2022

Statistical Learning For System Identification, Estimation, And Control

Anastasios Tsiamis
University of Pennsylvania

Statistical Learning For System Identification, Estimation, And Control

Abstract

Despite the recent widespread success of machine learning, we still do not fully understand its fundamental limitations. Going forward, it is crucial to better understand learning complexity, especially in critical decision making applications, where a wrong decision can lead to catastrophic consequences. In this thesis, we focus on the statistical complexity of learning unknown linear dynamical systems, with focus on the tasks of system identification, prediction, and control. We are interested in sample complexity, i.e. the minimum number of samples required to achieve satisfactory learning performance. Our goal is to provide finite-sample learning guarantees, explicitly highlighting how the learning objective depends on the number of samples. A fundamental question we are trying to answer is how system theoretic properties of the underlying process can affect sample complexity.

Using recent advances in statistical learning, high-dimensional statistics, and control theoretic tools, we provide finite-sample guarantees in the following settings. i) System Identification. We provide the first finite-sample guarantees for identifying a stochastic partially-observed system; this problem is also known as the stochastic system identification problem. ii) Prediction. We provide the first end-to-end guarantees for learning the Kalman Filter, i.e. for learning to predict, in an offline learning architecture. We also provide the first logarithmic regret guarantees for the problem of learning the Kalman Filter in an online learning architecture, where the data are revealed sequentially. iii) Difficulty of System Identification and Control. Focusing on fully-observed systems, we investigate when learning linear systems is statistically easy or hard, in the finite sample regime. Statistically easy to learn linear system classes have sample complexity that is polynomial with the system dimension. Statistically hard to learn linear system classes have worst-case sample complexity that is at least exponential with the system dimension. We show that there actually exist classes of linear systems, which are hard to learn. Such classes include indirectly excited systems with large degree of indirect excitation. Similar conclusions hold for both the problem of system identification and the problem of learning to control.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Electrical & Systems Engineering

First Advisor

George J. Pappas

Keywords

Online Control, Online Prediction, Statistical Learning, System Identification

Subject Categories

Computer Sciences | Electrical and Electronics | Statistics and Probability

STATISTICAL LEARNING FOR SYSTEM IDENTIFICATION, ESTIMATION, AND CONTROL

Anastasios Tsiamis

A DISSERTATION

in

Electrical and Systems Engineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2022

Supervisor of Dissertation

George J. Pappas, UPS Foundation Professor of Electrical and Systems Engineering

Graduate Group Chairperson

Alejandro Ribeiro, Professor of Electrical and Systems Engineering

Dissertation Committee

John Lygeros, Professor, Department of Information Technology and Electrical Engineering, ETH Zürich

Nikolai Matni, Assistant Professor, Electrical and Systems Engineering, University of Pennsylvania

Manfred Morari (Chair), Practice Professor, Peter and Susanne Armstrong Faculty Fellow, Department of Electrical and Systems Engineering, University of Pennsylvania

George J. Pappas, UPS Foundation Professor and Chair of the Electrical and Systems Engineering, University of Pennsylvania

Acknowledgements

I would really like to thank my advisor, George, for his invaluable support and guidance. I am really happy and proud to be your PhD student. Thank you for giving me freedom and trusting me to work on this topic although we didn't know if it would work and it took some time to get results. I am glad that I learned a lot under your supervision. I learned to be more independent and to always look at the big picture.

I would like to thank all my PhD committee members Nik, Manfred and John. Nik, your work has been really inspirational for me. I started thinking about the problems in this thesis when I first read your paper on the sample complexity of the Linear Quadratic Regulator. Through our collaboration and your course I learned a lot about the area. Manfred, it has always been a pleasure to discuss with you about current and future research. You have been a great mentor, always finding the weak points in my work and helping me to improve. John, thank you very much for your valuable feedback and time. Also thank you for giving me the opportunity to work with you as a Postdoc.

I would also like to thank all my remaining collaborators during my PhD, Sergio Pequito, Konstantinos Gatsis (KG), Andreea Alexandru, Dionysis Kalogerias, Luiz Chamon, Alejandro Ribeiro, Charis Stamouli, and Ingvar Ziemann. I enjoyed working with you and I really learned a lot. I would also like to thank all my remaining colleagues and lab-mates for useful discussions.

Of course it wouldn't be possible to write this thesis without the support of my friends in Philadelphia. I cannot really thank enough my roommates and friends Giorgos Pavlakos and Nikos Kolotouros. Giorgo, your advice and personal experience with PhD life were very

helpful to me and you were always eager to help me deal with all sorts of things. Niko, I enjoyed our deep pseudo-philosophical discussions, the bad jokes, and the silly stories from your past. See you soon in Zürich. I would like to thank Vasso, Spyros, Vassilis, Ioanna, and Vassilis Tzoumas for helping me settle in Philadelphia and making me feel like home. Many thanks to the Brunch group: Andreea, Achin, Alëna, Matei, Kuk, and Heejin. I had so much fun with you and I will miss you a lot. Andreea thank you for the engaging discussions about research and life in general. Achin and Alëna, thank you for your support during tough moments of my PhD. Next, I would like to thank the BBQ group Melissa, Luiz, Santi, Papi, Alp, Markos, Dimi, Antigoni and the newer recruits Juan, Kissas, Kallas, and Eliza. Melissa, Luiz, Santi, Papi, Dimi thank you for supporting me during good and tough times and thank you for tolerating me when I could never hear clearly what you were saying. Melissa, I miss so much our coffee breaks. My PhD life would be boring without any music. I am grateful to have played music in A lo Chorro Orchestra, alongside the world class musicians Dimi, Luiz, Kallas, Nacho, and Juan Jr. Many thanks to the L475 lab, Aritra, Thomas, Lars, and the honorary members Alex, Matthew, Rado. Also many thanks to my Catalan friends Mariona, Elena, Estela, Ingrid. Special thanks to Marina, Eliza, Dimi, Nikos and Kallas for their support during the last years of my PhD and during the quarantine. Marina, best of luck with your next steps. Eliza, I miss our fights and intense discussions already.

Thanks to all my friends in Greece Kalliopi, Drosia, Maria, Giorgos, Achilleas, Palyvos, Danae, Vassilis, Sokratis, Anthi, Charis and many others. Finally, I want to thank from the bottom of my heart my parents Giota, Lefteris, and my dear sister Antigoni. This thesis is dedicated to you.

Abstract

STATISTICAL LEARNING FOR SYSTEM IDENTIFICATION, ESTIMATION, AND CONTROL

Anastasios Tsiamis

George J. Pappas

Despite the recent widespread success of machine learning, we still do not fully understand its fundamental limitations. Going forward, it is crucial to better understand learning complexity, especially in critical decision making applications, where a wrong decision can lead to catastrophic consequences. In this thesis, we focus on the statistical complexity of learning unknown linear dynamical systems, with focus on the tasks of system identification, prediction, and control. We are interested in sample complexity, i.e. the minimum number of samples required to achieve satisfactory learning performance. Our goal is to provide finite-sample learning guarantees, explicitly highlighting how the learning objective depends on the number of samples. A fundamental question we are trying to answer is how system theoretic properties of the underlying process can affect sample complexity.

Using recent advances in statistical learning, high-dimensional statistics, and control theoretic tools, we provide finite-sample guarantees in the following settings. i) System Identification. We provide the first finite-sample guarantees for identifying a stochastic partially-observed system; this problem is also known as the stochastic system identification problem. ii) Prediction. We provide the first end-to-end guarantees for learning the Kalman Filter, i.e. for learning to predict, in an offline learning architecture. We also provide the first logarithmic regret guarantees for the problem of learning the Kalman Filter in an online learning architecture, where the data are revealed sequentially. iii) Difficulty of System Identification and Control. Focusing on fully-observed systems, we investigate when learning linear systems is statistically easy or hard, in the finite sample regime. Statistically easy to learn linear system classes have sample complexity that is polynomial with the system dimension. Statistically hard to learn linear system classes have worst-case sample complexity that is at least exponential with the system dimension. We show that there actually exist classes of linear systems, which are hard to learn. Such classes include indirectly excited systems with large degree of indirect excitation. Similar conclusions hold for both the problem of system identification and the problem of learning to control.

Contents

List of Tables

List of Illustrations

Chapter 1

Introduction

In the last decade, we have witnessed an unprecedented growth of machine learning and artificial intelligence applications. Renewed interest was sparked by the successful deployment of deep learning in tasks like computer vision (Krizhevsky et al., 2012), recommendation systems (Zhang et al., 2019), natural language processing (Young et al., 2018), biology (Jumper et al., 2021), and finance (Heaton et al., 2017). This success was also accelerated by the availability of large-scale data and specialized hardware. It is more accessible than ever to apply machine learning tools in any task with potential real-world impact.

However, despite its potential, it is still challenging to apply machine learning in certain settings, especially when dealing with critical physical applications involving decision making, e.g. self-driving cars, autonomous systems etc. Reinforcement learning has been effectively used to control systems in structured environments, e.g. in games (Silver et al., 2017), robotic manipulation in the lab (Levine et al., 2016), and simulated physics tasks (Lillicrap et al., 2015) to name a few. However, unless the environment is structured, e.g. games, simulators, or the lab, it is a very challenging problem to guarantee safe deployment of machine learning systems in real world decision and control systems. Without any guarantees, in the worst case there could be unpredictable consequences leading to system failure. Naturally, a question that arises is can we learn to control systems with guarantees? What are the fundamental limitations of learning in the context of control systems?

In this thesis, we take a step towards studying the limitations of learning in the context of control systems. We focus on linear systems driven by stochastic noise. Linear systems are simple enough to allow for an in-depth theoretical analysis, yet exhibit sufficiently rich behavior so that we can draw conclusions about control of more general system classes (Recht, 2019). Even in this simple setting, understanding learning limitations is quite challenging and highly non-trivial.

There are many different aspects of learning complexity. From a *computational* point of view, a learning task is difficult if it requires a lot of computational resources. From a *statistical* point of view, the difficulty of learning is captured by the amount of data needed to achieve satisfactory performance. In this thesis, we are interested in the latter, i.e. the statistical difficulty of learning in stochastic control systems.

Typically, there are two main tasks in stochastic control systems, the prediction and the regulation problem. In the former task, our goal is to predict the state of the underlying physical process based on noisy measurements, while in the latter our goal is to drive the process to a desired set-point or trajectory based on state feedback. Traditionally, the implementation of both procedures requires a model of the underlying system. However, in reality, the system might be unknown and we might not have access to such a model. This might happen if, for example, the physics are complex or some parameters are uncertain. In this case, we need to learn how to do control/prediction based on data.

Of course, the problem of learning *unknown* linear dynamical systems from data was studied extensively before in the context of system identification (Ljung, 1999) and adaptive control (Goodwin et al., 1981), long before the recent resurgence of machine learning. The convergence and statistical properties of such algorithms has been well understood in the *asymptotic regime*, i.e. when we have an *infinite* number of samples. However, in reality we only have access to a finite number of samples. In the *finite-sample regime*, tools like the Central Limit Theorem, can be only applied heuristically, while any existing rigorous results are conservative (Vershynin, 2018). Besides, under the big-O notation, it is not always clear how the learning performance depends on various system theoretic properties. Hence, there is a gap in our theoretical understanding of the statistical complexity of learning.

With the advances in statistical learning and high-dimensional statistics (Vershynin, 2018), there has been a recent shift of focus from asymptotic analysis with infinite data to statistical analysis with finite data (Matni & Tu, 2019; Matni et al., 2019). Inspired by this recent line of work, the goal of this thesis is to contribute towards a finite-sample complexity theory for dynamical systems learning. We aim to understand the relation between the number of samples and accuracy of control/prediction. We are also interested in understanding how system theoretic properties affect statistical learning difficulty. Towards this goal, we leverage modern statistical learning results (Vershynin, 2018) to bypass the limitations of asymptotic tools. Such tools should be applied carefully since the learning and control components interact with each other. Besides, time-series data are highly correlated, unlike the standard statistical learning setting where the collected samples are independent (Shalev-Shwartz & Ben-David, 2014).

1.1 Learning unknown linear systems

In this thesis, we focus on *unknown* linear systems of the form:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + v_k, \end{aligned} \tag{1.1}$$

where x_k is the state, u_k is the control input, and y_k is the measured output. The noise signals w_k, v_k represent the process and measurement noise respectively. Typically, we have access to an input-output data-set of length N, where N denotes the number of samples; the data can belong to either a single trajectory or multiple trajectories. Since system (1.1) is unknown, we have to use data in order to learn how to control the system or predict its evolution. Our main focus in this thesis is how the number of samples N affects learning performance. Of course learning performance can vary since there are many different learning methods that we can deploy. The selection of learning method depends on the nature of the data but also on our own engineering choices. Below, we list some of the most important conceptual and architectural distinctions.

Model-based versus model-free learning. One way to control or predict the evolution of system (1.1) is to first learn a model, i.e. estimate the unknown parameters A, B, C and potentially the noise profiles from data. This problem is of independent interest and has been known as the *system identification* problem Ljung (1999). After obtaining the model, we can design a controller or a filter. Any method that uses some kind of a model representation in order to control or predict system (1.1), will be called a *model-based* method. At the other end, we have *model-free* learning methods. In this case, we directly parameterize the control policy or the predictor based on some parameter, say θ, and use the data to find a value of θ that leads to good control/prediction performance. Of course, the distinction between model-based and model-free methods goes beyond the class of linear systems. It is a design choice whether to choose the former or the latter as both have advantages and disadvantages. For example, model-based methods can be more interpretable for some systems; also it might be easier to argue about robustness or guarantees once we have a model. However, if the assumed model class is wrong, the introduced bias might affect control performance significantly. Such methods are paired well with Model Predictive Control implementations. On the other hand, model-free methods might be more suitable for complex control tasks, but they might require more samples. Such methods have been studied in the model-free Reinforcement Learning literature. In this thesis we focus mostly on model-based methods.

Offline versus online learning. An important architectural distinction is whether learning is performed offline or online. In the *offline* setting, we have access to pre-collected input-output data. We perform learning once and we deploy the learned controller or predictor without using any of the new data that become available as the system is evolving. Learning and decision making happen in separate steps. In the *online* setting, at every time step the new measurement data y_t are revealed sequentially, and only after the controller applies input u_{t-1} or we make a prediction $\hat{y}_t$. Learning and decision making occur and interact simultaneously. The problem of learning to control/predict (1.1) online is also known as the *adaptive control/filtering* problem in the control theoretic literature. In principle, the online learning setting seems superior since the learning algorithm can adapt to changes in the dynamics or the environment. However, deploying adaptive algorithms in practice has proved to be challenging; because of their complexity, they might lead to unpredictable and abrupt behaviors (Anderson & Dehghani, 2007), unless we carefully tune their parameters (Moden, 1995).

Statistical versus adversarial learning. In the *stochastic noise* setting, the process and measurement noises w_k, v_k are randomly generated, typically independent from the control inputs u_k. In this case, the noise does not act against the control/prediction objective but it corrupts the measurement data, reducing learning accuracy. The name statistical learning suggests that we are dealing with stochastic noise. In the *adversarial noise* (adversarial learning) setting, the process and measurement noises w_k, v_k corrupt the process and the measurements to actively obstruct the learning algorithm. In practice, it is a design choice whether we assume that the noise is stochastic or adversarial. Typically, the former choice prioritizes performance, while the latter prioritizes safety.

1.2 Contributions

In this thesis, we focus on the model-based statistical learning paradigm, where the uncertainty is stochastic. Our contributions can be divided into two main themes: i) Statistical Learning of the Kalman Filter and ii) Statistical Difficulty of Learning Linear Systems.

