1, r > 0$ it holds

$$|\xi^{k+1} - \eta^\star| \le s|\xi^k - \eta^\star| \tag{V.240}$$

for all $\xi \in \mathbb{B}_r(\eta^\star)$. By using the same Lyapunov function for dynamics (V.237), we obtain

$$\begin{aligned} |\eta^{k+1} - \eta^\star| =& |\eta^k - \zeta^k + o(\zeta) - \eta^\star| \\ \le& s|\eta^k - \eta^\star| + |o(\zeta)|. \end{aligned} \tag{V.241}$$

Recalling (V.239), picking $\epsilon : s + \epsilon < 1$, we have that there exists $\delta_\epsilon > 0$ such that

$$\begin{aligned} |\eta^{k+1} - \eta^\star| =& |\xi^{k+1} + o(\zeta) - \eta^\star| \\ \le& (s + \epsilon)|\eta^k - \eta^\star| \end{aligned} \tag{V.242}$$

for all $\eta^k \in \mathbb{B}_{\delta_\epsilon}(\eta^\star) \cap \mathbb{B}_r(\eta^\star)$, which proves LES of $\eta^\star$ under dynamics (V.237). At last, notice that since $\eta^k \in \mathcal{T}$ for all k, we have $h(\eta^k) = 0$ and thus the update (V.237) can be rewritten as

$$\begin{aligned} \zeta^k &= Z_{11}(\eta^k)\nabla\ell(\eta^k) \\ \eta^{k+1} &= \mathcal{P}(\eta^k - \zeta^k), \end{aligned} \tag{V.243}$$

which is exactly dynamics (V.31). We can thus conclude Local Exponential Stability of $\eta^\star$ under dynamics (V.31) [89, Thm. 13.11]. The proof follows. ※

V.7.2 Proof of Lemma V.2

The proof is obtained by applying [56, Theorem 2.7]. First, notice that by taking the squares of (V.242), it is possible to use $V(\eta) := |\eta - \eta^\star|^2$ as C^1 Lyapunov function for the unperturbed dynamics (V.31). This means that Assumption B1 of [56, Theorem 2.7] is satisfied since the equilibrium $\eta^\star$ of unperturbed system is locally asymptotically stable with C^1 Lyapunov function $V(\eta) = |\eta - \eta^\star|^2$. Next, we rewrite the perturbed system (V.32) as

$$\begin{aligned} \eta^{k+1} &= \mathcal{P}(\eta^k + \zeta^k + \Delta\zeta^k) \\ &= \underbrace{\mathcal{P}(\eta^k + \zeta^k)}_{\text{unperturbed}} + \underbrace{\nabla\mathcal{P}^\top_{\eta^k+\zeta^k}\Delta\zeta^k + o\left((\Delta\zeta^k)^2\right)}_{\text{perturbation } d(\eta^k,\Delta\zeta^k)}, \end{aligned} \tag{V.244}$$

for $\Delta\zeta \to 0$, to explicit the disturbance as an additive term. Finally, to satisfy assumption B2 of [56, Theorem 2.7], we need a bound for the perturbation term. Being $d(\eta, \Delta\zeta)$ differentiable in its arguments and such that $d(\eta, 0) = 0$ for all $\eta \in \mathbb{R}^s$, for every ball $\mathbb{B}_r \subset \mathbb{R}^{s\times s}$ there exists $\bar{d}(r) > 0$ such that

$$|d(\eta^k, \Delta\zeta^k)| \le \bar{d}(r)|\Delta\zeta^k| \quad \forall\Delta\zeta^k \in \mathbb{B}_r, \eta^k \in \mathbb{R}^s. \tag{V.245}$$

Denote now as $B \subset \mathbb{R}^s$ the basin of attraction of the unperturbed dynamics (V.31). Then, by [56, Theorem 2.7], if $|\Delta\zeta^k| \le \delta_\zeta$ for all $k \in \mathbb{N}$, there exists $\bar{\delta}_\zeta > 0$ such that, if $\eta^0 \in B$ and $|\Delta\zeta^k| \le \delta_\zeta < \bar{\delta}_\zeta$ for all k, then there exists $K > 0$, class $\mathcal{KL}$ function ϕ and class $\mathcal{K}$ function b such that

$$\begin{aligned} |\Delta\eta^k| &\le \phi(|x_0|, k) && \forall k < K \\ |\Delta\eta^k| &\le b(\delta_\zeta) && \forall k \ge K. \end{aligned} \tag{V.246}$$

The proof follows. ※

V.7.3 Proof of Lemma V.3

Notice that each perturbation $\hat{\eta}^i = (\hat{\boldsymbol{x}}^i, \hat{\boldsymbol{u}}^i)$ is obtained via the closed-loop (V.21), which can be seen as a repeated composition of the functions π (controller), f (dynamics) and the dither injection. We thus write $\hat{\eta}^i = \hat{\eta}^i(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u)$, where the dependence on $\eta = (\boldsymbol{x}, \boldsymbol{u})$ takes into account the fact that π is tracking the current trajectory η. By composition of C^1 functions, under Assumptions V.1, V.3, we have then that for all iterations $i = 1, \dots, L$, $\hat{\eta}^i(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u)$ is a C^1 function of $\eta, \boldsymbol{d}_x, \boldsymbol{d}_u$. Next, notice that the data batches built as in (V.24) are stacks of differences between components of $\hat{\eta}^i$ and components of η, so the function Δ_{XU} can be seen as a smooth composition between functions $\hat{\eta}^i(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u)$, with $i = 1, \dots, L$, and η, which means $\Delta_{XU}(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u)$ is C^1. To conclude the proof, if all dithers $\boldsymbol{d}_u, \boldsymbol{d}_x$ are zero, then by Assumption V.3 we have that all perturbed trajectories $\hat{\eta}^i$ coincide with the current one

η, namely, $\hat{\eta}^i(\eta, 0, 0) = \eta$ for all $i = 1, \dots, L$. In turn, this means that the difference $\hat{\eta}^i(\eta, 0, 0) - \eta$ used to build Δ_{XU} is zero, and so $\Delta_{XU}(\eta, 0, 0) = 0$. The proof follows. ※

V.7.4 Proof of Lemma V.4

The proof goes through three main steps. First, we find a closed-form expression for the errors (V.38) on the Jacobians at each t. Second, we show that this expression is a differentiable function of the entries of $\Delta X_t, \Delta U_t$. Finally, we define a local linear bound on the norm of Δ_{AB}.

I) Closed-form expression for ΔA_t and ΔB_t. Define

$$\begin{aligned} \Delta x_t^i &= \hat{x}_t^i - x_t, \quad \Delta u_t^i = \hat{u}_t^i - u_t, \\ \Delta x_t^{+,i} &= f(\hat{x}_t^i, \hat{u}_t^i) - f(x_t, u_t) \end{aligned} \tag{V.247}$$

where the apex i denotes the i-th perturbation of the current trajectory $\eta = (\boldsymbol{x}, \boldsymbol{u})$. By expanding in Taylor series the dynamics f about x_t, u_t, we obtain:

$$\Delta x_t^{+,i} = A_t \Delta x_t^i + B_t \Delta u_t^i + o_\eta(\Delta x_t^i, \Delta u_t^i), \tag{V.248}$$

for $\Delta x_t^i \to 0, \Delta u_t^i \to 0$. The pedex η in $o_\eta(\cdot)$ highlights the fact that the little-o term depends on the trajectory η. We then build $\Delta X_t, \Delta U_t, \Delta X_t^+$ as in (V.24). It holds:

$$\Delta X_t^+ = \begin{bmatrix} A_t & B_t \end{bmatrix} \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix} + o_\eta\left(\begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}\right), \tag{V.249}$$

for $\Delta x_t^i \to 0, \Delta u_t^i \to 0$ for all $i = 1, \dots, L$, from which, under Assumption V.5, we obtain:

$$\begin{bmatrix} A_t & B_t \end{bmatrix} = \left(\Delta X_t^+ - o_\eta\left(\begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}\right)\right) \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\dagger. \tag{V.250}$$

Recall that we estimate the Jacobians with

$$\begin{bmatrix} \hat{A}_t & \hat{B}_t \end{bmatrix} = X_t^+ \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\dagger. \tag{V.251}$$

We can then subtract (V.250) from (V.251) to obtain

$$\begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} = o_\eta\left(\begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}\right) \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\dagger, \tag{V.252}$$

for all $t = 0, \dots, T-1$.

II) Continuous differentiability of $[\Delta A_t, \Delta B_t]$ in (V.252) Notice that $o_\eta(\cdot)$ in (V.252) must be a continuously differentiable function of $\Delta X_t, \Delta U_t$ by Assumption V.1 (since it is a stack of the $o_\eta(\cdot)$ in (V.248), which in turn must be C^1 in their argument being the remainder of a first-order approximation of a C^2 function). Furthermore, also the pseudoinverse in (V.252) is a C^1 function of the the entries of $\Delta X_t, \Delta U_t$, since, for $(\Delta X_t, \Delta U_t) \in \mathcal{F}_M$, it can be calculated as

$$\begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\dagger = \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\top \left(\begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix} \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}^\top \right)^{-1}, \tag{V.253}$$

and this is a product of C^1 functions of the entries of $\Delta X_t, \Delta U_t$ in the considered domain. This proves that, given $(\Delta X_t, \Delta U_t) \in \mathcal{F}_M$, for all $t = 1, \ldots, T-1$ the error matrix $[\Delta A_t, \Delta B_t]$ in (V.252) is a C^1 function of $\Delta X_t, \Delta U_t$. In turn, this implies that the function $\Delta_{AB}(\eta, \boldsymbol{\Delta X}, \boldsymbol{\Delta U})$ is a C^1 function of $\boldsymbol{\Delta X}, \boldsymbol{\Delta U}$, since it is the stack of all $[\Delta A_t, \Delta B_t]$. Furthermore, it is also C^1 in η since the little-o term $o_\eta(\cdot)$ is C^1 with respect to η by Assumption V.1.

III) Local linear bound on $|\Delta_{AB}|$ Denote, for simplicity,

$$\Delta_t := \begin{bmatrix} \Delta X_t \\ \Delta U_t \end{bmatrix}. \tag{V.254}$$

By substituting (V.253) in (V.252), we obtain

$$\begin{aligned} \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} &= o_\eta(\Delta_t)\Delta_t^\top (\Delta_t \Delta_t^\top)^{-1} \\ &= o_\eta(\Delta_t \Delta_t^\top)(\Delta_t \Delta_t^\top)^{-1}, \end{aligned} \tag{V.255}$$

and by taking the norms,

$$\begin{aligned} \left| \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} \right| &\leq |o_\eta(\Delta_t \Delta_t^\top)||(\Delta_t \Delta_t^\top)^{-1}| \\ &\leq |o_\eta(\Delta_t \Delta_t^\top)||(\Delta_t \Delta_t^\top)|^{-1} \kappa(\Delta_t \Delta_t^\top) \\ &\leq |o_\eta(\Delta_t \Delta_t^\top)||(\Delta_t \Delta_t^\top)|^{-1} M \end{aligned} \tag{V.256}$$

Notice that by definition, $o_\eta(\cdot)$ goes to zero faster than its argument, i.e., for any $\epsilon > 0$ we can find $\delta > 0$ such that, for all $t = 0, \ldots, T-1$, it holds

$$|\Delta_t \Delta_t^\top| \leq \delta \implies |o(\Delta_t \Delta_t^\top)||\Delta_t \Delta_t^\top|^{-1} \leq \epsilon. \tag{V.257}$$

Being $[\Delta A_t \ \Delta B_t]$ continuously differentiable in $\Delta X_t, \Delta U_t$ in the considered domain, we know that for any $\eta \in \mathcal{T}$ and any bounded set $\mathcal{K} \subset \mathcal{F}_M$, for all $\Delta_t := (\Delta X_t, \Delta U_t)$, $\Delta_t' := (\Delta X_t', \Delta U_t') \in \mathcal{K}$ there exist $k(\eta, \mathcal{K}) > 0$ such that

$$\begin{aligned} \left\| \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} - \begin{bmatrix} \Delta A'_t & \Delta B'_t \end{bmatrix} \right\| &\leq k(\eta, \mathcal{K}) \left| \Delta_t - \Delta'_t \right| \\ \left\| \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} \right\| \leq \left\| \begin{bmatrix} \Delta A'_t & \Delta B'_t \end{bmatrix} \right\| &+ k(\eta, \mathcal{K})(|\Delta_t| + |\Delta'_t|) \end{aligned} \tag{V.258}$$

Using (V.257), given any $\epsilon > 0$, there exists $\delta > 0$ for which, by picking $|\Delta'| \leq \delta$, we can to ensure $|[\Delta A'_t \ \Delta B'_t]| \leq \epsilon$, from which

$$\left\| \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} \right\| \leq k(\eta, \mathcal{K})|\Delta_t| + k(\eta, \mathcal{K})\delta + \epsilon. \tag{V.259}$$

Since ϵ can be picked arbitrarily small, we obtain it must hold

$$\left\| \begin{bmatrix} \Delta A_t & \Delta B_t \end{bmatrix} \right\| \leq k(\eta, \mathcal{K})|\Delta_t|, \tag{V.260}$$

and since this holds for all $t = 0, \ldots, T-1$, we have

$$|\Delta_{AB}(\eta, \mathbf{\Delta X}, \mathbf{\Delta U})| \leq k(\eta, \mathcal{K})T(|\mathbf{\Delta X}| + |\mathbf{\Delta U}|). \tag{V.261}$$

The proof follows. ※

V.7.5 Proof of Lemma V.5

Consider the exact problem (V.13). We reformulate it as

$$\begin{aligned} \min_{\zeta \in \mathbb{R}^s} \quad & \zeta^\top \nabla^2 \ell(\eta) \zeta + \nabla \ell(\eta)^\top \zeta \\ \text{s.t.} \quad & H(\eta)\zeta = 0, \end{aligned} \tag{V.262}$$

with $\zeta = (\mathbf{\Delta x}, \mathbf{\Delta u})$ and

$$\begin{aligned} H(\eta) &:= \begin{bmatrix} H_x(\eta) & H_u(\eta) \end{bmatrix}, \\ H_x(\eta) &:= \begin{bmatrix} I & 0 & 0 & 0 \\ -A_0 & I & 0 & 0 \\ 0 & \ldots & \ldots & 0 \\ 0 & 0 & -A_{T-1} & I \end{bmatrix} \in \mathbb{R}^{s_x \times s_x} \\ H_u(\eta) &:= \begin{bmatrix} 0 & 0 & 0 \\ -B_0 & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & -B_{T-1} \end{bmatrix} \in \mathbb{R}^{s_x \times s_u}. \end{aligned} \tag{V.263}$$

The estimated problem (V.18) differs from this only in the constraints, i.e., it can be written as

$$\begin{aligned} \min_{\zeta \in \mathbb{R}^s} \quad & \zeta^\top \nabla^2 \ell(\eta) \zeta + \nabla \ell(\eta)^\top \zeta \\ \text{s.t.} \quad & \hat{H}(\eta, \mathbf{\Delta A}, \mathbf{\Delta B}) \zeta = 0, \end{aligned} \tag{V.264}$$

where $\hat{H}(\eta, \mathbf{\Delta A}, \mathbf{\Delta B}) = H(\eta) + \tilde{H}(\mathbf{\Delta A}, \mathbf{\Delta B})$ given by

$$\begin{aligned} \tilde{H}(\mathbf{\Delta A}, \mathbf{\Delta B}) &:= \begin{bmatrix} \tilde{H}_x(\mathbf{\Delta A}, \mathbf{\Delta B}) & \tilde{H}_u(\mathbf{\Delta A}, \mathbf{\Delta B}) \end{bmatrix}, \\ \tilde{H}_x(\mathbf{\Delta A}, \mathbf{\Delta B}) &:= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\Delta A_0 & 0 & 0 & 0 \\ 0 & \dots & \dots & 0 \\ 0 & 0 & -\Delta A_{T-1} & 0 \end{bmatrix} \in \mathbb{R}^{s_x \times s_x} \\ \tilde{H}_u(\mathbf{\Delta A}, \mathbf{\Delta B}) &:= \begin{bmatrix} 0 & 0 & 0 \\ -\Delta B_0 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & -\Delta B_{T-1} \end{bmatrix} \in \mathbb{R}^{s_x \times s_u}. \end{aligned} \tag{V.265}$$

By Assumption (V.2), (V.262) is a strictly convex program and has a unique minimizer $\zeta^\star(\eta)$. Furthermore, given the structure of $H(\eta)$ LICQ hold for any $\eta \in \mathbb{R}^s$. This means that KKT conditions hold for the solution $\zeta^\star(\eta)$ of problem (V.262) [66, Thm. 3.14]. By [66, Thm. 3.18], having only equality constraints, also SOSC holds for $\zeta^\star(\eta)$. We can apply [66, Thm. 3.19] and obtain that there exists a $\delta_{AB}(\eta) > 0$ such that, for all $|\tilde{H}(\mathbf{\Delta A}, \mathbf{\Delta B})| \leq \delta_{AB}(\eta)$, there exists a unique solution $\hat{\zeta}^\star(\eta, \mathbf{\Delta A}, \mathbf{\Delta B})$ to (V.264) and it depends differentiably on $\tilde{H}(\eta, \mathbf{\Delta A}, \mathbf{\Delta B})$. Furthermore, $\delta_{AB}(\eta)$ must depend continuously on η, since $\nabla^2 \ell(\eta), \nabla \ell(\eta), H(\eta)$ describing the QP (V.262) are all continuous in η by Assumption V.1.