Statistical Learning of the Kalman Filter In the first part of this thesis, we study the finite-sample complexity of learning the Kalman Filter, i.e. we focus on the problem of *prediction*. Prediction is a critical component of decision making. It is also the main task of interest in the case of time-series data, i.e. when $B = 0$ in (1.1).

When system (1.1) is known with known noise statistics, we can employ the celebrated Kalman filter which is the optimal predictor, namely it minimizes the mean square prediction error. The stability and statistical properties of the Kalman filter have been well studied when the system model is known. However, when the model is unknown, we have to learn how to predict from data. Our goal is to learn a data-based predictor that competes with the Kalman Filter that has access to the true underlying model.

Here, we study this scenario, and provide *finite-data prediction guarantees*. We consider both an offline and an online learning architecture. In the offline architecture–see Chapters 3, 4, we perform system identification based on pre-collected data followed by a filter design stage. In the online learning architecture–see Chapter 5 we simultaneously predict and update the model continuously.

Our contributions are the following.

1. **Finite-Sample Analysis of Stochastic Identification**. We perform, to the best of our knowledge, the first finite sample analysis of identifying an unknown autonomous LTI system (1.1), with $B = 0$, also known as the *stochastic system identification* problem (Van Overschee & De Moor, 2012). We provide the first finite-sample guarantees for the estimation of matrices A, C as well as the Kalman filter gain of (1.1). We show that we can achieve a finite-sample statistical rate of $O(1/\sqrt{N})$, up to logarithmic factors, where N is the number of samples. More details can be found Chapter 3.

2. **Offline Learning Guarantees**. We provide, to the best of our knowledge, the first *end-to-end sample complexity bounds* for the Kalman Filtering of an unknown system, in the case of offline learning. In particular, we show that the mean square error between our data-driven predictor and the Kalman Filter with access to the true model is, with high probability, bounded by $O(1/\sqrt{N})$, where N is the number of samples collected in the system identification step. More details can be found Chapter 4.

3. **Online Learning Guarantees**. To capture the finite-sample suboptimality of online

prediction, we adopt the notion of regret (Cesa-Bianchi & Lugosi, 2006). It measures how far our online predictions are from the optimal Kalman Filter predictions that has access to the full system model. We define a notion of regret that has a natural, system theoretic interpretation. The prediction error of an online prediction algorithm is compared against the prediction error of the Kalman filter that has access to the exact model, which is allowed to be arbitrary. We present the first online prediction algorithm with provable *logarithmic regret upper bounds* for the classical Kalman Filter. In fact, we prove that with high probability the regret of our algorithm is of the order of $\tilde{O}(1)$, where $\tilde{O}$ hides poly $\log N$ terms, where N is the number of observations collected so far. More details can be found Chapter 5.

The material is based on our recent works Tsiamis & Pappas (2019); Tsiamis et al. (2020); Tsiamis & Pappas (2020).

Statistical Difficulty of Learning Linear Systems In our second part, we study when learning is statistically easy or hard in the case of fully observed linear systems (1.1) with $y_k = x_k$. In Chapter 6, we study the difficulty of system identification. We define as statistically easy, classes of systems whose finite-sample complexity is polynomial with the system dimension. Most prior research in the finite-sample analysis of fully observed systems falls in this category by assuming system (1.1) is fully excited by the process noise w_k. We define as statistically hard, classes of linear systems whose worst-case sample complexity is at least exponential with the system dimension, regardless of the learning algorithm. Using tools from minimax theory, we show that classes of linear systems which are statistically hard to learn do indeed exist. Such system classes include indirectly excited systems with large degree of indirect excitation, also known as controllability index. As we show in Chapter 6, structural properties like the controllability index can crucially affect learnability, determining whether a problem is hard or not. In Chapter 7, we extend these results to the problem of learning to control linear systems. In summary, our contributions are the following.

1. **Exponential sample complexity of system identification is possible.** We identify classes of under-actuated linear systems whose worst-case sample complexity increases exponentially with the state dimension n regardless of learning algorithm (see Chapter 6). These hardness results hold even for robustly controllable systems.

2. **Stabilization from data can exhibit exponential sample complexity.** We extend the previous results to the problem of stabilizing a system from data (see Chapter 7). We show that certain classes of under-actuated linear systems might exhibit worst-case sample complexity which increases exponentially with the state dimension n regardless of stabilization algorithm.

3. **The regret of the online LQR problem can be exponential** We also study the difficulty of the online learning of the Linear Quadratic Regulator (LQR). We show that the regret of online LQR can scale exponentially with the dimension as $\exp(n)\sqrt{T}$, where T is the number of samples collected so far (see Chapter 7) and n is the state dimension of the system. Once again, this behavior can arise in the case of under-actuated systems, which are in general hard to control.

4. **Controllability index affects sample-complexity/regret.** We prove that under some standard algorithms, the sample complexity of identification is upper-bounded by an exponential function of the system's controllability index (see Chapter 6). Similar results hold for the sample complexity of stabilization and the regret of the online LQR (see Chapter 7). This implies that if the controllability index is small $O(1)$ (with respect to the dimension n), then learning is guaranteed to be easy.

The material of Chapter 6 is based on our recent paper (Tsiamis & Pappas, 2021). At the time of writing this thesis, the material of Chapter 7 was submitted for publication (Tsiamis et al., 2022).

1.3 Related work

Asymptotic analyses Control theory has a long history studying the statistical properties of system identification (Ljung, 1999), adaptive control (Åström & Wittenmark, 1973; Goodwin et al., 1981), and adaptive filtering (Lai & Ying, 1991). Until recently, the main focus was providing guarantees in the *asymptotic regime*, when the number of collected samples N tends to infinity. In the *finite-sample regime*, tools like the Central Limit Theorem, can be only applied heuristically, while any existing rigorous results are conservative (Vershynin, 2018). Besides, under the big-O notation, it is not always clear how the learning performance depends on various system theoretic properties. Some more recent work in the asymptotic regime can be found in Wang & Janson (2021); Lu & Mo (2021).

Finite-sample analysis of system identification The first works on the finite-sample analysis of system identification appeared in the 90s (Dahleh et al., 1993; Poolla & Tikku, 1994; Weyer et al., 1999) and 2000s (Campi & Weyer, 2002; Vidyasagar & Karandikar, 2008). After the papers by Abbasi-Yadkori & Szepesvári (2011) and Dean et al. (2017), there has been a resurgence of interest. Over the past years there have been significant advances in understanding finite sample system identification for both fully-observed systems Simchowitz et al. (2018); Faradonbeh et al. (2018a); Sarkar & Rakhlin (2018); Fattahi et al. (2019); Jedra & Proutiere (2019); Wagenmaker & Jamieson (2020) as well as partially-observed systems Oymak & Ozay (2018); Sarkar et al. (2019); Simchowitz et al. (2019); Tsiamis & Pappas (2019); Lee & Lamperski (2020); Zheng & Li (2020); Lee (2020); Lale et al. (2020b). A tutorial can be found in Matni & Tu (2019). The above approaches offer mainly *data-independent* bounds which reveal how the state dimension n and other system theoretic parameters affect the sample complexity of system identification *qualitatively*. This is different from finite sample data-dependent bounds-see for example bootstrapping Dean et al. (2017) or Carè et al. (2018), which might be more tight and more suitable for applications but do not necessarily reveal this dependence.

Prediction In Tsiamis et al. (2020); Lee & Zhang (2020) the problem of end-to-end prediction guarantees was studied, in an offline learning architecture. Online prediction was studied in the case of systems without internal states (such as ARMA - autoregressive moving average) Anava et al. (2013). Prediction of observations generated by state space models in the case of exogenous inputs and adversarial noise but with a bounded budget was studied in Hazan et al. (2018). Recently, Kozdoba et al. (2019) introduced regret bounds for the Kalman Filter in the restricted context of scalar and bounded observations. The regret is shown to be of the order of $\sqrt{N}$ along with a small linear term. Here, we improve the state of the art to logarithmic bounds for general observations. Concurrently and independently Ghai et al. (2020) also proved logarithmic regret bounds for the Kalman Filter. Our analysis here is different focusing on persistency of excitation, which can also provide simultaneous parameter estimation guarantees. After our work Tsiamis & Pappas (2020), regret bounds were extended to the case where the Kalman Filter closed-loop matrix is close to instability Rashidinejad et al. (2020).

Sample Complexity of Learning Feedback Laws. The sample complexity of learning (stabilizing) feedback laws from data was studied before in the case of stochastic (Dean et al., 2017; Tu et al., 2017; Faradonbeh et al., 2018b; Mania et al., 2019) as well as adversarial (Chen & Hazan, 2021) disturbances. The standard paradigm has been to perform system identification, followed by a robust control or certainty equivalent gain design. Typically, apart from stability, another goal is to also achieve as good control performance as the optimal LQR controller that has access to the true model Dean et al. (2017); Mania et al. (2019). Similar issues were studied in the partially observed setting (Zheng et al., 2021).

Online Control Recently, there have been important results addressing the online learning of the Linear Quadratic Regulator (LQR) problem Abbasi-Yadkori & Szepesvári (2011); Faradonbeh et al. (2020b); Ouyang et al. (2017b); Dean et al. (2018); Mania et al. (2019); Cohen et al. (2019). The best regret for LQR is sublinear and of the order of $\tilde{O}(\sqrt{N})$, where

N is the numbers of state samples collected; an in-depth survey can be found in Matni et al. (2019). It was shown recently that this rate is in fact tight; in the worst case, the regret is indeed of the order of $\Omega(N)$ Simchowitz & Foster (2020); Ziemann & Sandberg (2020). Online control of linear systems has also been studied in the case of more general convex consts Plevrakis & Hazan (2020). When the system model is *known*, then the Kalman filter is the dual of the Linear Quadratic Regulator, suggesting that this duality can be exploited in deriving learning guarantees for the Kalman filter. However, when the system model is *unknown*, the Linear Quadratic Regular and the Kalman filter are not dual problems. As the state is fully observed in LQR, the system identification reduces to a simple least squares problem. In the Kalman filter case, the state is *partially* observed resulting in non-convex system identification problems requiring us to consider a different approach.

A different line of work studies adaptive stabilization of unstable systems in the setting of noiseless (Talebi et al., 2021) or noisy (Faradonbeh et al., 2018b) systems.

Part I

Statistical Learning of the Kalman Filter

Chapter 2

Introduction and Background

The celebrated Kalman Filter (Anderson & Moore, 2005; Kailath et al., 2000) has been a fundamental approach for estimation and prediction of time-series data, with diverse applications ranging from control systems and robotics (Bertsekas, 2017; Durrant-Whyte & Bailey, 2006) to computer vision (Coskun et al., 2017), economics (Harvey, 1990; Bauer & Wagner, 2002), and machine learning. Given a known system model with known noise statistics, the Kalman Filter predicts future observations of a *partially observable* dynamical process by filtering past observations. A well-studied setting is the case of autonomous linear time invariant (LTI) systems driven by Gaussian noise:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k, \end{aligned} \tag{2.1}$$

where x_k is the internal state, y_k are the observations (measurements), w_k is the Gaussian process noise, and v_k is the Gaussian measurement noise. In this setting, the Kalman Filter is optimal in the sense that it minimizes the mean square prediction error of the state/observations. Since Kalman's seminal paper (Kalman, 1960), the stability and statistical properties of the Kalman Filter have been well studied when system (2.1) is *known*.

In reality, in many practical cases of interest (e.g., tracking moving objects, stock price forecasting), the system model (2.1) is *unknown*, and we have to learn how to predict based on data. However, learning to predict unknown partially observable systems is a significantly more challenging problem. Even in the case of linear systems, learning directly the model parameters of the system results in nonlinear, non-convex problems (Yu et al., 2018). Besides, learning the system under finite samples, inevitably introduces parametric errors in model (2.1), which leads to a KF with suboptimal prediction performance (El Ghaoui & Calafiore, 2001).

In this part, we study exactly this scenario, and provide finite-sample prediction guarantees for the Kalman filtering of an unknown autonomous LTI system (2.1). We consider two possible architectures: a) an offline learning scheme, where we perform system identification based on batch data followed by a filter design stage-see Chapters 3, 4; b) an online learning scheme, where we simultaneously predict and update the model continuously–see Chapter 5.

In this chapter, we provide a brief high-level overview of the offline/online learning architectures. We also review for completeness, some well-known properties of the Kalman Filter in Section 2.3.

2.1 Offline learning architecture

In the offline learning architecture-see Fig. 2.1, we have access to batch observation data $y_0, \dots, y_N$ collected offline, where N is the number of data. Based on the collected samples, we learn a Kalman Filter for system (2.1), which predicts the system's state. In this architecture, the learning and the prediction phases are decoupled. As shown in Fig. 2.1, the online data are only used for prediction. For the learning phase, we consider the following simple two step procedure, which has been a standard paradigm in control theory. In the first step, using system identification tools rooted in subspace methods, we obtain finite-data estimates of the state-space parameters, and Kalman gain describing system (2.1). Then, in the second step, we use these approximate parameters to design a filter which predicts the system state.

Note that in this offline architecture, we only perform learning once. As a result, we

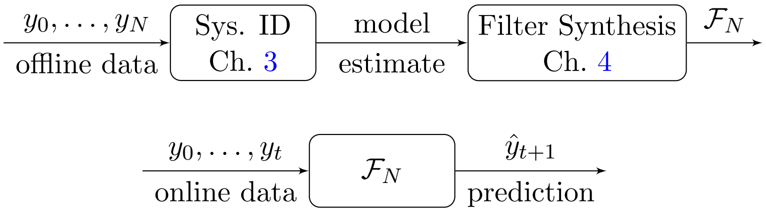


Figure 2.1: **Offline Learning Architecture.** In the offline architecture, we learn a filter $\mathcal{F}_N$ once based on batch offline collected data. First, we learn a model of the system based on system identification, and then we perform filter design to obtain a (suboptimal) filter $\mathcal{F}_N$. Second, we deploy the filter for prediction. The learning and prediction phases are decoupled. The online data are only used for prediction.

cannot adapt to changes in the dynamics or the environment. Hence, offline learning is more suitable for applications where the environment is in steady-state or does not change significantly. Nonetheless, offline learning algorithms are more simple, they are easier to analyze, and they lead to more predictable behaviors when deployed online.

Here, our goal is to provide finite-sample guarantees for the prediction performance of the offline learning architecture. We study the stages of system identification and filter design separately in Chapter 3 and Chapter 4 respectively. The material is based on the publications Tsiamis & Pappas (2019); Tsiamis et al. (2020). Our contributions are summarized as follows. More details can be found in Chapters 3, 4.

i) We perform the first finite-sample analysis of identifying the stochastic system (2.1), also known as the *stochastic system identification* problem (Van Overschee & De Moor, 2012). We provide the first non-asymptotic guarantees for the estimation of matrices A, C as well as the Kalman Filter gain of (2.1).

ii) We analyze the sensitivity to modeling error of the filter design phase. We show that if the system identification step produces sufficiently accurate estimates, or if the underlying true Kalman Filter is sufficiently robust, then the data-based Kalman Filter enjoys near optimal mean square prediction error.

iii) We integrate the above results with the finite-data system identification guarantees, to provide, to the best of our knowledge, the first end-to-end sample complexity bounds for the Kalman filtering of an unknown system. In particular, we show that the mean square

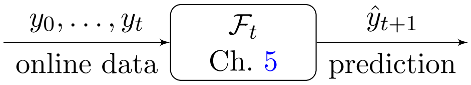


Figure 2.2: **Online Learning Architecture.** In the online architecture, the data are revealed sequentially. Learning and prediction are performed simultaneously and continuously. Since we continuously adapt to the data, the filter $\mathcal{F}_t$ changes with time t. Unlike the offline architecture, the learning data are also used for prediction.

prediction error of the data-based Kalman Filter produced by the two step offline procedure is, with high probability, bounded by $\tilde{O}(1/\sqrt{N})$, where N is the number of samples collected in the system identification step.