Next, as $H(\mathbf{\Delta A}, \mathbf{\Delta B})$ is linear in $\eta, \mathbf{\Delta A}, \mathbf{\Delta B}$, $\hat{\zeta}^\star(\eta, \mathbf{\Delta A}, \mathbf{\Delta B})$ is differentiable also in $\eta, \mathbf{\Delta A}, \mathbf{\Delta B}$. This means that $\zeta^\star(\eta) - \hat{\zeta}^\star(\eta, \mathbf{\Delta A}, \mathbf{\Delta B})$ is smooth in its arguments.

To conclude the statement, notice that if $\mathbf{\Delta A}, \mathbf{\Delta B} = 0$ then $\tilde{H}(\mathbf{\Delta A}, \mathbf{\Delta B}) = 0$, and thus, since problems (V.262) and (V.264) become identical, $\zeta^\star(\eta) - \hat{\zeta}^\star(\eta, 0, 0) = 0$. The proof follows. ※

V.7.6 Proof of Theorem V.1

We are now ready to prove the main result of this chapter. The proof goes through three main steps. First, we show that for sufficiently small exploration dithers problems (V.18) are well-posed and admit a solution, hence Δ_ζ smoothly depends on $\boldsymbol{d}_x, \boldsymbol{d}_u$. Second, we leverage on the smoothness properties of functions $\Delta_{XU}, \Delta_{AB}, \Delta_\zeta$ to obtain a linear bound on the maximum error on the descent direction. Third, we show how to use this bound to guarantee convergence of Algorithm 7.

I) Δ_ζ as smooth function of $\boldsymbol{d}_x, \boldsymbol{d}_u$ Consider an isolated local minima $\eta^\star$ of (V.8) and $\delta_{AB}(\eta^\star) > 0$, where $\delta_{AB}(\cdot)$ is given in Lemma V.5. By continuity and positivity of $\delta_{AB}(\cdot)$, we have that for any $\sigma > 0$ there exists

$$\delta_{AB} := \inf_{\eta \in \mathbb{B}_\sigma(\eta^\star) \cap \mathcal{T}} \delta_{AB}(\eta) > 0. \tag{V.266}$$

So, by Lemma V.5 it holds that for all $\eta \in \mathbb{B}_\sigma(\eta^\star) \cap \mathcal{T}$, if $|(\boldsymbol{\Delta A}, \boldsymbol{\Delta B})| \leq \delta_{AB}$, then $\Delta_\zeta(\eta, \boldsymbol{\Delta A}, \boldsymbol{\Delta B})$ is a C^1 function of its arguments.
In the following passages, we restrict our analysis only to those $\eta, \boldsymbol{d}_x, \boldsymbol{d}_u$ such that

$$(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u) \in \Delta_{XU}^{-1}(\mathbb{R}^s \times \mathcal{F}_M^T) \tag{V.267}$$

to guarantee that the composition of the functions Δ_{XU} and Δ_{AB} is well posed (this will be later guaranteed by Assumption V.5).
By applying Lemmas V.3 and V.4, it follows that there exists $\bar{\delta}_x, \bar{\delta}_u > 0$ such that, if $\delta_x \in (0, \bar{\delta}_x)$ and $\delta_u \in (0, \bar{\delta}_u)$ then $|(\boldsymbol{\Delta A}, \boldsymbol{\Delta B})| \leq \delta_{AB}$, where δ_x and δ_u are user-defined parameters of Algorithm 7, cf. Assumption V.4.

II) Linear bound for the maximum error on the descent direction Composing Lemmas V.3, V.4 and V.5, we have that (by continuous differentiability in a bounded set) if $\delta_x \in (0, \bar{\delta}_x)$ and $\delta_u \in (0, \bar{\delta}_u)$, then there exists $g = g(\sigma, \eta^\star) > 0$ such that

$$|\Delta_\zeta(\eta, \Delta_{AB}(\eta, \Delta_{XU}(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u))| \leq g(|\boldsymbol{d}_x| + |\boldsymbol{d}_u|) \tag{V.268}$$

for all $\eta \in \mathbb{B}_\sigma(\eta^\star) \cap \mathcal{T}$, from which we can define

$$\begin{aligned}\delta_\zeta(\delta_x, \delta_u) := & \sup_{\substack{\eta \in \mathbb{B}_\sigma(\eta^\star) \\ \eta \in \mathcal{T} \\ \boldsymbol{d}_x \in \mathbb{B}_{\delta_x} \\ \boldsymbol{d}_u \in \mathbb{B}_{\delta_u}}} |\Delta_\zeta(\eta, \Delta_{AB}(\eta, \Delta_{XU}(\eta, \boldsymbol{d}_x, \boldsymbol{d}_u))| \\ \leq & \; g(\delta_x + \delta_u). \end{aligned} \tag{V.269}$$

Notice that the bound in (V.269) is structural, i.e., it depends only on the algorithm parameters δ_x, δ_u and on $\eta^\star$, not on the algorithm iteration.

III) Convergence guarantees for Algorithm 7 By writing Alg. 7 (DATA-DRIVEN PRONTO) as a perturbed version of Alg. 6 (PRONTO), hence as the dynamical system (V.32) where the perturbations are introduced by the estimation of the dynamics Jacobians, we apply Lemma V.2 from which we define the quantities $\bar{\delta}_\zeta > 0$, $\phi(\cdot, \cdot)$, $b(\cdot)$. If $|\Delta\zeta^k| \leq \bar{\delta}_\zeta$ is satisfied at each iteration, then the algorithm evolution is constrained to the ball of radius $\phi(|\eta_0 - \eta^\star|, 0)$. Let $\eta_0 \in \mathbb{B}_\sigma(\eta^\star) \cap \mathcal{T}$, and define

$$p(\sigma) := \sup\{p' > 0 : \phi(p', 0) < \sigma\}. \tag{V.270}$$

We have by construction that for any $\eta_0 \in \mathbb{B}_{p(\sigma)}(\eta^\star) \cap \mathcal{T}$, if $\|\Delta\zeta^k\| \leq \bar{\delta}_\zeta$ is satisfied at each iteration, then the algorithm evolution is constrained to the ball $\mathbb{B}_\sigma(\eta^\star)$, namely, $\eta^k \in \mathbb{B}_\sigma(\eta^\star)$ for all k. Furthermore, thanks to the projection step in (V.32), it holds $\eta^k \in \mathbb{B}_\sigma(\eta^\star) \cap \mathcal{T}$ for all k. In this region, and thanks to Assumption V.5 which ensures condition (V.267) is satisfied for all k, we can use the bound in (V.269) to ensure that $|\Delta\zeta^k| < \bar{\delta}_\zeta$. Specifically, by picking $\bar{\delta}'_x = \frac{\bar{\delta}_\zeta}{4c_x(\eta^\star)}$, $\bar{\delta}'_u = \frac{\bar{\delta}_\zeta}{4c_u(\eta^\star)}$, and choosing $\delta_x \in (0, \min(\bar{\delta}_x, \bar{\delta}'_x))$, $\delta_u \in (0, \min(\bar{\delta}_u, \bar{\delta}'_u))$ it holds, from (V.269), that

$$|\Delta\zeta^k| \leq \delta_\zeta(\delta_x, \delta_u) \leq \frac{\bar{\delta}_\zeta}{2} \tag{V.271}$$

for all k. The bound in (V.271) then ensures that Algorithm 7 is Locally Uniformly Ultimately Bounded by Lemma V.2. Finally, we prove also that the bound in (V.33) is a strictly increasing function of δ_x, δ_u. Being δ_x, δ_u independent on the algorithm iteration k, it holds, from (V.269), that $\delta_\zeta(\delta_x, \delta_u)$ is bounded by a strictly increasing function of δ_x and δ_u. Hence, since $b(\delta_\zeta)$ is a class $\mathcal{K}$ function of its argument, we conclude using (V.269) that

$$b(\delta_\zeta(\delta_x, \delta_u)) \leq b(g(\delta_x + \delta_u)) \coloneqq b'(\delta_x, \delta_u), \tag{V.272}$$

from which the result follows with the function $b'(\delta_x, \delta_u)$.

※

Ringraziamenti

Per prima cosa, vorrei ringraziare Giuseppe per l'impegno e il sacrificio che ha messo e continua a mettere nel costruire un bel gruppo di ricerca attivo e coeso (il Notarnetwork), in cui non mancano opportunità di collaborazione e crescita con nessuno (né fuori né dentro al gruppo). Indubbiamente, il CASY di oggi è completamente diverso da quello in cui sono entrato, ormai più di tre anni fa: siamo passati dallo stare in stanze vuote, spoglie, scaldando uno alla volta il pranzo portato a casa al microonde, al non avere più spazio e pranzare tutti insieme con la pasta cucinata al momento. Questo è stato possibile solamente grazie al bel gruppo di cui Giuseppe è artefice.
Dal punto di vista tecnico e scientifico non posso che ringraziare Alessandro, che mi ha fatto scoprire un mondo di lati interessanti della teoria del controllo, primo di tutti l'adaptive control. A parer mio (e come ho cercato di riflettere in questa tesi), questi argomenti sono decisamente sottovalutati e sono più affascinanti, profondi e significativi di gran parte dei trend su cui la ricerca di oggi si concentra. È sicuramente grazie a questa guida che ho potuto appassionarmi a quello che faccio, eliminando del tutto la famigerata pressione del "publish or perish".
Per rimanere in tema guida scientifica, vorrei ringraziare anche Ivano, Andrea e Andrea per avermi introdotto ai temi dell'ottimizzazione, controllo ottimo e reinforcement learning, e per avermi insegnato, con Giuseppe -o provato a insegnarmi- come si prepara ed espone una presentazione fatta bene.
Grazie di cuore a Unibo e alla mensa di ingegneria, che insieme hanno fornito la spinta evolutiva necessaria allo sviluppo di una microeconomia interna al CASY (atta a sostenere la sopravvivenza dei poveri dottorandi). Insieme a loro, ringrazio i fautori di questa rivoluzione economica e politica (la Comune), identificati nei miei colleghi storici di stanza Guido, Lorenzo e Lorenzo. Ad oggi, dopo più di due anni, possiamo dire che ogni obiettivo della Comune è stato abbondantemente raggiunto e superato, che la qualità della vita al CASY è migliorata enormemente, e possiamo lamentarci che i nuovi dottorandi non sappiano "com'era prima".
Grazie a Simone e a malBiagio (e anche a AleCecc), perchè sono gli unici in lab che si prestano ad ascoltare le mie riflessioni deliranti sulle connessioni tra teoria dei sistemi, scienze cognitive ed intelligenza artificiale. Di nuovo grazie a Simone anche per il feedback, essenziale, che mi dà sulle dimostrazioni che scrivo (inesorabilmente sbagliate quando vado in open loop). Purtroppo mi accorgo di riuscire a

dimostrare sempre tutto quando sono da solo (il che non è sano in un sistema formale coerente). Grazie al Trama per i trekking che ha organizzato (Gradient Trekking). In queste escursioni e nella gestione del lab, il nostro gruppo ha finalmente trovato le mitologiche applicazioni pratiche dell'ottimizzazione distribuita. Grazie anche ad Alice e Riccardo e a tutti i successivi arrivati, che sembrano aver perfettamente colto lo spirito giusto, e che si apprestano a combattere le battaglie che ancora non sono vinte (acqua calda rossa, serpenti in laboratorio, luci che non vengono riparate, pioggia che filtra e stacca l'intonaco). Non posso citarvi tutti, altrimenti non rispetterei la scadenza della consegna.

Ci tengo anche a ringraziare Frank per avermi accolto -con decisamente poco preavviso- a Stoccarda, periodo in cui ho messo insieme i pezzi degli anni precedenti e che è stato per me di incredibile produttività accademica. Grazie anche a Yaryna e Jasmin, che mi hanno aiutato a gestire i passaggi burocratici per il periodo all'estero: non fosse per voi, non sarei partito, e starei ancora lottando contro quel mostro che sono le regole e la burocrazia del mio corso di dottorato (grazie, IBES). Grazie a tutti gli altri ragazzi dell'IST e alle persone con cui ho legato in quel periodo (e che mi hanno fatto mancare Bologna un po' meno). Grazie, Andrea, per avermi coinvolto anche nel tuo gruppo: da una semplice collaborazione stanno nascendo sviluppi molto interessanti, che pare incuriosiscano tante persone!

Passando alle amicizie, grazie a tutti gli storici componenti della Braga: Ace, Bruno, Cande, Dario, Giampi, Gio e Palo, compagni nei momenti migliori della mia vita da ormai un indecente numero di anni. Con loro, grazie anche ad Altea: anche se ci siamo conosciuti da poco, sei già per me un punto di riferimento. Grazie Boni, Mary, Nic e Sofi: con voi ho imparato cosa significa avere un progetto a lungo termine, delle responsabilità, e delle persone su cui contare.

Infine, ringrazio la mia famiglia: mamma e papà, Chicco e Silvia, gli zii e i nonni per avermi sempre supportato, se non da un punto di vista tecnico, da un punto di vista pratico. Senza la vostra solidità alle spalle, chissà dove sarei finito e con che fatica. E grazie a mio zio Francesco, che ha tracciato il mio futuro nel momento - tanto tempo fa ma che ancora mi ricordo - in cui mi ha spiegato cosa fosse un linguaggio di programmazione.

Concludo con un ringraziamento speciale a chi mi ha aiutato e sostenuto in maniera implicita (matematicamente parlando) e nascosta: so che ci siete. E un ringraziamento acausale a coloro che mi sosterranno in futuro: un giorno, forse, sarà chiara la ragione per cui ringraziarvi; nel frattempo, queste righe sono già per voi.

Marco

Bibliography

[1] P. Abbeel, M. Quigley, and A. Y Ng, *Using inaccurate models in reinforcement learning*, 23rd International Conference on Machine Learning, 2006, pp. 1–8.

[2] H.-S. Ahn, Y. Chen, and K. L Moore, *Iterative learning control: Brief survey and categorization*, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **37** (2007), no. 6, 1099–1121.

[3] S. Alarie, C. Audet, A. E Gheribi, M. Kokkolaras, and S. Le Digabel, *Two decades of blackbox optimization applications*, EURO Journal on Computational Optimization **9** (2021), 100011.

[4] S. Amaran, N. V Sahinidis, B. Sharda, and S. J Bury, *Simulation optimization: a review of algorithms and applications*, Annals of Operations Research **240** (2016), 351–380.

[5] B. D. Anderson, *Adaptive identification of multiple-input multiple-output plants*, 1974 ieee conference on decision and control including the 13th symposium on adaptive processes, 1974, pp. 273–281.

[6] ———, *An approach to multivariable system identification*, Automatica **13** (1977), no. 4, 401–408.

[7] ———, *Exponential stability of linear equations arising in adaptive identification*, IEEE Transactions on Automatic Control **22** (1977), no. 1, 83–88.

[8] B. D. Anderson, R. R. Bitmead, C. R. Johnsonm, P. V. Kokotovic, R. L. Kosut, I. Mareels, L. Praly, and B. D. Riedle, *Stability of adaptive systems*, Cambridge, MA: MIT Press, 1986.

[9] B. D. Anderson and C R. Johnson Jr, *Exponential convergence of adaptive identification and control algorithms*, Automatica **18** (1982), no. 1, 1–13.

[10] A. M Annaswamy and A. L Fradkov, *A historical perspective of adaptive control and learning*, Annual Reviews in Control **52** (2021), 18–41.

[11] P. J Antsaklis and A. N. Michel, *Linear systems*, Springer, 1997.

[12] M. ApS, *The mosek optimization toolbox for matlab manual. version 10.2*, 2024.

[13] Z. Artstein, *The limiting equations of nonautonomous ordinary differential equations*, Journal of Differential Equations **25** (1977), no. 2, 184–202.

[14] ———, *Uniform asymptotic stability via the limiting equations*, Journal of Differential Equations **27** (1978), no. 2, 172–189.

[15] K. J. Åström and P. Eykhoff, *System identification—a survey*, Automatica **7** (1971), no. 2, 123–162.

[16] K.-J. Åström and B. Torsten, *Numerical identification of linear dynamic systems from normal operating records*, IFAC Proceedings Volumes **2** (1965), no. 2, 96–111.

[17] E.-W. Bai and S. S. Sastry, *Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control*, Systems & Control Letters **6** (1985), no. 3, 153–163.

[18] F. A Bayer, G. Notarstefano, and F. Allgöwer, *A projected SQP method for nonlinear optimal control with quadratic convergence*, 52nd IEEE Conference on Decision and Control (CDC), 2013, pp. 6463–6468.

[19] A. S. Bazanella, M. Gevers, and L. Miškovic, *Closed-loop identification of mimo systems: a new look at identifiability and experiment design*, European Journal of Control **16** (2010), no. 3, 228–239.

[20] R. Bellman, *Dynamic programming*, Science **153** (1966), no. 3731, 34–37.

[21] R. E Bellman and S. E Dreyfus, *Applied dynamic programming*, Vol. 2050, Princeton university press, 2015.

[22] J. Berberich, A. Iannelli, A. Padoan, J. Coulson, F. Dörfler, and F. Allgöwer, *A quantitative and constructive proof of willems' fundamental lemma and its implications*, 2023 American Control Conference (ACC), 2023, pp. 4155–4160.

[23] J. Berberich, J. Köhler, M. A Müller, and F. Allgöwer, *Data-driven model predictive control with stability and robustness guarantees*, IEEE Transactions on Automatic Control **66** (2020), no. 4, 1702–1717.