2.2 Online learning architecture

In the online learning architecture-see Fig. 2.1, the data are revealed sequentially. At each time step we use the new information to both adapt and predict. Learning and prediction are performed simultaneously and continuously using the same data-set. From a machine learning perspective this problem has been known as the online learning problem, while from a control theoretic perspective this problem has been known as the adaptive (Kalman) filtering problem. Similar to the offline case, our learning algorithm is based system identification techniques, properly modified to account for the online adaptation. The advantage of the online architecture is that the learning algorithm can adapt to changes in the dynamics or the environment. However, online learning algorithms are more complex in general and might be more challenging to analyze.

In Chapter 5, we provide finite-sample guarantees for the prediction performance of the online architecture. The material is based on our preprint Tsiamis & Pappas (2020). To capture the finite-sample suboptimality of online prediction, we adopt the notion of regret Cesa-Bianchi & Lugosi (2006). It measures how far our online predictions are from the optimal Kalman Filter predictions that has access to the full system model. Our goal is to find an online prediction algorithm that has provably small regret. Our technical contributions are (for more details see Chapter 5):

i) System theoretic regret. We define a notion of regret that has a natural, system theoretic interpretation. The prediction error of an online prediction algorithm is compared against the prediction error of the Kalman Filter that has access to the exact model.

ii) Logarithmic regret for the Kalman Filter. We present the first online prediction algorithm with provable logarithmic regret upper bounds. In particular, we prove that with high probability the regret of our algorithm is of the order of $\tilde{O}(1)$, where $\tilde{O}$ hides poly log N terms, where N is the number of observations collected. Our regret guarantees hold for the class of non-explosive systems, which includes marginally stable linear systems.

iii) Learning gap between LQR and Kalman Filter: One of the implications of our bounds is that learning to predict observations like the Kalman Filter is provably easier than solving the online Linear Quadratic Regulator (LQR) problem, which in general requires $\Omega(\sqrt{N})$ regret Simchowitz & Foster (2020); Ziemann & Sandberg (2020). This might not be surprising due to the fact that, in the case of exogenous inputs, we need to inject exploratory signals into the system.

2.3 Kalman Filter preliminaries

Consider system (2.1), where $x_k \in \mathbb{R}^n$ is the system state, $y_k \in \mathbb{R}^m$ is the output, $A \in \mathbb{R}^{n\times n}$ is the system matrix, $C \in \mathbb{R}^{m\times n}$ is the output matrix, $w_k \in \mathbb{R}^n$ is the process noise, and $v_k \in \mathbb{R}^m$ is the measurement noise. The noises w_k, v_k are assumed to be i.i.d. zero mean Gaussian, with covariance matrices Q and R respectively, and independent of each other. The initial state x_0 is also assumed to be zero mean Gaussian, independent of the noises, with covariance Σ_0. Matrices A, C, Q, R, Σ_0 are initially unknown.

Let $\mathcal{F}_k \triangleq \{y_0, \dots, y_k\}$ be the filtration generated by the observations up to time k. Let $\mathcal{L}_2^d(\mathcal{F}_k)$ denote the space of square integrable, $\mathcal{F}_k$-measurable random vectors in $\mathbb{R}^d$, for some dimension d. Then, the minimum mean square error predictor is defined as:

$$\hat{y}_{k+1} \triangleq \arg \min_{z\in\mathcal{L}_2^m(\mathcal{F}_k)} \mathbb{E}\|y_{k+1} - z\|_2^2$$

$$\hat{x}_{k+1} \triangleq \arg \min_{z\in\mathcal{L}_2^n(\mathcal{F}_k)} \mathbb{E}\|x_{k+1} - z\|_2^2$$

In the case of Gaussian noise, the above optimal predictors admit a closed-form expression, which is the celebrated Kalman Filter

$$\begin{aligned}
\hat{x}_{k+1} &= A\hat{x}_k + K_k(y_k - \hat{y}_k), \quad \hat{x}_0 = 0\\
\hat{y}_{k+1} &= C\hat{x}_{k+1}\\
P_{k+1} &= AP_kA^* + Q - AP_kC^*(CP_kC^* + R)^{-1}CPA^*, \quad P_0 = \Sigma_0\\
K_k &= AP_kC^*(CP_kC^* + R)^{-1},
\end{aligned}$$

where K_k is the Kalman Filter gain. Matrix $P_k = \mathbb{E}(x_k - \hat{x}_k)(x_k - \hat{x}_k)^*$ is the state prediction error covariance and satisfies a standard Riccati recursion.

For the Kalman Filter to be stable and well-behaved we typically make the following standard observability assumption.

Assumption 2.1. *The pair (A, C) is detectable, $(A, Q^{1/2})$ is stabilizable and R is strictly positive definite.* ◇

Under the above assumption, the Kalman Filter converges exponentially to its steady-state, in the sense that the second order statistics converge.

Proposition 2.1 (KF convergence Anderson & Moore (2005))**.** *Consider system* (2.1) *under Assumption 2.1. The Kalman gain K_k and the covariance P_k converge exponentially fast to their respective limits, which are well-defined*

$$\lim_{k\to\infty} K_k = K, \quad \lim_{k\to\infty} P_k = P.$$

The steady-state gain is given by

$$K = APC^*\left(CPC^* + R\right)^{-1}, \tag{2.2}$$

while P is the unique stabilizing solution[1] *to*

$$P = (A - KC)P(A - KC)^* + Q + KRK^*. \tag{2.3}$$

As a result, the closed-loop matrix $A - KC$ is stable, i.e. it has spectral radius $\rho(A - KC) < 1$.

The above expressions can be simplified if the initial covariance Σ_0 is chosen such that we start from steady state.

Corollary 2.1 (KF steady-state Anderson & Moore (2005))**.** *Consider system* (2.1) *under Assumption 2.1. Let $\Sigma_0 = P$, where P is the solution to the algebraic Riccati equation* (2.3)*. Define the innovation process:*

$$e_k \triangleq y_k - \hat{y}_k \tag{2.4}$$

Then, the Kalman Filter recursion is simplified to:

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Ke_k, \quad \hat{x}_0 = 0 \\ y_k &= C\hat{x}_k + e_k. \end{aligned} \tag{2.5}$$

We conclude this brief section with some interesting observations about the Kalman Filter and we also highlight two of its fundamental properties.

Remark 1 (equivalent representation)**.** The Kalman Filter (2.5) is also sometimes called predictor form (or innovation form) of system (2.1) in the system identification literature. Interestingly, under Assumption 2.1 and $\Sigma_0 = P$, systems (2.1), (2.5) are indistinguishable from the perspective of the outputs/observations; both systems can produce observations with identical distributions. Hence, we can view (2.5) as an alternative system representation. This implies that if system (2.1) is completely unknown, then we might have multiple systems that explain the same observations. In other words, the representation of the statistics of the observations is not unique, even if we keep A the same.

Remark 2 (KF properties)**.** Consider the steady-state Kalman Filter (2.5). Two of the

[1] A stabilizing solution P to the Riccati equation defines a Kalman gain K such that $\rho(A - KC) < 1$.

properties that make it desirable are i) stability, and ii) the orthogonality principle. Stability of the Kalman Filter follows from observability (Assumption 2.1) and implies that the innovation sequence e_k can reach steady-state. The orthogonality principle states that the innovation process e_k is orthogonal, i.e. $\mathbb{E}e_k e_t^* = 0$ for $k \neq t$. However, by Gaussianity of e_k, this will also imply that the sequence e_k is independent. In addition, since we are at steady state, the innovation e_k is an i.i.d. sequence.

Chapter 3

Finite Sample Analysis of Stochastic System Identification

3.1 Introduction

In this chapter, we study the finite-sample complexity of system identification in the case of stochastic systems. Recall from Chapter 2 that stochastic systems have the following form:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k, \end{aligned} \tag{3.1}$$

namely the state is only driven by noise and we have no exogenous inputs. Our approach will serve as the backbone of our analysis for both the offline and the online architecture. Before we present the problem formulation, let us give a brief historical overview.

Most identification methods (Ljung, 2010) for linear systems either follow the prediction error approach (Ljung, 1999) or the subspace method (Van Overschee & De Moor, 2012; Verhaegen & Verdult, 2007). The prediction error approach is usually non-convex and directly searches over the system parameters A, B, C, D by minimizing a prediction error cost. The subspace approach is a convex one; first, Hankel matrices of the system are estimated, then, the parameters are realized via steps involving singular value decomposition (SVD). Methods inspired by machine learning have also been employed (Chiuso & Pillonetto, 2019). In this thesis, we focus on the *subspace identification* approach–see (Qin, 2006) for an overview.

The asymptotic statistical properties of subspace algorithms have been well-studied in the stationary regime (Deistler et al., 1995; Peternell et al., 1996; Viberg et al., 1997; Jansson & Wahlberg, 1998; Knudsen, 2001; Bauer et al., 1999; Chiuso & Picci, 2004). In (Deistler et al., 1995; Peternell et al., 1996), it is shown that the identification error can decay as fast as $O(1/\sqrt{N})$ up to logarithmic factors, where N is the number of data. In (Bauer et al., 1999; Chiuso & Picci, 2004) Central Limit Theorems for the identification errors are established. The aforementioned results rely on the assumption of asymptotic stability (spectral radius $\rho(A) < 1$) and hold as the number of data N grows to infinity. In the non-stationary case, subspace identification for a subclass of marginally stable systems was considered in Bauer & Wagner (2002), where it is shown that consistency can be guaranteed asymptotically if the unit circle eigenvalues of A are all equal to 1 with simple Jordan blocks.

From a machine learning perspective, finite sample analysis has been a standard tool for comparing algorithms in the non-asymptotic regime. A series of papers Faradonbeh et al. (2018a); Simchowitz et al. (2018); Sarkar & Rakhlin (2018); Fattahi et al. (2019) studied the finite sample properties of system identification from a single trajectory, when the system state is fully observed ($C = I$). Finite sample results for partially observed systems ($C \neq I$), which is a more challenging problem, appeared recently in Oymak & Ozay (2018); Simchowitz et al. (2019); Sarkar et al. (2019). These papers provide a non-asymptotic convergence rate of $1/\sqrt{N}$ (up to logarithmic factors) for the recovery of matrices A, B, C, D up to a similarity transformation. The results rely on the assumption that the system can be driven by external inputs, i.e. $B, D \neq 0$. In Simchowitz et al. (2019), it was shown that consistency can be achieved even for arbitrary marginally stable systems, where $\rho(A) \leq 1$. Sample complexity of prediction error methods has also been considered Weyer et al. (1999); Campi & Weyer (2002); Vidyasagar & Karandikar (2008); Hazan et al. (2018); Hardt et al. (2018), where the main metric is prediction performance. Finite sample properties of system identification algorithms have also been used in robust and adaptive control Dean et al. (2017); Rantzer (2018). The dual problem of Kalman filtering has not been studied yet in this context; preliminary results for scalar observations appeared in Kozdoba et al. (2019).

In this thesis, we perform the first finite sample analysis of identifying system (3.1) in the case $B, D = 0$, when we have no inputs, also known as *stochastic system identification* (SSI) Van Overschee & De Moor (2012). We provide the first non-asymptotic guarantees for the estimation of matrices A, C as well as the Kalman filter gain of (3.1). Similar to Sarkar & Rakhlin (2018); Oymak & Ozay (2018), the analysis is based on new tools from machine learning and statistics Vershynin (2018); Abbasi-Yadkori et al. (2011); Tu et al. (2016). As in Weyer et al. (1999); Campi & Weyer (2002); Vidyasagar & Karandikar (2008); Faradonbeh et al. (2018a); Simchowitz et al. (2018); Sarkar & Rakhlin (2018); Fattahi et al. (2019); Oymak & Ozay (2018); Simchowitz et al. (2019); Sarkar et al. (2019), we focus on data-independent bounds, i.e. bounds which reveal how the identification error depends on the number of data N, and the system's and algorithm's parameters. An alternative approach is to derive data-dependent bounds, see for example Carè et al. (2018). Such bounds could potentially be more tight, however it is not yet clear how they vary with the number of data N. In summary, our main contributions are:

- To the best of our knowledge, we provide the first finite sample upper bounds in the case of stochastic system identification, where we have no inputs and the system is only driven by noise. We also provide the first finite sample guarantees for the estimation error of the Kalman filter gain.

- We prove that the outputs of the system satisfy persistence of excitation in finite time with high probability. This result is fundamental for the analysis of many subspace identification algorithms which use outputs as regressors.

- We show that we can achieve a non-asymptotic learning rate of $O(\sqrt{1/N})$ up to logarithmic factors in the case of general marginally stable systems $\rho(A) = 1$, generalizing the asymptotic results of Bauer & Wagner (2002). The learning rate is also valid in the case of repeated unit circle eigenvalues, when the system is unstable but non-explosive. For stable systems ($\rho(A) < 1$), the non-asymptotic learning rate is consistent with classical asymptotic results Deistler et al. (1995).

3.2 Problem formulation

Consider system (3.1), where $x_k \in \mathbb{R}^n$ is the system state, $y_k \in \mathbb{R}^m$ is the output, $A \in \mathbb{R}^{n\times n}$ is the system matrix, $C \in \mathbb{R}^{m\times n}$ is the output matrix, $w_k \in \mathbb{R}^n$ is the process noise, and $v_k \in \mathbb{R}^m$ is the measurement noise. The noises w_k, v_k are assumed to be i.i.d. zero mean Gaussian, with covariance matrices Q and R respectively, and independent of each other. The initial state x_0 is also assumed to be zero mean Gaussian, independent of the noises, with covariance Σ_0.

Assumption 3.1. *Matrices A, C, Q, R, Σ_0 are initially unknown. The order of the system n is known* [2]. *The spectral radius $\rho(A)$ of A is $\rho(A) \leq 1$. The pair (A, C) is observable, $(A, Q^{1/2})$ is controllable and R is strictly positive definite.* ◇

The assumption $\rho(A) \leq 1$ includes marginally stable systems as well as non-explosive unstable systems with repeated unit circle roots. It is more general than the stricter condition $\rho(A) < 1$ found in previous works, see Deistler et al. (1995); Peternell et al. (1996); Viberg et al. (1997); Jansson & Wahlberg (1998); Knudsen (2001); Bauer et al. (1999); Chiuso & Picci (2004). The remaining conditions in Assumption 3.1, are standard for Kalman filter to be well-behaved–see also Section 2.3. They are slightly stricter compared to the conditions of Assumption 2.1 to guarantee that the system we try to learn is minimal. Next, we also assume that the Kalman filter has reached its steady state.

Assumption 3.2. *We assume that the initial state covariance is equal to the steady-state covariance of the Kalman filter $\Sigma_0 = P$, where P is defined in* (2.3). ◇

Since the Kalman filter converges exponentially fast to the steady-state gain, this assumption is reasonable in many situations; it is also standard Deistler et al. (1995); Knudsen

[2] The results of Section 3.4 do not depend on the order n being known.

(2001). Based on the above assumptions, by Corollary 2.1 the steady-state Kalman filter of system (3.1) takes the following form:

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Ke_k, \quad \hat{x}_0 = 0 \\ y_k &= C\hat{x}_k + e_k, \end{aligned} \tag{3.2}$$

where $\hat{x}_k$ is the Kalman filter predicted state, $e_k = y_k - \hat{y}_k$ is the innovation process, and K is the filter gain, given by (2.2). We denote the covariance matrix of the prediction $\hat{x}_k$ by:

$$\Gamma_k = \mathbb{E}\left[\hat{x}_k \hat{x}_k^*\right]. \tag{3.3}$$

The innovation error sequence e_k has covariance

$$\bar{R} \triangleq \mathbb{E}\left[e_k e_k^*\right] = CPC^* + R. \tag{3.4}$$

Since the original errors are Gaussian i.i.d., by the orthogonality principle the innovation error sequence e_k is also Gaussian and i.i.d.