[24] J. Berberich, C. W. Scherer, and F. Allgöwer, *Combining prior knowledge and data for robust controller design*, IEEE Transactions on Automatic Control **68** (2022), no. 8, 4618–4633.

[25] J. Berberich, S. Wildhagen, M. Hertneck, and F. Allgöwer, *Data-driven analysis and control of continuous-time systems under aperiodic sampling*, IFAC-PapersOnLine **54** (2021), no. 7, 210–215.

[26] P. Bernard, V. Andrieu, and D. Astolfi, *Observer design for continuous-time dynamical systems*, Annual Reviews in Control **53** (2022), 224–248.

[27] D. Bertsekas, *Dynamic programming and optimal control: Volume i*, Vol. 4, Athena scientific, 2012.

[28] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, 1999.

[29] T. Bian and Z.-P. Jiang, *Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design*, Automatica **71** (2016), 348–360.

[30] M. Bianchi, S. Grammatico, and J. Cortés, *Data-driven stabilization of switched and constrained linear systems*, Automatica **171** (2025), 111974.

[31] A. Bisoffi, C. De Persis, and P. Tesi, *Data-based stabilization of unknown bilinear systems with guaranteed basin of attraction*, Systems & Control Letters **145** (2020), 104788.

[32] R Bitmead, *Persistence of excitation conditions and the convergence of adaptive schemes*, IEEE Transactions on Information Theory **30** (1984), no. 2, 183–191.

[33] R Bitmead and B. D. Anderson, *Lyapunov techniques for the exponential stability of linear difference equations with random coefficients*, IEEE Transactions on Automatic Control **25** (1980), no. 4, 782–787.

[34] H. Bohr, *Zur Theorie der Fastperiodischen Funktionen: III. Dirichletentwicklung analytischer Funktionen*, Acta Mathematica **47** (1926), no. 3, 237 –281.

[35] M. Borghesi, S. Baroncini, G. Carnevale, A. Bosso, and G. Notarstefano, *On sufficient richness for linear time-invariant systems*, arXiv preprint arXiv:2502.04062 (2025).

[36] M. Borghesi, A. Bosso, and G. Notarstefano, *On-policy data-driven linear quadratic regulator via model reference adaptive reinforcement learning*, 2023 62nd IEEE Conference on Decision and Control (CDC), 2023, pp. 32–37.

[37] ———, *MR-ARL: Model reference adaptive reinforcement learning for robustly stable on-policy data-driven LQR*, arXiv preprint arXiv:2402.14483 (2024).

[38] A. Bosso, M. Borghesi, A. Iannelli, G. Notarstefano, and A. R. Teel, *Derivative-free data-driven control of continuous-time linear time-invariant systems*, arXiv preprint arXiv:2410.24167 (2024).

[39] L. Böttcher, N. Antulov-Fantulin, and T. Asikis, *AI Pontryagin or how artificial neural networks learn to control dynamical systems*, Nature Communications **13** (2022), no. 1, 333.

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

[41] S. Boyd and S. Sastry, *On parameter convergence in adaptive control*, Systems & Control Letters **3** (1983), no. 6, 311–319.

[42] ———, *Necessary and sufficient conditions for parameter convergence in adaptive control*, Automatica **22** (1986), no. 6, 629–639.

[43] S. J Bradtke, B E. Ydstie, and A. G Barto, *Adaptive linear quadratic control using policy iteration*, Proceedings of 1994 American Control Conference-acc'94, 1994, pp. 3475–3479.

[44] D. A Bristow, M. Tharayil, and A. G Alleyne, *A survey of iterative learning control*, IEEE Control Systems Magazine **26** (2006), no. 3, 96–114.

[45] L. Brivadis, V. Andrieu, P. Bernard, and U. Serres, *Further remarks on kkl observers*, Systems & Control Letters **172** (2023), 105429.

[46] A. E. Bryson, *Applied optimal control: optimization, estimation and control*, Routledge, 2018.

[47] R. S Bucy, *Global theory of the Riccati equation*, Journal of computer and system sciences **1** (1967), no. 4, 349–361.

[48] R. Carroll and D. Lindorff, *An adaptive observer for single-input single-output linear systems*, IEEE Transactions on Automatic Control **18** (1973), no. 5, 428–435.

[49] F. Celi, G. Baggio, and F. Pasqualetti, *Closed-form estimates of the LQR gain from finite data*, 2022 IEEE 61st Conference on Decision and Control (CDC), 2022, pp. 4016–4021.

[50] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, *Adversarial attacks and defences: A survey*, arXiv preprint arXiv:1810.00069 (2018).

[51] A. Charpentier, R. Elie, and C. Remlinger, *Reinforcement learning in economics and finance*, Computational Economics (2021), 1–38.

[52] Y. Chen and D. J Braun, *Hardware-in-the-loop iterative optimal feedback control without model-based future prediction*, IEEE Transactions on Robotics **35** (2019), no. 6, 1419–1434.

[53] A. R Conn, K. Scheinberg, and L. N Vicente, *Introduction to derivative-free optimization*, SIAM, 2009.

[54] R. R Costa, L. Hsu, A. K Imai, and P. Kokotović, *Lyapunov-based adaptive control of mimo systems*, Automatica **39** (2003), no. 7, 1251–1257.

[55] J. Coulson, H. J Van Waarde, J. Lygeros, and F. Dörfler, *A quantitative notion of persistency of excitation and the robust fundamental lemma*, IEEE Control Systems Letters **7** (2022), 1243–1248.

[56] C. Cruz-Hernández, J. Alvarez-Gallegos, and R. Castro-Linares, *Stability of discrete nonlinear systems under novanishing pertuabations: application to a nonlinear model-matching problem*, IMA Journal of Mathematical Control and Information **16** (1999), no. 1, 23–41.

[57] G. R G. da Silva, A. S Bazanella, C. Lorenzini, and L. Campestrini, *Data-driven LQR control design*, IEEE Control Systems Letters **3** (2018), no. 1, 180–185.

[58] D. P. De Farias and B. Van Roy, *The linear programming approach to approximate dynamic programming*, Operations Research **51** (2003), no. 6, 850–865.

[59] ———, *On constraint sampling in the linear programming approach to approximate dynamic programming*, Mathematics of operations research **29** (2004), no. 3, 462–478.

[60] J. DE O. PANTOJA, *Differential dynamic programming and newton's method*, International Journal of Control **47** (1988), no. 5, 1539–1553.

[61] C. De Persis, R. Postoyan, and P. Tesi, *Event-triggered control from data*, IEEE Transactions on Automatic Control (2023).

[62] C. De Persis and P. Tesi, *Formulas for data-driven control: Stabilization, optimality, and robustness*, IEEE Transactions on Automatic Control **65** (2019), no. 3, 909–924.

[63] ———, *Low-complexity learning of linear quadratic regulators from noisy data*, Automatica **128** (2021), 109548.

[64] ———, *Learning controllers for nonlinear systems from data*, Annual Reviews in Control (2023), 100915.

[65] S. Dean, S. Tu, N. Matni, and B. Recht, *Safely learning to control the constrained linear quadratic regulator*, 2019 American Control Conference (ACC), 2019, pp. 5582–5588.

[66] M. Diehl and S. Gros, *Numerical optimal control*, Optimization in Engineering Center (OPTEC) (2011).

[67] F. Dörfler, P. Tesi, and C. De Persis, *On the role of regularization in direct data-driven LQR control*, 2022 IEEE 61st Conference on Decision and Control (CDC), 2022, pp. 1091–1098.

[68] ———, *On the certainty-equivalence approach to direct data-driven LQR design*, IEEE Transactions on Automatic Control **68** (2023), no. 12, 7989–7996.

[69] J. C Dunn and D. P Bertsekas, *Efficient dynamic programming implementations of newton's method for unconstrained optimal control problems*, Journal of Optimization Theory and Applications **63** (1989), no. 1, 23–38.

[70] O. Eberhard, C. Vernade, and M. Muehlebach, *A Pontryagin perspective on reinforcement learning*, arXiv preprint arXiv:2405.18100 (2024).

[71] F. Edition, A. Papoulis, and S U. Pillai, *Probability, random variables, and stochastic processes*, McGraw-Hill Europe: New York, NY, USA, 2002.

[72] J. Eising and J. Cortes, *When sampling works in data-driven control: Informativity for stabilization in continuous time*, IEEE Transactions on Automatic Control (2024).

[73] H. Elliott, R. Cristi, and M. Das, *Global stability of adaptive pole placement algorithms*, IEEE Transactions on Automatic Control **30** (1985), no. 4, 348–356.

[74] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, *Global convergence of policy gradient methods for the linear quadratic regulator*, International Conference on Machine Learning, 2018, pp. 1467–1476.

[75] R. Galbiati, L. Sabug, F. Ruiz, and L. Fagiano, *Direct control design using a set membership-based black-box optimization approach*, 2022 IEEE Conference on Control Technology and Applications (CCTA), 2022, pp. 1259–1264.