In the classical stochastic subspace identification problem, the main goal is to identify the Kalman filter parameters A, C, K from output samples $y_0 \dots, y_N$, see for example Chapter 3 of Van Overschee & De Moor (2012). The problem is ill-posed in general since the outputs are invariant under any similarity transformation $\bar{A} = S^{-1}AS$, $\bar{C} = CS$, $\bar{K} = S^{-1}K$. Thus, we can only estimate A, C, K up to a similarity transformation.

Here, we will analyze the finite sample properties of a subspace identification algorithm, which is based on least squares.

Problem 3.1 (Finite Sample Analysis of SSI). *Consider a finite number N of output samples $y_0, \dots, y_{N-1}$, which follow model* (3.1)*, and an algorithm $\mathcal{A}$, which returns estimates $\hat{A}, \hat{C}, \hat{K}$ of the true parameters. Given a confidence level δ provide upper bounds $\epsilon_A(\delta, N)$,*

$\epsilon_C(\delta, N)$, $\epsilon_K(\delta, N)$ such that with probability at least $1-\delta$:

$$\begin{aligned}\left\|\hat{A} - S^{-1}AS\right\|_2 &\leq \epsilon_A(\delta, N)\\ \left\|\hat{C} - CS\right\|_2 &\leq \epsilon_C(\delta, N)\\ \left\|\hat{K} - S^{-1}K\right\|_2 &\leq \epsilon_K(\delta, N),\end{aligned} \tag{3.5}$$

for some invertible matrix S, where $\|\cdot\|_2$ denotes the spectral norm. The bounds ϵ can also depend on the model parameters n, A, C, R, Q as well as the identification algorithm used. $\diamond$

3.3 Subspace Identification Algorithm

The procedure of estimating the parameters A, C, K is based on a least squares approach, see for example Deistler et al. (1995); Knudsen (2001). It involves two stages. First, we regress future outputs to past outputs to obtain a Hankel-like matrix, which is a product of an observability and a controllability matrix. Second, we perform a balanced realization step, similar to the Ho-Kalman algorithm, to obtain estimates for A, C, K.

Before describing the algorithm, we need some definitions. Let p, f, with $p, f \geq n$ be two design parameters that define the horizons of the past and the future respectively. Assume that the total number of output samples is $\bar{N} = N + p + f - 1$. Then, the future outputs $Y_k \in \mathbb{R}^{mf}$ and past outputs $Z_k \in \mathbb{R}^{mp}$ at time $k \geq p$ are defined as follows:

$$Y_k \triangleq \begin{bmatrix} y_k \\ \vdots \\ y_{k+f-1} \end{bmatrix}, \quad Z_k \triangleq \begin{bmatrix} y_{k-p} \\ \vdots \\ y_{k-1} \end{bmatrix}, k \geq p \tag{3.6}$$

By stacking the outputs for all sample sequences, over all times $p \leq k \leq N + p - 1$, we form the batch outputs:

$$\bar{Y} \triangleq \left[\begin{array}{ccc} Y_p & \dots & Y_{N+p-1} \end{array} \right],$$
$$\bar{Z} \triangleq \left[\begin{array}{ccc} Z_p & \dots & Z_{N+p-1} \end{array} \right],$$

The past and future noises E_k, E_k^+ and the respective batch noises $\bar{E}$, $\bar{E}_+$ are defined similarly as

$$\begin{array}{ll} E_k^+ \triangleq \left[\begin{array}{c} e_k \\ \vdots \\ e_{k+f-1} \end{array} \right], & E_k \triangleq \left[\begin{array}{c} e_{k-p} \\ \vdots \\ e_{k-1} \end{array} \right], \ k \geq p \\ \bar{E}_+ \triangleq \left[\begin{array}{ccc} E_p^+ & \dots & E_{N+p-1}^+ \end{array} \right], & \bar{E} \triangleq \left[\begin{array}{ccc} E_p & \dots & E_{N+p-1} \end{array} \right] \end{array}, \tag{3.7}$$

where we hide the dependence on the past and future horizons $p\,f$, as well as the dependence on the number of samples N. Next, define the batch states:

$$\bar{X} \triangleq \left[\begin{array}{ccc} \hat{x}_0 & \dots & \hat{x}_{N-1} \end{array} \right].$$

The (extended) observability matrix $\mathcal{O}_k \in \mathbb{R}^{mk \times n}$ and the reversed (extended) controllability matrix $\mathcal{K}_k \in \mathbb{R}^{n \times mk}$ associated to system (3.2) are defined as:

$$\mathcal{O}_k \triangleq \left[\begin{array}{cccc} C^* & A^*C^* & \cdots & (A^*)^{k-1}C^* \end{array} \right]^*, \tag{3.8}$$

$$\mathcal{K}_k \triangleq \left[\begin{array}{cccc} (A-KC)^{k-1}K & \dots & (A-KC)K & K \end{array} \right] \tag{3.9}$$

respectively. We denote the Hankel-like matrix $\mathcal{O}_f\mathcal{K}_p$ by:

$$G \triangleq \mathcal{O}_f\mathcal{K}_p. \tag{3.10}$$

Finally, for any $s \geq 2$, define the block-Toeplitz matrix:

$$\mathcal{T}_s \triangleq \begin{bmatrix} I_m & 0 & & 0 \\ CK & I_m & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ CA^{s-2}K & CA^{s-3}K & \cdots & I_m \end{bmatrix}. \tag{3.11}$$

Based on the definition of the Toeplitz matrix, we can define the covariance matrices of the weighted past and future noises in a compact way:

$$\Sigma_E^+ \triangleq \mathbb{E}\left(\mathcal{T}_f E_k^+ E_k^{+*} \mathcal{T}_f^*\right) = \mathcal{T}_f \operatorname{diag}(\bar{R}, \dots, \bar{R}) \mathcal{T}_f^* \tag{3.12}$$

$$\Sigma_E \triangleq \mathbb{E}\left(\mathcal{T}_p E_k E_k^* \mathcal{T}_p^*\right) = \mathcal{T}_p \operatorname{diag}(\bar{R}, \dots, \bar{R}) \mathcal{T}_p^*. \tag{3.13}$$

3.3.1 Regression for Hankel Matrix Estimation

First, we establish a linear relation between the future and past outputs. From (3.2), for every k, the future outputs can be written as a linear combination of the initial predicted state $\hat{x}_k$ at time k and the future noises:

$$Y_k = \mathcal{O}_f \hat{x}_k + \mathcal{T}_f E_k^+.$$

Meanwhile, from (3.2), the predicted state $\hat{x}_k$ can be expressed in terms of the past outputs:

$$\hat{x}_k = K y_{k-1} + \cdots + (A - KC)^{p-1} K y_{k-p} + (A - KC)^p \hat{x}_{k-p}.$$

Combining the above expressions in their batch form, we arrive at

$$\bar{Y} = G\bar{Z} + \mathcal{O}_f (A - KC)^p \bar{X} + \mathcal{T}_f \bar{E}_+, \tag{3.14}$$

which expresses the future outputs as a linear combination of past outputs, the initial predicted states, and the future noises. Note that the regressors $\bar{Z}$ and the residuals $\bar{E}_+$ are independent from each other column-wise. Hence, equation (3.14) resembles the standard linear regression setting. However, the term $\mathcal{O}_f(A-KC)^p\bar{X}$ introduces a bias due to the Kalman filter truncation, where we use only p past outputs instead of all of them. Based on (3.14), we compute the least squares estimate

$$\hat{G} = \bar{Y}\bar{Z}^*(\bar{Z}\bar{Z}^*)^{-1}. \tag{3.15}$$

The Hankel matrix G can be interpreted as a (truncated) Kalman filter which predicts future outputs directly from past outputs, independently of the internal state-space representation Van Overschee & De Moor (2012). In this sense, the estimate $\hat{G}$ is a "data-driven" Kalman filter. Notice that persistence of excitation of the outputs (invertibility of $\bar{Z}\bar{Z}^*$) is required in order to compute the least squares estimate $\hat{G}$.

3.3.2 Balanced Realization

This step determines a balanced realization of the state-space, which is only one of the possibly infinite state-space representations–see Section 3.6 for comparison with other subspace methods. First, we compute a rank-n factorization of the full rank matrix $\hat{G}$. Let the SVD of $\hat{G}$ be:

$$\hat{G} = \begin{bmatrix} \hat{U}_1 & \hat{U}_2 \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_1 & 0 \\ 0 & \hat{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \hat{V}_1^* \\ \hat{V}_2^* \end{bmatrix}, \tag{3.16}$$

where $\hat{\Sigma}_1 \in \mathbb{R}^{n\times n}$ contains the $n-$largest singular values. Then, a standard realization of $\mathcal{O}_f$, $\mathcal{K}_p$ is:

$$\hat{\mathcal{O}}_f = \hat{U}_1\hat{\Sigma}_1^{1/2},\ \hat{\mathcal{K}}_p = \hat{\Sigma}_1^{1/2}\hat{V}_1^*. \tag{3.17}$$

This step assumed knowing the order n of the system, see Assumption 3.1. In addition, matrix $\mathcal{K}_p$ should have full rank n. This is equivalent to the pair (A, K) being controllable. Otherwise, $\mathcal{O}_f\mathcal{K}_p$ will have rank less than n making it impossible to accurately estimate $\mathcal{O}_f$.

Assumption 3.3. *The pair (A, K) is controllable.* ◇

The above assumption is standard–see for example Knudsen (2001).

Based on the estimated observability/controllability matrices, we can approximate the system parameters as follows:

$$\hat{C} = \hat{\mathcal{O}}_f\left(1:m,:\right), \quad \hat{K} = \hat{\mathcal{K}}_p\left(:,(p-1)m+1:pm\right),$$

where the notation $\hat{\mathcal{O}}_f\left(1:m,:\right)$ means we pick the first m rows and all columns. The notation for $\hat{\mathcal{K}}_p$ has similar interpretation. For simplicity, define

$$\hat{\mathcal{O}}_f^u \triangleq \hat{\mathcal{O}}_f\left(1:m(f-1),:\right),$$

which includes the $m(f-1)$ "upper" rows of matrix $\hat{\mathcal{O}}_f$. Similarly, we define the lower part $\hat{\mathcal{O}}_f^l$. For matrix A we exploit the structure of the extended observability matrix and solve $\hat{\mathcal{O}}_f^u\hat{A} = \hat{\mathcal{O}}_f^l$ in the least squares sense by computing

$$\hat{A} = \left(\hat{\mathcal{O}}_f^u\right)^{\dagger}\hat{\mathcal{O}}_f^l,$$

where † denotes the pseudoinverse.

The finite sample analysis of the above algorithm is divided in two parts. First, in Section 3.4, we provide high probability upper bounds for the error $\|G-\hat{G}\|_2$ in the regression step. Then, in Section 3.5, we analyze the robustness of the balanced realization step.

3.4 Finite Sample Analysis of Regression

In this section, we provide the finite sample analysis of the linear regression step of the identification algorithm. We provide high-probability upper bounds for the estimation error $\|G-\hat{G}\|_2$ of the Hankel-like matrix G. Before we state the main result, recall the definition of the covariance matrix $\bar{R}$ in (3.4). Recall the definition of the past noises' weighted covariance

$$\Sigma_E = \mathbb{E}\left[\mathcal{T}_p E_k E_k^* \mathcal{T}_p^*\right] = \mathcal{T}_p \operatorname{diag}(\bar{R},\ldots,\bar{R})\mathcal{T}_p^*.$$

The least singular value of the above matrix is denoted by:

$$\sigma_E \triangleq \sigma_{\min}\left(\Sigma_E\right). \tag{3.18}$$

Lemma 3.3 in Section 3.7 proves that $\sigma_E \geq \sigma_{\min}\left(R\right) > 0$.

Theorem 3.1 (Regression Step). *Consider system* (3.2) *and let Assumptions 3.1, 3.2, 3.3 be in effect. Let $\hat{G}$ be the estimate* (3.15) *of the subspace identification algorithm given an output trajectory $y_0, \dots, y_{N+p+f-1}$ and let G be as in* (3.10)*. Fix a confidence $\delta > 0$ and define:*

$$\delta_N \triangleq (2(N+p-1)m)^{-\log^2(2pm)\log(2(N+p-1)m)}. \tag{3.19}$$

There exist N_0, N_1, N_2 such that if $N \geq N_0, N_1, N_2$, (see definitions (3.37), (3.41), (3.44) *in Section 3.7), then with probability at least $1-\delta_N-6\delta$:*

$$\|G-\hat{G}\|_2 \leq \underbrace{\mathcal{C}_1\sqrt{\frac{fmp}{N}\log\frac{5f\kappa_N}{\delta}}}_{O\left(\sqrt{p\log N/N}\right)} + \underbrace{\mathcal{C}_2\left\|(A-KC)^p\right\|_2}_{O(\rho(A-KC)^p)}, \tag{3.20}$$

where

$$\kappa_N = \frac{4}{\sigma_E}\left(\|\mathcal{O}_p\|_2^2 \operatorname{tr}\Gamma_{N-1} + \operatorname{tr}\Sigma_E\right) + \delta \tag{3.21}$$

over-approximates the condition number of $\mathbb{E}\left[\bar{Z}\bar{Z}^*\right]$ and

$$\mathcal{C}_1 = 8\sqrt{\frac{\|\Sigma_E^+\|_2}{\sigma_E}}, \quad \mathcal{C}_2 = 4\|\mathcal{O}_f\|_2\|\mathcal{O}_p^\dagger\|_2, \tag{3.22}$$

are system-dependent constants. ◇

Remark 3 (Interpretation). From (3.14), (3.15) the estimation error consists of two terms:

$$\hat{G}-G = \underbrace{\mathcal{T}_f\bar{E}_+\bar{Z}^*\left(\bar{Z}\bar{Z}^*\right)^{-1}}_{\text{Cross term}} + \underbrace{\mathcal{O}_f\left(A-KC\right)^p\bar{X}\bar{Z}^*\left(\bar{Z}\bar{Z}^*\right)^{-1}}_{\text{Kalman filter truncation bias term}}. \tag{3.23}$$

The first term in (3.20) corresponds to the cross-term error, while the second term corresponds to the Kalman filter truncation bias term. To obtain consistency for $\hat{G}$, we have to let the term $\|(A-KC)^p\|_2$ go to zero with N. Recall that the matrix $A-KC$ has spectral radius less than one, thus, the second term decreases exponentially with p. By selecting $p=\beta\log N$, for some sufficiently large β, we can force the Kalman truncation error term to decrease at least as fast as the first one, see for example Deistler et al. (1995). In this sense, the dominant term is the first one, i.e. the cross-term–this is formalized in the following corollary. Notice that f can be kept bounded as long as it is larger than n. ◇

The following result directly follows from Theorem 3.1.

Corollary 3.1 (Consistency)**.** *Consider the conditions of Theorem 3.1 and the definition of δ_N in (3.19). Fix a confidence $\delta>0$ and let $\rho>\rho(A-KC)$. Select*

$$p=\beta\log N,\ \beta>-1/2\frac{1}{\log\rho} \tag{3.24}$$

If $N\geq N_0,N_1,N_2$, (see definitions (3.37), (3.41), (3.44) in Section 3.7), then with probability at least $1-\delta_N-6\delta$:

$$\left\|G-\hat{G}\right\|_2\leq\sqrt{\frac{\|\Sigma_E^+\|_2}{\sigma_E}}\sqrt{fmp}\tilde{O}\left(\sqrt{\frac{\log 1/\delta}{N}}\right), \tag{3.25}$$

where $\tilde{O}$ hides logarithmic terms of N, constants, and other system parameters. ◇

The condition (3.24) guarantees that if we select a sufficiently large β, then the truncation term will decay as fast as $\|(A-KC)^p\|_2=o(1/\sqrt{N})$, i.e. the truncation error will decay faster than the statistical error.