[76] F. R. Gantmakher, *The theory of matrices*, Vol. 131, American Mathematical Soc., 2000.

[77] H. Garnier, M. Mensler, and A. Richard, *Continuous-time model identification from sampled data: implementation issues and performance evaluation*, International journal of Control **76** (2003), no. 13, 1337–1357.

[78] M. Gevers, A. S. Bazanella, X. Bombois, and L. Miskovic, *Identification and the information matrix: How to get just sufficiently rich?*, IEEE Transactions on Automatic Control **54** (2009), no. 12, 2828–2840.

[79] R. Goebel, R. G Sanfelice, and A. R Teel, *Hybrid dynamical systems: Modeling stability, and robustness*, Princeton University Press, Princeton, NJ, 2012.

[80] G. H Golub and C. F Van Loan, *Matrix computations*, JHU press, 2013.

[81] I. Goodfellow, *Deep learning*, MIT press, 2016.

[82] G. Goodwin and E. K. Teoh, *Persistency of excitation in the presence of possibly unbounded signals*, IEEE transactions on automatic control **30** (1985), no. 6, 595–597.

[83] G. Goodwin, P. Ramadge, and P. Caines, *Discrete-time multivariable adaptive control*, IEEE Transactions on Automatic Control **25** (1980), no. 3, 449–456.

[84] D. Gorinevsky, *On the persistency of excitation in radial basis function network identification of nonlinear systems*, IEEE Transactions on Neural Networks **6** (1995), no. 5, 1237–1244.

[85] M. Green and D. J. Limebeer, *Linear robust control*, Courier Corporation, 2012.

[86] M. Green and J. B Moore, *Persistence of excitation in linear systems*, Systems & Control Letters **7** (1986), no. 5, 351–360.

[87] A. Guha and A. M Annaswamy, *Online policies for real-time control using MRAC-RL*, 2021 60th IEEE Conference on Decision and Control (CDC), 2021, pp. 1808–1813.

[88] L. Guo and H. Zhao, *Online adaptive optimal control algorithm based on synchronous integral reinforcement learning with explorations*, Neurocomputing **520** (2023), 250–261.

[89] W. M Haddad and V. Chellaboina, *Nonlinear dynamical systems and control: a lyapunov-based approach*, Princeton university press, 2008.

[90] J. Hauser, *A projection operator approach to the optimization of trajectory functionals*, IFAC Proceedings Volumes **35** (2002), no. 1, 377–382.

[91] J. Hauser and D. G Meyer, *The trajectory manifold of a nonlinear control system*, 37th IEEE Conference on Decision and Control (CDC), 1998, pp. 1034–1039.

[92] R. Hermann and A. Krener, *Nonlinear controllability and observability*, IEEE Transactions on automatic control **22** (1977), no. 5, 728–740.

[93] A. Heydari and S. N Balakrishnan, *Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics*, IEEE Transactions on Neural Networks and Learning Systems **24** (2012), no. 1, 145–157.

[94] G. Hillerström and K. Walgama, *Repetitive control theory and applications-a survey*, IFAC Proceedings Volumes **29** (1996), no. 1, 1446–1451.

[95] D. Hinrichsen and D. Prätzel-Wolters, *A jordan canonical form for reachable linear systems*, Linear algebra and its applications **122** (1989), 489–524.

[96] S. Huang, N. Papernot, I. Goodfellow, Y. Duan, and P. Abbeel, *Adversarial attacks on neural network policies*, arXiv preprint arXiv:1702.02284 (2017).

[97] A. Inoue, Z. Iwai, M. Ishitobi, and H. Tokumaru, *Adaptive observer with exponential rate of convergence for multi-input multi-output systems*, IFAC Proceedings Volumes **17** (1984), no. 2, 1017–1022.

[98] P. A Ioannou and J. Sun, *Robust adaptive control*, Vol. 1, PTR Prentice-Hall Upper Saddle River, NJ, 1996.

[99] A. Isidori, *Nonlinear control systems ii*, Springer, 1999.

[100] D. H Jacobson, *New second-order and first-order algorithms for determining optimal control: A differential dynamic programming approach*, Journal of Optimization Theory and Applications **2** (1968), 411–440.

[101] B. M Jenkins, A. M Annaswamy, E. Lavretsky, and T. E Gibson, *Convergence properties of adaptive systems and the definition of exponential stability*, SIAM journal on control and optimization **56** (2018), no. 4, 2463–2484.

[102] Y. Jiang and Z.-P. Jiang, *Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics*, Automatica **48** (2012), no. 10, 2699–2704.

[103] R. E. Kalman, *On the general theory of control systems*, Proceedings first international conference on automatic control, moscow, ussr, 1960, pp. 481–492.

[104] R. E Kalman and R. S. Bucy, *New results in linear filtering and prediction theory* (1961).

[105] R. E. Kalman, *A new approach to linear filtering and prediction problems* (1960).

[106] N. Kazantzis and C. Kravaris, *Nonlinear observer design using lyapunov's auxiliary theorem*, Systems & Control Letters **34** (1998), no. 5, 241–247.

[107] H. Khalil, *Nonlinear systems third edition*, Prentice Hall, 2002.

[108] B. Kiumarsi, F. L Lewis, M.-B. Naghibi-Sistani, and A. Karimpour, *Optimal tracking control of unknown discrete-time linear systems using input-output measured data*, IEEE Transactions on Cybernetics **45** (2015), no. 12, 2770–2779.

[109] B. Kiumarsi, K. G Vamvoudakis, H. Modares, and F. L Lewis, *Optimal and autonomous control using reinforcement learning: A survey*, IEEE Transactions on Neural Networks and Learning Systems **29** (2017), no. 6, 2042–2062.

[110] K. Krauth, S. Tu, and B. Recht, *Finite-time analysis of approximate policy iteration for the linear quadratic regulator*, Advances in Neural Information Processing Systems **32** (2019).

[111] G. Kreisselmeier, *Adaptive observers with exponential rate of convergence*, IEEE Transactions on Automatic Control **22** (1977), no. 1, 2–8.

[112] M. Krstic, P. V Kokotovic, and I. Kanellakopoulos, *Nonlinear and adaptive control design*, John Wiley & Sons, Inc., 1995.

[113] V. Kučera, *A review of the matrix riccati equation*, Kybernetika **9** (1973), no. 1, 42–61.

[114] P. Kudva and K Narendra, *An identification procedure for discrete multivariable systems*, IEEE Transactions on Automatic Control **19** (1974), no. 5, 549–552.

[115] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, nature **521** (2015), no. 7553, 436–444.

[116] J. Y. Lee, J. B. Park, and Y. H. Choi, *Integral reinforcement learning for continuous-time input-affine nonlinear systems with simultaneous invariant explorations*, IEEE Transactions on Neural Networks and Learning Systems **26** (2014), no. 5, 916–932.

[117] W. Leonhard, *Control of electrical drives*, Springer Science & Business Media, 2001.

[118] W. Li and E. Todorov, *Iterative linear quadratic regulator design for nonlinear biological movement systems*, First International Conference on Informatics in Control, Automation and Robotics, 2004, pp. 222–229.

[119] L.-z. Liao and C. A Shoemaker, *Advantages of differential dynamic programming over newton's method for discrete-time optimal control problems*, Cornell University, 1992.

[120] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, *Controllability of complex networks*, nature **473** (2011), no. 7346, 167–173.

[121] L. Ljung and T. Söderström, *Theory and practice of recursive identification*, MIT press Cambridge, MA, 1983.

[122] L. Ljung, *Characterization of the concept of'persistently exciting'in the frequency domain* (1971).

[123] ———, *System identification*, Signal analysis and prediction, 1998, pp. 163–173.

[124] L. Ljung and S. Gunnarsson, *Adaptation and tracking in system identification—a survey*, Automatica **26** (1990), no. 1, 7–21.

[125] J. Löfberg, *YALMIP : A toolbox for modeling and optimization in MATLAB*, In proceedings of the cacsd conference, 2004.

[126] R. W Longman, *Iterative learning control and repetitive control for engineering practice*, International Journal of Control **73** (2000), no. 10, 930–954.

[127] V. G Lopez, M. Alsalti, and M. A Müller, *Efficient off-policy Q-learning for data-based discrete-time LQR problems*, IEEE Transactions on Automatic Control (2023).

[128] V. G Lopez and M. A Müller, *On a continuous-time version of Willems' lemma*, 2022 IEEE 61st Conference on Decision and Control (CDC), 2022, pp. 2759–2764.

[129] A. Loria, A. Chaillet, G. Besançon, and Y. Chitour, *On the PE stabilization of time-varying systems: open questions and preliminary answers*, Proceedings of the 44th IEEE Conference on Decision and Control, 2005, pp. 6847–6852.