Remark 4 (Statistical rates)**.** For **marginally stable** systems or non-explosive unstable systems ($\rho(A)=1$) and $p=\beta\log N$, we have $\log\kappa_N=O(\log N)$, since $\|\mathcal{O}_p\|_2, \operatorname{tr}\Gamma_N$ depend at most polynomially on p,N. In this case, (3.20) results in a rate of:

$$\|G-\hat{G}\|_2=O\left(\frac{\log N}{\sqrt{N}}+\sqrt{\frac{\log N}{N}\log\frac{1}{\delta}}\right).$$

To the best of our knowledge, there have not been any bounds for subspace algorithms in the general case of $\rho(A)=1$.

In the case of **asymptotically stable** systems ($\rho(A)<1$), we have $\kappa_N=O\left(p\right)$, since $\left\|\mathcal{O}_p\right\|_2, \operatorname{tr}\Gamma_N, \left\|\mathcal{T}_p\right\|_2$ are now $O(1)$. Hence, if $p=c\log N$, we obtain a rate of:

$$\left\|G-\hat{G}\right\|_2=O\left(\sqrt{\frac{\log N\log\log N}{N}}+\sqrt{\frac{\log N}{N}\log\frac{1}{\delta}}\right).$$

As a result, our finite sample bound (3.20) resembles the asymptotic bound in equation (14) of Deistler et al. (1995). ◇

In the absence of inputs ($B, D=0$), the noise both helps and obstructs identification. Larger noise leads to better excitation of the outputs, but also worsens the convergence of the least squares estimator. To see how our finite sample bounds capture that, observe that large noise leads to more excitation σ_E but also to larger future noise $\|\Sigma_E^+\|_2$. This trade-off is captured by the condition number of the noise $\mathcal{C}_1$.

If N is sufficiently large (condition $N\geq N_0, N_1$), the outputs are guaranteed to be persistently exciting in finite time; more details can be found in Sections 3.4.1, 3.7. Meanwhile, condition $N\geq N_2$ is not necessary; it just leads to a simplified expression for the bound of the Kalman filter truncation error–see Sections 3.4.3, 3.7. The definitions of N_0, N_1, N_2 can be found in (3.37), (3.41), (3.44). Their existence is guaranteed even if p varies slowly with N, i.e. logarithmically.

Obtaining the bound on the error $\|G-\hat{G}\|_2$ in (3.20) of Theorem 3.1 requires the following three steps:

1. Proving persistence of excitation (PE) for the past outputs, i.e. invertibility of $\bar{Z}\bar{Z}^*$.

2. Establishing bounds for the cross-term error in (3.23).

3. Establishing bounds for the the truncation term in (3.23).

In the following subsections, we sketch the proof steps.

3.4.1 Persistence of Excitation in Finite Time

The next theorem shows that with high probability the past outputs and noises are persistently exciting in finite time. The result is of independent interest and is fundamental since many subspace algorithms use past outputs as regressors.

Theorem 3.2 (Persistence of Excitation)**.** *Consider the conditions of Theorem 3.1 and* N_0*,* N_1 *as in* (3.37)*,* (3.41)*. If* $N \geq N_0, N_1$*, then with probability at least* $1 - \delta_N - 2\delta$ *both of the following events occur:*

$$\mathcal{E}_{PE} = \left\{ \bar{Z}\bar{Z}^* \succeq \frac{1}{2}\mathcal{O}_p\bar{X}\bar{X}^*\mathcal{O}_p^* + \frac{1}{2}\mathcal{T}_p\bar{E}\bar{E}^*\mathcal{T}_p^* \right\} \tag{3.26}$$

$$\mathcal{E}_E = \left\{ \mathcal{T}_p\bar{E}\bar{E}^*\mathcal{T}_p^* \succeq \frac{N}{2}\Sigma_E \right\}, \tag{3.27}$$

where $\succeq$ *denotes comparison in the positive semidefinite cone. Hence, with probability at least* $1 - \delta_N - 2\delta$ *the outputs satisfy the PE condition:*

$$\bar{Z}\bar{Z}^* \succeq \frac{N}{4}\sigma_E I_{mp},$$

where $\sigma_E > 0$ *is defined in* (3.18)*.* ◇

The above result implies that if the past noises satisfy a PE condition, then PE for the outputs is also guaranteed; the noises are the only way to excite the system in the absence of control inputs. The see why the outputs are persistently exciting, notice that the past output correlations satisfy:

$$\begin{aligned} \bar{Z}\bar{Z}^* =& \mathcal{O}_p\bar{X}\bar{X}^*\mathcal{O}_p^* + \mathcal{T}_p\bar{E}\bar{E}^*\mathcal{T}_p^* + \\ & \mathcal{O}_p\bar{X}\bar{E}^*\mathcal{T}_p^* + \mathcal{T}_p\bar{E}\bar{X}^*\mathcal{O}_p^*. \end{aligned} \tag{3.28}$$

We can first show PE for the noise correlations $\mathcal{T}_p\bar{E}\bar{E}^*\mathcal{T}_p^*$, i.e. show that the event $\mathcal{E}_E$ occurs with high probability when N is sufficiently large (condition $N \geq N_0$). This behavior is due to the fact that $\mathbb{E}\left[\mathcal{T}_p\bar{E}\bar{E}^*\mathcal{T}_p^*\right] = N\Sigma_E$ and the sequence $\bar{E}_k$ is component-wise i.i.d.

To prove this step, we use Lemma C.2 from Oymak & Ozay (2018)–see Lemma 3.2 in Section 3.7.

Meanwhile, the cross terms $\bar{X}\bar{E}^*$ are much smaller and their norm increases with a rate of at most $O(\sqrt{N})$ up to logarithmic terms. This is since $\mathbb{E}\left[\bar{X}\bar{E}^*\right] = 0$ and the product $\bar{X}\bar{E}^*$ has martingale structure (see Section 3.7 and Theorem 3.3 below). Eventually, if the number of samples N is large enough (condition $N \geq N_1$), the cross-terms will be dominated by the noise and state correlations with high probability, which establishes output PE.

3.4.2 Cross-term error

To bound the cross-term error, we express it as a product of $\bar{E}_+\bar{Z}^*\left(\bar{Z}\bar{Z}^*\right)^{-1/2}$, $(\bar{Z}\bar{Z}^*)^{-1/2}$, as in Sarkar & Rakhlin (2018). The second term of the product can be bounded by applying Theorem 3.2. The first term is self-normalized and has martingale structure component-wise. In particular, the product $\bar{Z}\bar{E}_+^*$ is equal to:

$$\bar{Z}\bar{E}_+^* = \begin{bmatrix} \sum_{k=p}^{N+p-1} Z_k e_k^* & \dots & \sum_{k=p}^{N+p-1} Z_k e_{k+f-1}^* \end{bmatrix},$$

where every sum above is a martingale. To bound it, we apply the next theorem, which generalizes Theorem 1 in Abbasi-Yadkori et al. (2011) and Proposition 8.2 in Sarkar & Rakhlin (2018).

Theorem 3.3 (Cross terms)**.** *Let $\{\mathcal{F}_t\}_{t=0}^{\infty}$ be a filtration. Let $\eta_t \in \mathbb{R}^m$, $t \geq 0$ be $\mathcal{F}_t$-measurable, independent of $\mathcal{F}_{t-1}$. Suppose also that η_t has independent components $\eta_{t,i}$ $i = 1, \dots, m$, which are $1-$sub-Gaussian:*

$$\mathbb{E}\left[e^{\lambda \eta_{t,i}} | \mathcal{F}_{t-1}\right] = \mathbb{E}\left[e^{\lambda \eta_{t,i}}\right] \leq e^{\lambda^2/2}, \textit{ for all } \lambda \in \mathbb{R}.$$

Let $X_t \in \mathbb{R}^d$, $t \geq 0$ be $\mathcal{F}_{t-1}-$measurable. Assume that V is a $d \times d$ positive definite matrix.

For any $t \geq 0$, define:

$$\bar{V}_t = V + \sum_{s=1}^{t} X_s X_s^*, \qquad S_t = \sum_{s=1}^{t} X_s H_s^*,$$

where

$$H_s^* = \begin{bmatrix} \eta_s^* & \dots & \eta_{s+r-1}^* \end{bmatrix} \in \mathbb{R}^{rm},$$

for some integer r. Then, for any $\delta > 0$, with probability at least $1 - \delta$, for all $t \geq 0$

$$\left\| \bar{V}_t^{-1/2} S_t \right\|_2^2 \leq 8r \left(\log \frac{r5^m}{\delta} + \frac{1}{2} \log \det \bar{V}_t V^{-1} \right). \qquad \diamond$$

The above theorem along with a Markov upper bound on $\bar{Z}\bar{Z}^*$ (see Lemma 3.4 in Section 3.7) are used to bound $\bar{E}_+ \bar{Z}^* \left(\bar{Z}\bar{Z}^* \right)^{-1/2}$.

3.4.3 Kalman truncation error

For the Kalman truncation error term, we need to bound the term $\bar{X}\bar{Z}^*(\bar{Z}\bar{Z}^*)^{-1}$, which is $O(1)$. Using the identities $\mathcal{O}_p^\dagger \mathcal{O}_p \bar{X} = \bar{X}$, and $\bar{Z} = \mathcal{O}_p \bar{X} + \mathcal{T}_p \bar{E}$, we derive the following equality:

$$\begin{aligned} \bar{X}\bar{Z}^*(\bar{Z}\bar{Z}^*)^{-1} = \mathcal{O}_p^\dagger \big(I_{mp} - \mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^* (\bar{Z}\bar{Z}^*)^{-1} \\ - \mathcal{T}_p \bar{E}\bar{X}^* \mathcal{O}_p^* (\bar{Z}\bar{Z}^*)^{-1} \big) \end{aligned} \tag{3.29}$$

From Theorem 3.2, we obtain $\left\| \mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \leq 2$. The last term in (3.29) can be treated like the cross-term in Section 3.4.2, by applying Theorems 3.2, 3.3 and Lemma 3.4. It decreases with a rate of $O\left(1/\sqrt{N}\right)$ up to logarithmic terms, so it is much smaller than the other terms in (3.29). To keep the final bound simple, we select N_2 such that

$$\left\| \mathcal{T}_p \bar{E}\bar{X}^* \mathcal{O}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \leq 1 \tag{3.30}$$

with high probability–see also (3.44) for the definition of N_2.

3.5 Robustness of Balanced Realization

In this section, we analyze the robustness of the balanced realization. In particular, we upper bound the estimation errors of matrices A, C, K in terms of the estimation error $\|G - \hat{G}\|_2$ obtained by Theorem 3.1.

Assume that we knew G exactly. Then, the SVD of the true G, would be:

$$G = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} = U_1 \Sigma_1 V_1^*,$$

for some $\Sigma_1 \in \mathbb{R}^{n \times n}$. Hence, if we knew G exactly, the output of the balanced realization would be:

$$\bar{\mathcal{O}}_f = U_1 \Sigma_1^{1/2},\ \bar{\mathcal{K}}_p = \Sigma_1^{1/2} V_1^*. \tag{3.31}$$

The respective matrices $\bar{C}, \bar{K}, \bar{A}$ are defined similarly, based on $\bar{\mathcal{O}}_f$, $\bar{\mathcal{K}}_p$, as described in Section 3.3. The original matrices $\mathcal{O}_f, \mathcal{K}_p$ and $\bar{\mathcal{O}}_f, \bar{\mathcal{K}}_p$ are equivalent up to the similarity transformation $\mathcal{O}_f S = \bar{\mathcal{O}}_f$, $S^{-1}\mathcal{K}_p = \bar{\mathcal{K}}_p$ where

$$S \triangleq \mathcal{O}_f^{\dagger} \bar{\mathcal{O}}_f. \tag{3.32}$$

The system matrices $\bar{C}, \bar{K}, \bar{A}$ are also equivalent to the original matrices C, K, A up to the same similarity transformation $\bar{C} = CS$, $\bar{K} = S^{-1}K$, $\bar{A} = S^{-1}AS$, with S defined as above.

For simplicity, we will quantify the estimation errors in terms of the similar $\bar{A}, \bar{C}, \bar{K}$ instead of the original A, C, K. The next result follows the steps of Oymak & Ozay (2018) and relies on Lemma 5.14 of Tu et al. (2016) and Theorem 4.1 of Wedin (1973). Let $\sigma_n(\cdot)$ denote the $n-$th largest singular value.

Theorem 3.4 (Realization robustness). *Consider the true Hankel-like matrix G defined in (3.10) and the noisy estimate $\hat{G}$ defined in (3.15). Let $\hat{A}, \hat{C}, \hat{K}, \hat{\mathcal{O}}_f$, $\hat{\mathcal{K}}_p$ be the output of the balanced realization algorithm based on $\hat{G}$. Let $\bar{A}, \bar{C}, \bar{K}, \bar{\mathcal{O}}_f$, $\bar{\mathcal{K}}_p$ be the output of the balanced realization algorithm based on the true G. If G has rank n and the following* *robustness condition is satisfied:*

$$\left\|\hat{G}-G\right\|_2 \leq \frac{\sigma_n\left(G\right)}{4}, \tag{3.33}$$

then there exists an orthonormal matrix $T \in \mathbb{R}^{n\times n}$ *such that:*

$$\begin{aligned}
&\left\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\right\|_2 \leq 2\sqrt{\frac{10n}{\sigma_n\left(G\right)}}\left\|G-\hat{G}\right\|_2\\
&\left\|\hat{C}-\bar{C}T\right\|_2 \leq \left\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\right\|_2\\
&\left\|\hat{A}-T^*\bar{A}T\right\|_2 \leq \underbrace{\frac{\sqrt{\|G\|_2}+\sigma_o}{\sigma_o^2}}_{O(1)}\left\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\right\|_2\\
&\left\|\hat{K}-T^*\bar{K}\right\|_2 \leq 2\sqrt{\frac{10n}{\sigma_n\left(G\right)}}\left\|G-\hat{G}\right\|_2,
\end{aligned}$$

where $\sigma_o = \min\left(\sigma_n\left(\hat{\mathcal{O}}_f^u\right), \sigma_n\left(\bar{\mathcal{O}}_f^u\right)\right)$*. The notation* $\hat{\mathcal{O}}_f^u, \bar{\mathcal{O}}_f^u$*, refers to the upper part of the respective matrix (first* $(f-1)m$ *rows)–see Section 3.3.2.* ◇

Remark 5. The result states that if the error of the regression step is small enough, then the realization is robust. The singular value $\sigma_n\left(G\right)$ can be quite small. Hence, the robustness condition (3.33) can be restrictive in practice. However, such a condition is a fundamental limitation of the SVD procedure; it guarantees that the singular vectors related to small singular values of G are separated from the singular vectors coming from the noise $G-\hat{G}$, which can be arbitrary. See also Wedin's theorem Wedin (1972). Such robustness conditions have also appeared in model reduction theory Pernebo & Silverman (1982). ◇

The term $\frac{\sqrt{\|G\|_2}+\sigma_o}{\sigma_o^2}$ which appears in the bound of A is $O\left(1\right)$. Although, the value of σ_o^{-1} is random and depends on $\hat{\mathcal{O}}_f^h$, we could replace it by a deterministic bound. From

$$\sigma_n\left(\hat{\mathcal{O}}_f^h\right) \geq \sigma_n\left(\bar{\mathcal{O}}_f^h\right) - \left\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\right\|_2,$$

σ_o will eventually be lower bounded by $\sigma_n\left(\bar{\mathcal{O}}_f^h\right)/2$ if the error $\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\|_2$ is small enough. The norm $\|G\|_2 \leq \|\mathcal{O}_f\|_2\,\|\mathcal{K}_p\|_2$ is upper bounded for all p, since $A-KC$ is asymptotically stable and f is fixed.

Remark 6 (Total bounds)**.** The final upper bounds for the estimation of the system parameters A, C, K, as stated in Problem 3.1, can be found by combining the finite sample guarantees of the regression step (Theorem 3.1) with the robustness analysis of the realization step (Theorem 3.4). All matrix estimation errors depend linearly on the Hankel matrix estimation error $\|G - \hat{G}\|_2$. As a result, all matrix errors have the same statistical rate as the error of G, i.e. their estimation error decreases at least as fast as $O\left(1/\sqrt{N}\right)$ up to logarithmic factors. ◇

3.6 Open Problems

One of the differences between the subspace algorithm considered in our work and other subspace identification algorithms is the SVD step. Other algorithms perform SVD on $W_1 G W_2$ instead of G, where W_1, W_2 are full rank weighting matrices, usually data dependent Van Overschee & De Moor (1995); Ljung (1999); Van Overschee & De Moor (2012). From this point of view, the results of Section 3.4 (upper bound for $\|G - \hat{G}\|$ in Theorem 3.1 and persistence of excitation in Theorem 3.2) are fundamental for understanding the finite sample properties of other subspace identification algorithms. Here, we studied the case $W_1 = I, W_2 = I$, which is not the standard choice Knudsen (2001). Our results can be extended to the case where $W_2 = (\bar{Z}\bar{Z}^*)^{1/2}/\sqrt{N}$, which is a variation of the MOESP algorithm Qin (2006). Since we proved persistency of excitation, matrix W_2 will be well-behaved and all steps of our analysis carry through with minor changes. It is subject of future work to formally treat the above choice as well as explore how other choices of W_1, W_2 affect the realization step, especially the robustness condition of the SVD step.

3.7 Proofs

3.7.1 Proof of Theorem 3.3

Let us first state a result which follows from the arguments of Chapter 4 of Vershynin (2018). See also Proposition 8.1 of Sarkar & Rakhlin (2018).

Proposition 3.1 (Vershynin (2018))**.** *Consider a matrix $M \in \mathbb{R}^{m\times n}$. Let S^{n-1} denote the unit sphere. Then for any $\epsilon > 0$:*

$$\mathbb{P}\left(\|M\|_2 \geq t\right) \leq \left(1+\frac{2}{\epsilon}\right)^n \max_{x\in S^{n-1}} \mathbb{P}\left(\|Mx\|_2 \geq t\left(1-\epsilon\right)\right)$$

◇

Second, we state Theorem 1 of Abbasi-Yadkori et al. (2011), which upper bounds self-normalized martingales.

Theorem 3.5 (Theorem 1 in Abbasi-Yadkori et al. (2011))**.** *Let $\{\mathcal{F}_t\}_{t=0}^{\infty}$ be a filtration. Let $\eta_t \in \mathbb{R}$, $t \geq 0$ be real-valued, $\mathcal{F}_t$-measurable, and conditionally $1-$sub-Gaussian:*

$$\mathbb{E}\left[e^{\lambda\eta_t}|\mathcal{F}_{t-1}\right] \leq e^{\lambda^2/2},\ \textit{for all}\ \lambda \in \mathbb{R}. \tag{3.34}$$

Let $X_t \in \mathbb{R}^d$, $t \geq 0$ be vector valued and $\mathcal{F}_{t-1}-$measurable. Assume that V is a $d \times d$ positive definite matrix. For any $t \geq 0$, define:

$$\bar{V}_t = V + \sum_{s=1}^{t} X_s X_s^*, \qquad S_t = \sum_{s=1}^{t} \eta_s X_s.$$

Then, for any $\delta > 0$, with probability at least $1-\delta$, for all $t \geq 0$

$$\left\|\bar{V}_t^{-1/2} S_t\right\|_2^2 \leq 2\log\frac{\det\left(\bar{V}_t\right)^{1/2}\det\left(V\right)^{-1/2}}{\delta} \tag{3.35}$$

◇

Finally, we state a standard linear algebra result. We include the proof for completeness.

Lemma 3.1 (Block Matrix Norm)**.** *Assume*

$$M = \begin{bmatrix} M_1 & M_2 & \dots & M_r \end{bmatrix},$$

for matrices of appropriate dimensions. Then:

$$\|M\|_2 \leq \sqrt{r} \max_{i=1,\dots,r} \|M_i\|_2$$

◇

Proof. Consider a vector x such that Mx is defined. Then, from triangle inequality, the definition of matrix norm and Cauchy-Schwartz:

$$\begin{aligned} \|Mx\|_2 = \|M_1 x_1 + \dots + M_r x_r\|_2 &\leq \|M_1\|_2 \|x_1\|_2 + \dots + \|M_r\|_2 \|x_r\|_2 \\ &\leq \sqrt{\|M_1\|_2^2 + \dots + \|M_r\|_2^2} \, \|x\|_2 \\ &\leq \sqrt{r} \max_{i=1,\dots,r} \|M_i\|_2 \|x\|_2 \,, \end{aligned}$$

where we used that $\|x\|_2^2 = \|x_1\|_2^2 + \dots + \|x_r\|^2$. □

Now, we can prove Theorem 3.3. Notice that:

$$S_t = \begin{bmatrix} \sum_{s=1}^{t} X_s \eta_s^* & \dots & \sum_{s=1}^{t} X_s \eta_{s+r-1}^* \end{bmatrix}.$$

We can analyze each component $\bar{V}_t^{-1/2} \sum_{s=1}^{t} X_s \eta_{k+i}^*$ separately and apply a union bound afterwards, since by Lemma 3.1:

$$\left\| \bar{V}_t^{-1/2} S_t \right\|_2 \leq \sqrt{r} \max_{i=0,\dots,r-1} \left\| \bar{V}_t^{-1/2} \sum_{s=1}^{t} X_s \eta_{s+i}^* \right\|_2 . \tag{3.36}$$

Now, fix a $0 \leq i < r$ and let

$$S_t^i \triangleq \sum_{s=1}^{t} X_s \eta_{s+i}^* .$$

Consider an arbitrary element of the unit sphere $\xi \in S^{m-1}$. The scalar $\eta^*_{s+i}\xi$ is conditionally 1-sub-Gaussian and satisfies the conditions of Theorem 3.5. Thus, with probability at least $1 - \frac{\delta}{r5^m}$:

$$\left\| \bar{V}_t^{-1/2} S_t^i \xi \right\|_2^2 \leq \mathcal{C}_{XH} \triangleq 2\left(\log \frac{r5^m}{\delta} + \frac{1}{2} \log \det \bar{V}_t V^{-1} \right)$$

Now, we apply Proposition 3.1 for $\epsilon = 1/2$:

$$\begin{aligned} &\mathbb{P}\left(\left\| \bar{V}_t^{-1/2} S_t^i \right\|_2 \geq 2\sqrt{\mathcal{C}_{XH}} \right) \\ &\leq 5^m \max_{\xi \in S^{m-1}} \mathbb{P}\left(\left\| \bar{V}_t^{-1/2} S_t^i \xi \right\|_2 \geq \sqrt{\mathcal{C}_{XH}} \right) \leq \frac{\delta}{r}. \end{aligned}$$

Finally, by (3.36) and a union bound over all components:

$$\begin{aligned} \mathbb{P}\left(\left\| \bar{V}_t^{-1/2} S_t \right\|_2 \geq 2\sqrt{r}\sqrt{\mathcal{C}_{XH}} \right) &\leq \mathbb{P}\left(\max_{i=0,\dots,r-1} \left\| \bar{V}_t^{-1/2} S_t^i \right\|_2 \geq 2\sqrt{\mathcal{C}_{XH}} \right) \\ &\leq \sum_{i=0}^{r-1} \mathbb{P}\left(\left\| \bar{V}_t^{-1/2} S_t^i \right\|_2 \geq 2\sqrt{\mathcal{C}_{XH}} \right) \leq \delta. \end{aligned}$$

□

3.7.2 Persistence of Excitation

The main focus of this section is the proof of Theorem 3.2, which provides finite sample guarantees for the PE of the past noises and the past outputs. We also include upper bounds for the sample correlations of the past outputs and the states $\hat{x}_k$. Finally, we provide the definition of N_0, N_1 that we hided in the main theorem statements.

The following result shows that with high probability, the past noises are persistently exciting. It follows from Lemma C.2 of Oymak & Ozay (2018), which in turn is based on results for random circulant matrices Krahmer et al. (2014).

Lemma 3.2 (Noise PE)**.** *Consider the conditions of Theorem 3.2 and the definition of* δ_N

in (3.19)*:*

$$\delta_N \triangleq (2(N+p-1)m)^{-\log^2(2pm)\log(2(N+p-1)m)}.$$

There exists a universal constant c (independent of system and algorithm parameters) such that if $N \geq 2cpm \log 1/\delta_N$, then with probability at least $1-\delta_N$ the event:

$$\mathcal{E}_E = \left\{ \frac{1}{2}\Sigma_E \preceq \frac{1}{N}\mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^* \right\},$$

occurs, where Σ_E is defined in (3.12)*.* ⋄

Proof. We can rewrite $E_k = \operatorname{diag}(\bar{R}^{1/2}, \ldots, \bar{R}^{1/2})U_k$, where U_k is defined similarly to E_k but has components with unit covariance. Now from Lemma C.2 of Oymak & Ozay (2018) applied on U_k we obtain that with probability at least $1-\delta_N$:

$$\frac{1}{N}\sum_{k=p}^{N+p-1} U_k U_k^* \succeq I_{mp}/2.$$

Multiplying by $\mathcal{T}_p \operatorname{diag}(\bar{R}^{1/2}, \ldots)$ from the left and $\operatorname{diag}(\bar{R}^{1/2}, \ldots)\mathcal{T}_p^*$ from the right gives the desired result. □

From the above lemma it follows that N should be large enough to guarantee PE for the noises. In particular, N should be larger than N_0, where

$$\begin{aligned} N_0 &= \min\left\{N :\ N \geq 2cpm \log 1/\delta_N\right\} \\ &= \min\left\{N :\ N \geq 2cpm \log^2(2pm)\log^2(2(N+p-1)m)\right\}. \end{aligned} \tag{3.37}$$

Such a N_0 exists since the term $2cpm\log^2(2pm)\log^2(2(N+p-1)m)$ depends logarithmically on N.

To guarantee PE for the noises, we also need to show that the smallest singular value $\sigma_E = \sigma_{\min}(\Sigma_E)$ is positive.

Lemma 3.3. *Let Σ_E be as in (3.12). Then:*

$$\sigma_E \geq \sigma_{\min}(R) > 0.$$

◇

Proof. The past errors $\mathcal{T}_p E_k$ can be rewritten as:

$$\mathcal{T}_p E_k = Z_k - \mathcal{O}_p \hat{x}_{k-p} = \mathcal{O}_p (x_{k-p} - \hat{x}_{k-p}) + V_k + \mathcal{T} W_k,$$

where V_k, W_k are defined similarly to Z_k and consist of the past measurement and process noises respectively. Matrix $\mathcal{T}$ is a block Toeplitz matrix; we omit its analytical expression. By independence of the measurement noise, the process noise and $x_{k-p} - \hat{x}_{k-p}$:

$$\Sigma_E \succeq \mathbb{E}[V_k V_k^*] = \mathrm{diag}(R, \ldots, R).$$

This in turn implies $\sigma_E \geq \sigma_{\min}(R)$. By Assumption 3.1, $R \succ 0$ and $\sigma_{\min}(R) > 0$. □

Lemma 3.4 (Markov upper bounds)**.** *Consider system (3.2) and recall the definition of Γ_k, Σ_E in (3.3), (3.12). We have the following upper bounds:*

$$\mathbb{P}\left(\left\|\bar{X}\bar{X}^*\right\|_2 \geq N \frac{\operatorname{tr}\Gamma_{N-1}}{\delta}\right) \leq \delta \tag{3.38}$$

$$\mathbb{P}\left(\left\|\bar{Z}\bar{Z}^*\right\|_2 \geq N \frac{\|\mathcal{O}_p\|_2^2 \operatorname{tr}\Gamma_{N-1} + \operatorname{tr}\Sigma_E}{\delta}\right) \leq \delta. \tag{3.39}$$

Proof. We only show (3.39). The proof of (3.38) is similar. From Markov's inequality we obtain:

$$\mathbb{P}\left(\|\bar{Z}\bar{Z}^*\|_2 \geq \epsilon\right) \leq \frac{\mathbb{E}\|\bar{Z}\bar{Z}^*\|_2}{\epsilon}.$$

What remains is to bound the expectation in the right-hand side. Notice that

$$\|Z_k (Z_k)^*\|_2 = \operatorname{tr} Z_k Z_k^*,$$

since Z_k has unit rank. Hence, by the triangle inequality:

$$\mathbb{E}\|\bar{Z}\bar{Z}^*\|_2 \leq \sum_{k=p}^{N+p-1} \operatorname{tr}\mathbb{E}\left[Z_k {Z_k}^*\right].$$

The right hand side of the above inequality is

$$\operatorname{tr}\mathbb{E}\left[Z_k {Z_k}^*\right] = \operatorname{tr}\left(\mathcal{O}_p\Gamma_{k-p}\mathcal{O}_p^* + \Sigma_E\right)$$

The result follows from $\operatorname{tr}\left(\mathcal{O}_p\Gamma_{k-p}\mathcal{O}_p^*\right) \leq \|\mathcal{O}_p\|_2^2 \operatorname{tr}\Gamma_{k-p}$ along with

$$\operatorname{tr}\sum_{k=p}^{N+p-1}\Gamma_{k-p} \leq N\operatorname{tr}\Gamma_{N-1}$$

since the sequence Γ_k is monotone (see the following Lemma). □

The following result is standard and we include it for completeness.

Lemma 3.5 (Monotonicity of Γ_k)**.** *Consider system* (3.2)*, under Assumption 3.1. The sequence* $\Gamma_k = \mathbb{E}\left[\hat{x}_k\hat{x}_k^*\right]$ *is monotone:* $\Gamma_k \succeq \Gamma_{k-1}$.

Proof. Notice that since $\hat{x}_0 = 0$, we have $\Gamma_0 = 0$. Define $\bar{Q} = K\bar{R}K^*$. By the orthogonality principle, $\hat{x}_k$ and e_k are uncorrelated. Hence

$$\Gamma_k = \mathcal{L}\left(\Gamma_{k-1}\right) \triangleq A\Gamma_{k-1}A^* + \bar{Q}.$$

For $k = 1$ we obtain:

$$\Gamma_1 = \bar{Q} \succeq 0 = \Gamma_0.$$

But the operator $\mathcal{L}$ is monotone, which implies:

$$\Gamma_2 = \mathcal{L}\left(\Gamma_1\right) \succeq \mathcal{L}\left(\Gamma_0\right) = \Gamma_1$$

The result $\Gamma_k \succeq \Gamma_{k-1}$ follows by induction. □

Proof of Theorem 3.2

Some arguments are similar to Section 9 of Sarkar & Rakhlin (2018).

Step 1: Noise PE. Under the condition $N \geq N_0$, from Lemma 3.2 the event $\mathcal{E}_E$ occurs with probability at least $1 - \delta_N$.