[130] A. Lorıa and E. Panteley, *Uniform exponential stability of linear time-varying systems: revisited*, Systems & Control Letters **47** (2002), no. 1, 13–24.

[131] A. Loria, E. Panteley, D. Popovic, and A. R Teel, *Persistency of excitation for uniform convergence in nonlinear control systems*, arXiv preprint math (2003).

[132] A. Loría, E. Panteley, D. Popovic, and A. R Teel, *A nested Matrosov theorem and persistency of excitation for uniform convergence in stable nonautonomous systems*, IEEE Transactions on Automatic Control **50** (2005), no. 2, 183–198.

[133] A. Loría, E. Panteley, and A Teel, *Ugas of nonlinear time-varying systems: A/spl delta/-persistency of excitation approach*, Proceedings of the 39th IEEE Conference on Decision and Control (cat. no. 00ch37187), 2000, pp. 3489–3494.

[134] A. Loría, E. Panteley, and A. Teel, *A new notion of persistency-of-excitation for ugas of nltv systems: Application to stabilisation of nonholonomic systems*, 1999 European Control Conference (ECC), 1999, pp. 1363–1368.

[135] G. Luders and K. Narendra, *An adaptive observer and identifier for a linear system*, IEEE Transactions on Automatic Control **18** (1973), no. 5, 496–499.

[136] ______, *A new canonical form for an adaptive observer*, IEEE Transactions on Automatic Control **19** (1974), no. 2, 117–119.

[137] D. Luenberger, *Observers for multivariable systems*, IEEE Transactions on Automatic Control **11** (1966), no. 2, 190–197.

[138] ______, *Canonical forms for linear multivariable systems*, IEEE Transactions on Automatic Control **12** (1967), no. 3, 290–293.

[139] ______, *An introduction to observers*, IEEE Transactions on Automatic Control **16** (1971), no. 6, 596–602.

[140] D. G. Luenberger, *Observing the state of a linear system*, IEEE Transactions on Military Electronics **8** (1964), no. 2, 74–80.

[141] H. Ma, D. Büchler, B. Schölkopf, and M. Muehlebach, *Reinforcement learning with model-based feedforward inputs for robotic table tennis*, Autonomous Robots **47** (2023), no. 8, 1387–1403.

[142] I. M. Mareels and M. Gevers, *Persistency of excitation criteria for linear, multivariable, time-varying systems*, Mathematics of Control, Signals and Systems **1** (1988), 203–226.

[143] I. Markovsky, E. Prieto-Araujo, and F. Dörfler, *On the persistency of excitation*, Automatica **147** (2023), 110657.

[144] D. Mayne, *A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems*, International Journal of Control **3** (1966), no. 1, 85–95.

[145] L. Menini, C. Possieri, and A. Tornambè, *Algebraic analysis of the structural properties of parametric linear time-invariant systems*, IET Control Theory & Applications **14** (2020), no. 20, 3568–3579.

[146] J. Miller and M. Sznaier, *Data-driven gain scheduling control of linear parameter-varying systems using quadratic matrix inequalities*, IEEE Control Systems Letters **7** (2022), 835–840.

[147] S. Miller and D. Childers, *Probability and random processes: With applications to signal processing and communications*, Academic Press, 2012.

[148] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, *Playing Atari with deep reinforcement learning*, arXiv preprint arXiv:1312.5602 (2013).

[149] H. Modares and F. L Lewis, *Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning*, IEEE Transactions on Automatic Control **59** (2014), no. 11, 3051–3056.

[150] H. Modares, F. L Lewis, and Z.-P. Jiang, *Optimal output-feedback control of unknown continuous-time linear systems using off-policy reinforcement learning*, IEEE Transactions on Cybernetics **46** (2016), no. 11, 2401–2410.

[151] H. Modares, F. L Lewis, and M.-B. Naghibi-Sistani, *Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks*, IEEE Transactions on Neural Networks and Learning Systems **24** (2013), no. 10, 1513–1525.

[152] J Moore, *Persistence of excitation in extended least squares*, IEEE Transactions on Automatic Control **28** (1983), no. 1, 60–68.

[153] K. L Moore, *Iterative learning control for deterministic systems*, Springer Science & Business Media, 2012.

[154] A. Morgan and K. Narendra, *On the stability of nonautonomous differential equations* $\dot{x} = A + B(t)x$*, with skew symmetric matrix B(t)*, SIAM Journal on Control and Optimization **15** (1977), no. 1, 163–176.

[155] ______, *On the uniform asymptotic stability of certain linear nonautonomous differential equations*, SIAM Journal on Control and Optimization **15** (1977), no. 1, 5–24.

[156] A S. Morse, *Structural invariants of linear multivariable systems*, SIAM Journal on Control **11** (1973), no. 3, 446–465.

[157] E. Mosca, *Optimal, predictive, and adaptive control*, Prentice-Hall, Inc., 1994.

[158] D. M Murray and S. J Yakowitz, *Differential dynamic programming and newton's method for discrete optimal control problems*, Journal of Optimization Theory and Applications **43** (1984), no. 3, 395–414.

[159] K. Nar and S S. Sastry, *Persistency of excitation for robustness of neural networks*, arXiv preprint arXiv:1911.01043 (2019).

[160] K. Narendra and Y.-H. Lin, *Stable discrete adaptive control*, IEEE Transactions on Automatic Control **25** (1980), no. 3, 456–461.

[161] K. Narendra and A. M Annaswamy, *Persistent excitation in adaptive systems*, International Journal of Control **45** (1987), no. 1, 127–160.

[162] K. S Narendra and A. M Annaswamy, *Stable adaptive systems*, Courier Corporation, 2012.

[163] P. Nikiforuk, H Ohta, and M. Gupta, *Adaptive observer and identifier design for multi-input multi-output systems*, IFAC Proceedings Volumes **10** (1977), no. 6, 189–196.

[164] J. Nocedal and S. J Wright, *Numerical optimization*, Springer, 1999.

[165] N. Nordström and S. S. Sastry, *Persistency of excitation in possibly unstable continuous time systems and parameter convergence in adaptive identification*, Adaptive systems in control and signal processing 1986, 1987, pp. 347–352.

[166] B. Nortmann and T. Mylvaganam, *Data-driven control of linear time-varying systems*, 2020 59th ieee conference on decision and control, 2020, pp. 3939–3944.

[167] J. Numper, R. Evans, et al., *Highly accurate protein structure prediction with alphafold*.

[168] A. Padoan, G. Scarciotti, and A. Astolfi, *A geometric characterization of the persistence of excitation condition for the solutions of autonomous systems*, IEEE Transactions on Automatic Control **62** (2017), no. 11, 5666–5677.

[169] B. Pang, T. Bian, and Z.-P. Jiang, *Data-driven finite-horizon optimal control for linear time-varying discrete-time systems*, 2018 IEEE Conference on Decision and Control (CDC), 2018, pp. 861–866.

[170] ———, *Robust policy iteration for continuous-time linear quadratic regulation*, IEEE Transactions on Automatic Control **67** (2021), no. 1, 504–511.

[171] E. Panteley, A. Loria, and A. Teel, *Relaxed persistency of excitation for uniform asymptotic stability*, IEEE Transactions on Automatic Control **46** (2001), no. 12, 1874–1886.

[172] C. Possieri and M. Sassano, *Q-Learning for continuous-time linear systems: A data-driven implementation of the Kleinman algorithm*, IEEE Transactions on Systems, Man, and Cybernetics: Systems **52** (2022), no. 10, 6487–6497.

[173] ———, *Value iteration for continuous-time linear time-invariant systems*, IEEE Transactions on Automatic Control (2022).

[174] K. Prag, M. Woolway, and T. Celik, *Toward data-driven optimal control: A systematic review of the landscape*, IEEE Access **10** (2022), 32190–32212.

[175] A. C. Ran and L. Rodman, *On parameter dependence of solutions of algebraic Riccati equations*, Mathematics of Control, Signals and Systems **1** (1988), 269–284.

[176] P. Rapisarda, M K. Camlibel, and H. Van Waarde, *A persistency of excitation condition for continuous-time systems*, IEEE Control Systems Letters **7** (2022), 589–594.

[177] B. Recht, *A tour of reinforcement learning: The view from continuous control*, Annual Review of Control, Robotics, and Autonomous Systems **2** (2019), 253–279.

[178] L. M. Rios and N. V Sahinidis, *Derivative-free optimization: a review of algorithms and comparison of software implementations*, Journal of Global Optimization **56** (2013), no. 3, 1247–1293.

[179] S. A. A. Rizvi and Z. Lin, *Output feedback reinforcement Q-learning control for the discrete-time linear quadratic regulator problem*, 2017 IEEE 56th annual Conference on Decision and Control (CDC), 2017, pp. 1311–1316.

[180] ———, *Reinforcement learning-based linear quadratic regulation of continuous-time systems using dynamic output feedback*, IEEE Transactions on Cybernetics **50** (2019), no. 11, 4670–4679.