Step 2: Cross terms are small. Next, we show that the cross terms $\bar{X}\bar{E}^*$ are small. As in Lemma 3.2, express $\bar{E} = \mathrm{diag}(\bar{R}^{1/2}, \dots, \bar{R}^{1/2})\bar{H}$, where $\bar{H}$ is defined similarly to $\bar{E}$ but has unit variance components. Define:

$$\bar{V}_N = \bar{X}\bar{X}^* + \frac{N}{\|\mathcal{O}_p\|_2^2} I_n, \quad V = \frac{N}{\|\mathcal{O}_p\|_2^2} I_n, \quad V_N = \bar{X}\bar{X}^*, \quad S_N = \bar{X}\bar{E}^*.$$

Notice that

$$\left\|\bar{V}_N^{-1/2} S_N\right\|_2^2 \leq \left\|\bar{R}\right\|_2 \left\|\bar{V}_N^{-1/2} \bar{X}\bar{H}^*\right\|_2^2$$

Hence, by Theorem 3.3 applied to $\bar{V}_N$, $\bar{X}\bar{H}^*$ the event:

$$\mathcal{E}_1 = \left\{ \left\|\bar{V}_N^{-1/2} S_N\right\|_2^2 \leq 8p \left\|\bar{R}\right\|_2 \left(\log \frac{p5^m}{\delta} + \frac{1}{2} \log\det \bar{V}_N V^{-1} \right) \right\}$$

occurs with probability at least $1 - \delta$.

Next we upper bound term $\bar{V}_N$. From Lemma 3.4, the event:

$$\mathcal{E}_2 = \left\{ V_N \preceq N \frac{\mathrm{tr}\,\Gamma_{N-1}}{\delta} I_n \right\}$$

occurs with probability at least $1 - \delta$. This implies that:

$$\begin{aligned} \log\det \bar{V}_N V^{-1} &\leq \log\det \left[\left(\frac{N\,\mathrm{tr}\,\Gamma_{N-1}}{\delta} + \frac{N}{\|\mathcal{O}_p\|_2^2} \right) I_n \frac{\|\mathcal{O}_p\|_2^2}{N} \right] \\ &= \log \left(\frac{\|\mathcal{O}_p\|_2^2\,\mathrm{tr}\,\Gamma_{N-1}}{\delta} + 1 \right)^n = n \log \left(\frac{\|\mathcal{O}_p\|_2^2\,\mathrm{tr}\,\Gamma_{N-1}}{\delta} + 1 \right) \end{aligned}$$

Combining the two events $\mathcal{E}_1, \mathcal{E}_2$ and by a union bound, with probability at least $1 - 2\delta$

the event:

$$\mathcal{E}_{XE} = \left\{ \left\| \bar{V}_N^{-1/2} S_N \right\|_2^2 \leq \mathcal{C}_{XE} \left\| \bar{R} \right\|_2 \right\},$$

occurs, where

$$\mathcal{C}_{XE} \triangleq 8p \left(\frac{n}{2} \log \left(\frac{\|\mathcal{O}_p\|^2 \operatorname{tr} \Gamma_{N-1}}{\delta} + 1 \right) + \log \frac{p5^m}{\delta} \right).$$

As a consequence, if $u \in \mathbb{R}^{mp}$, $\|u\|_2 = 1$ is an arbitrary unit vector:

$$\begin{aligned}
&|u^* \mathcal{O}_p \bar{X} \bar{E}^* \mathcal{T}_p^* u| \leq \|u^* \mathcal{O}_p \bar{V}_N^{1/2} \bar{V}_N^{-1/2} S_N \mathcal{T}_p^*\|_2 \\
&\leq \sqrt{u^* \mathcal{O}_p \bar{X} \bar{X}^* \mathcal{O}_p^* u + N \frac{u^* \mathcal{O}_p \mathcal{O}_p^* u}{\|\mathcal{O}_p\|_2^2}} \sqrt{\mathcal{C}_{XE} \left\| \bar{R} \right\|_2} \|\mathcal{T}_p\|_2 \\
&\leq \sqrt{u^* \mathcal{O}_p \bar{X} \bar{X}^* \mathcal{O}_p^* u + N} \sqrt{\mathcal{C}_{XE} \left\| \bar{R} \right\|_2} \left\| \mathcal{T}_p \right\|_2, \text{ conditioned on } \mathcal{E}_{XE} \qquad (3.40)
\end{aligned}$$

Step 3: Output PE Consider an arbitrary unit vector $u \in \mathbb{R}^{mp}$, $\|u\|_2 = 1$. Consider the events $\mathcal{E}_E$ and $\mathcal{E}_{XE}$ from steps 1,2. With probability $1 - \delta_N - 2\delta$, since $N \geq N_0$ the event $\mathcal{E}_E \cap \mathcal{E}_{XE}$ occurs. It remains to show that on $\mathcal{E}_E \cap \mathcal{E}_{XE}$ the outputs satisfy PE for sufficiently large N. Define

$$\alpha \triangleq \frac{1}{N} u^* \mathcal{O}_p \bar{X} \bar{X}^* \mathcal{O}_p^* u, \ \beta \triangleq \frac{1}{N} u^* \mathcal{T}_p \bar{E} \bar{E}^* \mathcal{T}_p^* u$$

From (3.28), (3.40) for $N \geq N_0$ on $\mathcal{E}_E \cap \mathcal{E}_{XE}$:

$$\frac{1}{N} u^* \bar{Z} \bar{Z}^* u \geq \alpha + \beta - \underbrace{2 \left\| \mathcal{T}_p \right\|_2 \sqrt{\frac{\mathcal{C}_{XE} \left\| \bar{R} \right\|_2}{N}}}_{\gamma_N} \sqrt{\alpha + 1}$$

with $\beta \geq \sigma_E/2$. Now let N_1 be such that:

$$N_1 = \min \left\{ N : \ \gamma_N \leq \min \left\{ 1, \frac{\sigma_E}{4} \right\} \right\}. \qquad (3.41)$$

Since $\mathcal{C}_{XE}$ grows at most logarithmically with N, N_1 always exists. Now, since $N \geq N_1$

and $\beta \geq \sigma_E/2$:

$$\alpha + \beta - \gamma_N \sqrt{\alpha + 1} \geq \frac{\alpha + \beta}{2}.$$

The above inequality follows from the following lemma. □

Lemma 3.6 (Minimum of function)**.** *Let $\beta \geq b > 0$, for some $b > 0$ and consider the function:*

$$f(\alpha, \beta) = \frac{\alpha + \beta}{2} - \gamma\sqrt{\alpha + 1}, \text{ for } \alpha \geq 0,\ \beta \geq b > 0$$

If $\gamma \leq 1, \frac{b}{2}$, then

$$f(\alpha, \gamma) \geq 0, \text{ for all } \alpha \geq 0,\ \beta \geq b > 0.$$

◇

Proof. By elementary calculus:

$$\min_{\alpha \geq 0} f(\alpha, \beta) = \begin{matrix} \frac{\beta - 1 - \gamma^2}{2}, & \text{if } \gamma \geq 1 \\ \frac{\beta}{2} - \gamma, & \text{if } \gamma < 1 \end{matrix}$$

Thus, if $\gamma \leq 1, b/2$, we have $f(\alpha, \beta) \geq 0$. □

3.7.3 Proof of Theorem 3.1

Step 1: Since $N \geq N_0, N_1$, from Theorem 3.2, with probability at least $1 - \delta_N - 2\delta$, the event $\mathcal{E}_Y \cap \mathcal{E}_E$ occurs.

Step 2: Next, we analyze the cross term. Define:

$$\Psi_N = \bar{Z}\bar{Z}^*, \quad \bar{\Psi}_N = \Psi_N + N\frac{\sigma_E}{4} I_{mp}, \quad S_N = \mathcal{T}_f \bar{E}_+ \bar{Z}^*$$

Notice that on the event $\mathcal{E}_Y \cap \mathcal{E}_E$, we have that $\bar{\Psi}_N \preceq 2\Psi_N$ since $\Psi_N \succeq N\frac{\sigma_E}{4} I_{mp}$. Hence,

$$\left\| S_N \Psi_N^{-1/2} \right\|_2 \leq \sqrt{2} \left\| S_N \bar{\Psi}_N^{-1/2} \right\|_2.$$

Now the proof continues as in the case of cross-terms in the proof of Theorem 3.2. We apply Theorem 3.3 to $\bar{\Psi}_N, S_N$ and use Lemma 3.4 to upper bound $\bar{\Psi}_N$. Then, conditioned on $\mathcal{E}_Y \cap \mathcal{E}_E$, with probability $1-2\delta$:

$$\left\| S_N \Psi_N^{-1/2} \right\|_2^2 \leq 16 \left\| \Sigma_E^{+} \right\|_2 \left(\frac{fmp}{2} \log \frac{\kappa_N}{\delta} + f \log \frac{5^m f}{\delta} \right), \tag{3.42}$$

where

$$\kappa_N = \frac{4}{\sigma_E} \left(\|\mathcal{O}_p\|_2^2 \operatorname{tr} \Gamma_{N-1} + \operatorname{tr} \Sigma_E \right) + \delta.$$

Next, we bound $\|\Psi_N^{-1/2}\|$ separately on the event $\mathcal{E}_Y \cap \mathcal{E}_E$ by $\frac{2}{\sqrt{N\sigma_E}}$.

Finally, conditioned on $\mathcal{E}_E \cap \mathcal{E}_Y$, with probability at least $1-2\delta$:

$$\left\| \mathcal{T}_f \bar{E}_+ \bar{Z}^* \left(\bar{Z}\bar{Z}^* \right)^{-1} \right\| \leq \frac{\mathcal{C}_1}{\sqrt{N}} \sqrt{\frac{fmp}{2} \log \frac{\kappa_N}{\delta} + f \log \frac{5^m f}{\delta}}. \tag{3.43}$$

Step 3: We bound the Kalman truncation term. Recall that:

$$\bar{X}\bar{Z}^*(\bar{Z}\bar{Z}^*)^{-1} = \mathcal{O}_p^{\dagger} \left(I_{mp} - \mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^* (\bar{Z}\bar{Z}^*)^{-1} - \mathcal{T}_p \bar{E}\bar{X}^* \mathcal{O}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right)$$

On the event $\mathcal{E}_E \cap \mathcal{E}_Y$, we have:

$$\left\| \mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \leq 2.$$

since $\bar{Z}\bar{Z}^* \succeq \frac{1}{2} \mathcal{T}_p \bar{E}\bar{E}^* \mathcal{T}_p^*$. Hence we obtain:

$$\left\| \bar{X}\bar{Z}^*(\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \leq \left\| \mathcal{O}_p^{\dagger} \right\|_2 \left(3 + \|\mathcal{T}_p\|_2 \left\| \bar{E}\bar{X}^* \mathcal{O}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \right)$$

From the discussion in Section 3.4, we only need to find N_2 such that for $N \geq N_2$ with high probability:

$$\|\mathcal{T}_p\|_2 \left\| \bar{E}\bar{X}^* \mathcal{O}_p^* (\bar{Z}\bar{Z}^*)^{-1} \right\|_2 \leq 1.$$

Define

$$B_N = \mathcal{O}_p \bar{X}\bar{X}\mathcal{O}_p,\ B = \frac{\sigma_E}{2} N I_{mp},\ \bar{B}_N = B_N + B.$$

Notice that on the event $\mathcal{E}_E \cap \mathcal{E}_Y$, we have $\bar{Z}\bar{Z}^* \succeq \frac{1}{2}\bar{B}_N$, which implies:

$$\|\bar{E}\bar{X}^*\mathcal{O}_p^*(\bar{Z}\bar{Z}^*)^{-1/2}\| \leq \sqrt{2}\|\bar{E}\bar{X}^*\mathcal{O}_p^*\bar{B}_N^{-1/2}\|$$

Now we can treat the right-hand side in the same way as the cross-term above. By an application of Theorem 3.3 and Lemma 3.4, we obtain that conditioned on $\mathcal{E}_Y \cap \mathcal{E}_E$, with probability $1 - 2\delta$:

$$\begin{aligned}\left\|\bar{E}\bar{X}^*\mathcal{O}_p^*(\bar{Z}\bar{Z}^*)^{-1}\right\| &\leq \sqrt{2}\left\|\bar{E}\bar{X}^*\mathcal{O}_p^*\bar{B}_N^{-1/2}\right\|\left\|(\bar{Z}\bar{Z}^*)^{-1/2}\right\|_2 \\ &\leq 8\sqrt{\frac{\|\bar{R}\|_2}{\sigma_E}}\frac{\mathcal{C}_N}{\sqrt{N}},\end{aligned}$$

where

$$\mathcal{C}_N = \sqrt{\frac{mp^2}{2}\log\left(\frac{2\left\|\mathcal{O}_p\right\|_2^2 \operatorname{tr}\Gamma_{N-1}}{\delta\sigma_E} + 1\right) + p\log\frac{p5^m}{\delta}}$$

Thus, we define:

$$N_2 = \min\left\{N :\ 8\sqrt{\frac{\|\bar{R}\|_2}{\sigma_E}}\|\mathcal{T}_p\|_2\frac{\mathcal{C}_N}{\sqrt{N}} \leq 1\right\}. \tag{3.44}$$

Such an N_2 exists since $\mathcal{C}_N$ grows at most logarithmically with N.

Step 4: Final expression From the previous step and a union bound, for $N \geq N_0, N_1, N_2$ with probability at least $1 - \delta_N - 6\delta$:

$$\left\|G - \hat{G}\right\|_2 \leq \frac{\mathcal{C}_1}{\sqrt{N}}\sqrt{\frac{fmp}{2}\log\frac{\kappa_N}{\delta} + f\log\frac{5^m f}{\delta}} + \mathcal{C}_2\left\|A - KC\right\|_2^p, \tag{3.45}$$

Step 5: Simplification of final expression To simplify the final expression, we use

$$\frac{fmp}{2}\log\kappa_N + f\log\left(5^m f\right) \leq fmp\left(\log\kappa_N + \log\left(5f\right)\right) = fmp\log\left(5f\kappa_N\right)$$

and

$$\frac{fmp}{2}\log\frac{1}{\delta} + f\log\frac{1}{\delta} \leq fmp\log\frac{1}{\delta}$$

since $p \geq n+1 \geq 2$.

3.7.4 Proof of Theorem 3.4

The proof follows the steps of Oymak & Ozay (2018).