[181] A. Rosenblueth, N. Wiener, and J. Bigelow, *Behavior, purpose and teleology*, Philosophy of science **10** (1943), no. 1, 18–24.

[182] M. Rotulo, C. De Persis, and P. Tesi, *Data-driven linear quadratic regulation via semidefinite programming*, IFAC-PapersOnLine **53** (2020), no. 2, 3995–4000.

[183] L. Sabug, F. Ruiz, and L. Fagiano, *On the use of set membership theory for global optimization of black-box functions*, 2020 59th IEEE Conference on Decision and Control (CDC), 2020, pp. 3586–3591.

[184] L. Sabug Jr, F. Ruiz, and L. Fagiano, *SMGO: A set membership approach to data-driven global optimization*, Automatica **133** (2021), 109890.

[185] S. Sagara and Z.-Y. Zhao, *Numerical integration approach to on-line identification of continuous-time systems*, Automatica **26** (1990), no. 1, 63–74.

[186] R. G. Sanfelice and A. R. Teel, *Lyapunov analysis of sample-and-hold hybrid feedbacks*, Proceedings of the 45th ieee conference on decision and control, 2006, pp. 4879–4884.

[187] ———, *On singular perturbations due to fast actuators in hybrid control systems*, Automatica **47** (2011), no. 4, 692–701.

[188] A. Saoud, M. Maghenem, A. Loria, and R. G. Sanfelice, *Hybrid persistency of excitation in adaptive estimation for hybrid systems*, IEEE Transactions on Automatic Control (2024).

[189] R. Sargent, *Optimal control*, Journal of Computational and Applied Mathematics **124** (2000), no. 1-2, 361–371.

[190] A. Sarker, P. Fisher, J. E Gaudio, and A. M Annaswamy, *Accurate parameter estimation for safety-critical systems with unmodeled dynamics*, Artificial Intelligence **316** (2023), 103857.

[191] S. Sastry and M. Bodson, *Adaptive control: stability, convergence, and robustness*, Acoustical Society of America, 1990.

[192] L. Sforni, G. Carnevale, I. Notarnicola, and G. Notarstefano, *On-policy data-driven linear quadratic regulator via combined policy iteration and recursive least squares*, 2023 62nd IEEE Conference on Decision and Control (CDC), 2023, pp. 5047–5052.

[193] ———, *Stability-certified on-policy data-driven LQR via recursive learning and policy gradient*, arXiv preprint arXiv:2403.05367 (2024).

[194] C. E. Shannon, *A mathematical theory of communication*, The Bell system technical journal **27** (1948), no. 3, 379–423.

[195] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Force control*, Springer, 2009.

[196] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, *A general reinforcement learning algorithm that masters chess, shogi, and go through self-play*, Science **362** (2018), no. 6419, 1140–1144.

[197] T. Söderström and P. Stoica, *System identification*, Prentice-Hall international series in systems and control engineering, Prentice Hall, 1989.

[198] M. Sondhi and D. Mitra, *New results on the performance of a well-known class of adaptive filters*, Proceedings of the IEEE **64** (1976), no. 11, 1583–1597.

[199] Y. Song, A. Romero, M. Müller, V. Koltun, and D. Scaramuzza, *Reaching the limit in autonomous racing: Optimal control versus reinforcement learning*, Science Robotics **8** (2023), no. 82, eadg1462.

[200] K. Sridhar, O. Sokolsky, I. Lee, and J. Weimer, *Improving neural network robustness via persistency of excitation*, 2022 American Control Conference (ACC), 2022, pp. 1521–1526.

[201] B. L Stevens, F. L Lewis, and E. N Johnson, *Aircraft control and simulation: dynamics, controls design, and autonomous systems*, John Wiley & Sons, 2015.

[202] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, The MIT Press, 2018.

[203] Y. Tan, W. H Moase, C. Manzie, D. Nešić, and I. M. Mareels, *Extremum seeking from 1922 to 2010*, Proceedings of the 29th Chinese Control Conference, 2010, pp. 14–26.

[204] G. Tao, *Adaptive control design and analysis*, Vol. 37, John Wiley & Sons, 2003.

[205] ———, *Multivariable adaptive control: A survey*, Automatica **50** (2014), no. 11, 2737–2764.

[206] A. R Teel, L. Moreau, and D. Nesic, *A unified framework for input-to-state stability in systems with two time scales*, IEEE Transactions on Automatic Control **48** (2003), no. 9, 1526–1544.

[207] G. P Tolstov, *Fourier series*, Courier Corporation, 2012.

[208] A. M Turing, *Computing machinery and intelligence*, Springer, 2009.

[209] H Unbehauen and G. Rao, *A review of identification in continuous-time systems*, Annual reviews in Control **22** (1998), 145–171.

[210] H. Unbehauen and G. Rao, *Continuous-time approaches to system identification—a survey*, Automatica **26** (1990), no. 1, 23–35.

[211] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—implementation—applications*, Springer Science & Business Media, 2012.

[212] H. J. van Waarde, M. K. Camlibel, and M. Mesbahi, *From noisy data to feedback controllers: Nonconservative design via a matrix S-lemma*, IEEE Transactions on Automatic Control **67** (2020), no. 1, 162–175.

[213] H. J Van Waarde, C. De Persis, M K. Camlibel, and P. Tesi, *Willems' fundamental lemma for state-space systems and its extension to multiple datasets*, IEEE Control Systems Letters **4** (2020), no. 3, 602–607.

[214] H. J Van Waarde, J. Eising, H. L Trentelman, and M. K. Camlibel, *Data informativity: A new perspective on data-driven analysis and control*, IEEE Transactions on Automatic Control **65** (2020), no. 11, 4753–4768.

[215] J. Venkatasubramanian, J. Köhler, J. Berberich, and F. Allgöwer, *Sequential learning and control: Targeted exploration for robust performance*, IEEE Transactions on Automatic Control (2024).

[216] J. Venkatasubramanian, J. Köhler, M. Cannon, and F. Allgöwer, *Towards targeted exploration for non-stochastic disturbances*, IFAC-PapersOnLine **58** (2024), no. 15, 556–561.

[217] F. Vicentini, *Collaborative robotics: a survey*, Journal of Mechanical Design **143** (2021), no. 4, 040802.

[218] D. Vrabie, O Pastravanu, M. Abu-Khalaf, and F. L Lewis, *Adaptive optimal control for continuous-time linear systems based on policy iteration*, Automatica **45** (2009), no. 2, 477–484.

[219] G. C. Walsh and H. Ye, *Scheduling of networked control systems*, IEEE control systems magazine **21** (2001), no. 1, 57–65.

[220] W G. Walter, *An imitation of life*, Scientific american **182** (1950), no. 5, 42–45.

[221] ———, *A machine that learns*, Scientific American **185** (1951), no. 2, 60–64.

[222] C. J. C. H. Watkins, *Learning from delayed rewards* (1989).

[223] A. Weiss and D. Mitra, *Digital adaptive filters: Conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation*, IEEE Transactions on Information Theory **25** (1979), no. 6, 637–652.

[224] N. Wiener, *Generalized harmonic analysis*, Acta mathematica **55** (1930), no. 1, 117–258.

[225] ———, *Cybernetics or control and communication in the animal and the machine*, MIT press, 2019.

[226] J. C Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, *A note on persistency of excitation*, Systems & Control Letters **54** (2005), no. 4, 325–329.

[227] W. M. Wonham, *On a matrix riccati equation of stochastic control*, SIAM Journal on Control **6** (1968), no. 4, 681–697.

[228] ———, *Linear multivariable control: a geometric approach*, Springer-Verlag, Berlin; New York, 1974.

[229] W. M. Wonham and A. Morse, *Feedback invariants of linear multivariable systems*, Automatica **8** (1972), no. 1, 93–100.

[230] H. Yang, H. Itakura, and Y. Nishikawa, *Two approaches for adaptive observer in multi-output systems*, IFAC Proceedings Volumes **14** (1981), no. 2, 1093–1098.

[231] P. Young, *Parameter estimation for continuous-time models—a survey*, Automatica **17** (1981), no. 1, 23–39.

[232] J Yuan and W. M. Wonham, *Probing signals for model reference identification*, IEEE Transactions on Automatic Control **22** (1977), no. 4, 530–538.

[233] M. Zhang and T.-J. Tarn, *Hybrid control of the pendubot*, IEEE/ASME Transactions on Mechatronics **7** (2002), no. 1, 79–86.

[234] K. Zhou and J. C. Doyle, *Essentials of robust control*, Vol. 104, Prentice hall Upper Saddle River, NJ, 1998.

[235] I. Ziemann, A. Tsiamis, H. Sandberg, and N. Matni, *How are policy gradient methods affected by the limits of control?*, 2022 IEEE 61st Conference on Decision and Control (CDC), 2022, pp. 5992–5999.