Step 1: Bounds for observability/controllability matrix

Denote the rank n approximation of $\hat{G}$ by

$$\hat{G}_n \triangleq \hat{\mathcal{O}}_f \hat{\mathcal{K}}_p$$

By definition, $\hat{G}_n$ is the matrix which minimizes $\left\|\hat{G}-M\right\|_2$, among all rank n matrices M. Thus, by optimality:

$$\left\|\hat{G}-\hat{G}_n\right\|_2 \leq \left\|\hat{G}-G\right\|_2,$$

since G has also rank n. As a result, we have:

$$\left\|G-\hat{G}_n\right\|_2 \leq \left\|G-\hat{G}\right\|_2 + \left\|\hat{G}-\hat{G}_n\right\|_2 \leq 2\left\|\hat{G}-G\right\|_2 \tag{3.46}$$

From (3.46) and the robustness condition (3.33) we have:

$$\left\|G-\hat{G}_n\right\|_2 \leq 2\left\|\hat{G}-G\right\|_2 \leq \frac{\sigma_n(G)}{2}. \tag{3.47}$$

Hence, we can now apply Theorem 5.14 of Tu et al. (2016), which states that there exists an orthonormal matrix T such that:

$$\sqrt{\left\|\hat{\mathcal{O}}_f-\bar{\mathcal{O}}_f T\right\|_F^2 + \left\|\hat{\mathcal{K}}_p - T^*\bar{\mathcal{K}}_p\right\|_F^2} \leq \sqrt{\frac{2}{\left(\sqrt{2}-1\right)\sigma_n(G)}}\left\|G-\hat{G}_n\right\|_F, \tag{3.48}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Since matrices $G, \hat{G}_n$ have rank-n, the sum $G-\hat{G}_n$ has rank at most $2n$. Thus, we can bound the Frobenius norm $\left\|G - \hat{G}_n\right\|_F$ in terms of spectral norm:

$$\left\|G - \hat{G}_n\right\|_F \leq \sqrt{2n}\left\|G - \hat{G}_n\right\|_2 \leq 2\sqrt{2n}\left\|\hat{G} - G\right\|_2 \tag{3.49}$$

where the second inequality follows from (3.46). For simplicity, we also use $\frac{2}{\sqrt{2}-1} \leq 5$. Thus, from (3.48) and the above inequalities:

$$\sqrt{\left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_F^2 + \left\|\hat{\mathcal{K}}_p - T^* \bar{\mathcal{K}}_p\right\|_F^2} \leq 2\sqrt{\frac{10n}{\sigma_n(G)}}\left\|G - \hat{G}\right\|_2.$$

Finally, since the spectral norm is always smaller than the Frobenius one:

$$\sqrt{\left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2^2 + \left\|\hat{\mathcal{K}}_p - T^* \bar{\mathcal{K}}_p\right\|_2^2} \leq 2\sqrt{\frac{10n}{\sigma_n(G)}}\left\|G - \hat{G}\right\|_2, \tag{3.50}$$

As a corollary,

$$\max\left\{\left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2, \left\|\hat{\mathcal{K}}_p - T^* \bar{\mathcal{K}}_p\right\|_2\right\} \leq 2\sqrt{\frac{10n}{\sigma_n(G)}}\left\|G - \hat{G}\right\|_2,$$

Step 2: Bounds for system parameters

Bounds for C, K

Since $\hat{C} - \bar{C}T$ is a sub-matrix of $\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T$, we immediately obtain:

$$\left\|\hat{C} - \bar{C}T\right\|_2 \leq \left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2 \leq 2\sqrt{\frac{10n}{\sigma_n(G)}}\left\|G - \hat{G}\right\|_2.$$

Similarly, for K we have:

$$\left\|\hat{K} - T^* \bar{K}\right\|_2 \leq \left\|\hat{\mathcal{K}}_p - T^* \bar{\mathcal{K}}_p\right\|_2 \leq 2\sqrt{\frac{10n}{\sigma_n(G)}}\left\|G - \hat{G}\right\|_2$$

Bounds for A

For simplicity, denote $\hat{M} \triangleq \hat{\mathcal{O}}_f^h$, $\bar{M} \triangleq \bar{\mathcal{O}}_f^h$ and $\hat{N} \triangleq \hat{\mathcal{O}}_f^l$, $\bar{N} \triangleq \bar{\mathcal{O}}_f^l$. Based on this notation:

$$\hat{A} = \hat{M}^\dagger \hat{N}, \quad \bar{A} = \bar{M}^\dagger \bar{N}.$$

After some algebraic manipulations:

$$\hat{A} - T^* \bar{A} T = \left(\hat{M}^\dagger - T^* \bar{M}^\dagger\right) \bar{N} T + \hat{M}^\dagger \left(\hat{N} - \bar{N} T\right).$$

First, notice that $\left\|\bar{N}\right\|_2 \leq \left\|\bar{\mathcal{O}}_f\right\|_2 = \sqrt{\|G\|}$, where the inequality follows from the fact that $\bar{N}$ is a submatrix of $\bar{\mathcal{O}}_f$; equality follows from the definition of $\bar{\mathcal{O}}_f = U_1 \Sigma_1^{1/2}$. Second, $\left\|\hat{M}^\dagger\right\|_2 = \frac{1}{\sigma_n(\hat{M})} \leq \frac{1}{\sigma_o}$ and third

$$\left\|\hat{N} - \bar{N} T\right\|_2 \leq \left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2,$$

since $\hat{N} - \bar{N} T$ is a submatrix of $\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T$. Finally, from Theorem 4.1 of Wedin (1972)

$$\left\|\hat{M}^\dagger - T^* \bar{M}^\dagger\right\|_2 \leq \left\|\hat{M} - T^* \bar{M}\right\|_2 \max\left\{\frac{1}{\sigma_n^2\left(\hat{M}\right)}, \frac{1}{\sigma_n^2\left(\bar{M}\right)}\right\} \leq \left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2 \frac{1}{\sigma_o^2}.$$

Combining all previous bounds, we obtain

$$\left\|\hat{A} - T^* \bar{A} T\right\|_2 \leq \left(\frac{\sqrt{G}}{\sigma_o^2} + \frac{1}{\sigma_o}\right) \left\|\hat{\mathcal{O}}_f - \bar{\mathcal{O}}_f T\right\|_2.$$

3.7.5 Bounds for system matrices

In this section, we formally prove bounds for $\mathcal{O}_k, \mathcal{T}_k, \Gamma_k$, which we implicitly used in Section 3.4.

First, we prove a standard result for the norm of (block) Toeplitz matrices.

Lemma 3.7 (Toeplitz norm)**.** *Let $M \in \mathbb{R}^{m_1 n \times m_2 n}$, for some integers n, m_1, m_2 be an* *(upper) block triangular Toeplitz matrix:*

$$M = \begin{bmatrix} M_1 & M_2 & M_3 & \cdots & \cdots & M_n \\ 0 & M_1 & M_2 & & & M_{n-1} \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & & & \\ & & & & M_1 & M_2 \\ 0 & 0 & \cdots & & 0 & M_1 \end{bmatrix},$$

where $M_i \in \mathbb{R}^{m_1 \times m_2}$, $i = 1, \ldots, n$. Then:

$$\|M\|_2 \leq \sum_{i=1}^{n} \|M_i\|_2$$

Proof. A standard technique in the analysis of Toeplitz matrices is to write them in terms of linear combinations of powers of companion matrices (see Horn & Johnson (2012), equation (0.9.7)). We can write M as:

$$M = I_n \otimes M_1 + \sum_{i=1}^{n-1} J_n^i \otimes M_i$$

where $\otimes$ is the Kronecker product, I_n is the identity matrix of dimension n, and J_n is the companion matrix:

$$J_n = J = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ & & \ddots & & \\ & & & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

But we have $\|D \otimes F\|_2 = \|D\|_2 \|F\|_2$ for any matrices D, F (see Theorem 4.2.15 in Horn & Johnson (1994)). Also the companion matrix has norm one $\|J_n\|_2 = 1$ since $J_n J_n^* = \text{diag}(1, \ldots, 1, 0)$. The result follows from the triangle inequality. □

Next, we provide a bound for the sum of powers of A.

Lemma 3.8. *Consider the series $S_t = \sum_{i=0}^{t} \left\| A^i \right\|_2$. We have the following two cases:*

- *If the system is asymptotically stable $\rho(A) < 1$, then $\|S_t\|_2 = \mathcal{O}(1)$*
- *If the system is marginally stable ($\rho(A) = 1$), then $\|S_t\|_2 = \mathcal{O}(t^{\kappa})$, where κ is the largest Jordan block of A corresponding to a unit circle eigenvalue $|\lambda| = 1$.*

Proof. **Proof of first part.** By Gelfand's formula Horn & Johnson (2012), for every $\epsilon > 0$, there exists a $i_0 = i_0(\epsilon)$ such that $\left\| A^i \right\| \leq (\rho(A) + \epsilon)^i$, for all $i \geq i_0$. Just pick ϵ such that $\rho(A) + \epsilon < 1$. Then,

$$S_t \leq \sum_{i=0}^{i_0} \left\| A^i \right\|_2 + \frac{1}{1 - \rho(A) - \epsilon} = \mathcal{O}(1).$$

Proof of second part. Assume that A is equal to a $n \times n$ Jordan block corresponding to $\lambda = 1$. The proof for the other cases is similar. Then we have that:

$$A^i = \begin{bmatrix} 1 & \binom{i}{1} & \dots & \binom{i}{n-1} \\ 0 & 1 & \dots & \binom{i}{n-2} \\ & & \ddots & \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

By Lemma 3.7, we obtain:

$$\left\| A^i \right\|_2 \leq \sum_{k=0}^{n-1} \binom{i}{k} \leq \left(\frac{ei}{n-1} \right)^{n-1}$$

where the second inequality is classical, see Exercise 0.0.5 in Vershynin (2018).

Finally, we have:

$$S_t \leq t \left(\frac{et}{n-1} \right)^{n-1} = \mathcal{O}(t^n)$$

□

Corollary 3.2. *The norms $\|\mathcal{O}_k\|_2$, $\|\mathcal{T}_k\|_2$, $\|\Gamma_k\|_2$ depend at most polynomially in k if $\rho(A) \leq 1$ (they are $\mathcal{O}(poly(k))$). If the system is asymptotically stable they are upper* *bounded for all k (they are $\mathcal{O}(1)$).*

Proof. **Observability matrix:** Consider the sum $S_t = \sum_{i=0}^{t} \left\|A^i\right\|_2$ of the previous lemma. We have:

$$\|\mathcal{O}_k\|_2 \leq \|C\|_2 S_{k-1} = \mathcal{O}(S_k).$$

The result follows by applying the previous lemma.

Block Toeplitz matrix Consider the sum $S_t = \sum_{i=0}^{t} \left\|A^i\right\|_2$ of the previous lemma. By Lemma 3.7 we have:

$$\|\mathcal{T}_k\|_2 \leq 1 + \|C\|_2 \|K\|_2 S_{k-2} = \mathcal{O}(S_k).$$

The result follows by applying the previous lemma.

Covariance matrix We have

$$\|\Gamma_k\|_2 \leq \left\|K\bar{R}K^*\right\|_2 \sum_{i=0}^{k-1} \left\|A^i\right\|^2.$$

The result follows by using a similar argument as in the previous lemma for $\sum_{i=0}^{k-1} \left\|A^i\right\|^2$. □

Lemma 3.9 (Lest non-zero singular values of G, $\mathcal{O}_p$ are increasing)**.** *The $n-th$ singular value of $\mathcal{O}_p$ is increasing with p:*

$$\sigma_n(\mathcal{O}_{p_1}) \geq \sigma_n(\mathcal{O}_{p_2}), \text{ for } p_1 \geq p_2.$$

The same is true for the $n-th$ singular value of G:

$$\sigma_n(\mathcal{O}_f \mathcal{K}_{p_1}) \geq \sigma_n(\mathcal{O}_f \mathcal{K}_{p_2}), \text{ for } p_1 \geq p_2.$$

Proof. We only prove the first part. The other proof is similar. Notice that $\mathcal{O}_{p_1}$ can be rewritten as:

$$\mathcal{O}_{p_1} = \begin{bmatrix} \mathcal{O}_{p_2} \\ M \end{bmatrix},$$

for some matrix M. Hence

$$\mathcal{O}_{p_1}^* \mathcal{O}_{p_1} \succeq \mathcal{O}_{p_2}^* \mathcal{O}_{p_2}.$$

Thus $\sigma_n^2\left(\mathcal{O}_{p_1}\right) \geq \sigma_n^2\left(\mathcal{O}_{p_2}\right)$. □

Chapter 4

Offline Learning of the Kalman Filter

4.1 Introduction

In this chapter, we study the problem of learning the Kalman Filter (KF) for unknown systems. Similar to the previous chapters, we focus on autonomous LTI systems

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k. \end{aligned} \tag{4.1}$$

We provide finite-sample guarantees for the performance of the offline learning architecture, where we learn the filter once based on batch data of finite size.

We consider a simple two step procedure. In the first step, using system identification tools rooted in subspace methods (see Chapter 3), we obtain finite-data estimates of the state-space parameters, and Kalman gain describing system (4.1). Then, in the second step, we use these approximate parameters to design a filter which predicts the system state. We provide an end-to-end analysis of this two-step procedure, and characterize the sub-optimality of the resulting filter in terms of the number of samples used during the system identification step. The sub-optimality is measured in terms of the mean square prediction error of the filter. A key insight that emerges from our analysis is that using a Certainty Equivalent (CE) Kalman Filter, i.e., using a Kalman Filter computed directly from estimated parameters, can yield poor estimation performance if the resulting Certainty Equivalent Kalman Filter has eigenvalues close to the unit circle. To address this issue, we propose a Robust Kalman Filter that mitigates these effects and that still enjoys provable sub-optimality guarantees.

Our main contributions are that: i) we show that if the system identification step produces sufficiently accurate estimates, or if the underlying true Kalman Filter is sufficiently robust, then the Certainty Equivalent Kalman Filter has near optimal mean square prediction error, ii) we show when the Certainty Equivalent Kalman Filter is marginally stable, i.e., when it has eigenvalues close to the unit circle, that a Robust Kalman Filter synthesized by explicitly imposing bounds on the magnitude of certain closed loop maps of the system enjoys similar mean square prediction error bounds as the Certainty Equivalent Kalman Filter, while demonstrating improved stability properties, and iii) we integrate the above results with the finite-data system identification guarantees of Chapter 3, to provide, to the best of our knowledge, the first end-to-end sample complexity bounds for the Kalman filtering of an unknown system. In particular, we show that the mean square estimation error of both the Certainty Equivalent and Robust Kalman Filter produced by the two step procedure described above is, with high probability, bounded by $\tilde{O}(1/\sqrt{N})$, where N is the number of samples collected in the system identification step.

Related work. A similar two step process was studied for the Linear Quadratic (LQ) control of an unknown system in Dean et al. (2017); Mania et al. (2019). While LQ optimal control and Kalman filtering are known to be dual problems, this duality breaks down when the state-space parameters describing the system dynamics are not known. In particular, the LQ optimal control problem assumes full state information, making the system identification step much simpler – in particular, it reduces to a simple least-squares problem. In contrast, in the Kalman Filter setting, as only partial observations are available, the additional challenge of finding an appropriate system order and state-space realization must be addressed. On the other hand, in the Kalman Filter problem one can directly estimate the filter gain from data, which makes analyzing performance of the Certainty Equivalent Kalman Filter simpler than the performance of the Certainty Equivalent LQ optimal controller (Mania et al., 2019).

System identification of autonomous LTI systems (4.1) is referred to as stochastic system identification (Van Overschee & De Moor, 2012). Classical results consider the asymptotic consistency of stochastic subspace system identification, as in Deistler et al. (1995); Bauer et al. (1999), whereas contemporary results seek to provide finite data guarantees (Tsiamis & Pappas, 2019; Lee & Lamperski, 2020). Finite data guarantees for system identification of partially observed systems can also be found in Oymak & Ozay (2018); Simchowitz et al. (2019); Sarkar et al. (2019), but these results focus on learning the non-stochastic part of the system, assuming that a user specified input is used to persistently excite the dynamics. More references can be found in Chapter 3.

Classical approaches to robust Kalman filtering can be found in El Ghaoui & Calafiore (2001); Sayed et al. (2001); Levy & Nikoukhah (2012), where parametric uncertainty is explicitly taken into account during the filter synthesis procedure. Although similar in spirit to our robust Kalman Filter procedure, these approaches assume fixed parametric uncertainty, and do not characterize the effects of parametric uncertainty on estimation performance, with this latter step being key in providing end-to-end sample complexity bounds. We also note that although not directly comparable to our work, the filtering problem for an unknown LTI system was also recently studied in the adversarial noise setting in Hazan et al. (2018), where a spectral filtering technique is used to directly predict the output bypassing the system identification step. In the stochastic noise case, online-learning of the Kalman Filter was studied in Kozdoba et al. (2019), where the goal is to predict a scalar output

Notation. We let bold symbols denote the frequency representation of signals. For example, $\mathbf{\Phi} = \sum_{t=0}^{\infty} \Phi_t z^{-t}$. If M is stable with spectral radius $\rho(M) < 1$, then we denote its resolvent by $\mathfrak{R}_M \triangleq (zI - M)^{-1}$. The $\mathcal{H}_2$ system norm is defined by $\|\mathbf{\Phi}\|_{\mathcal{H}_2}^2 \triangleq$

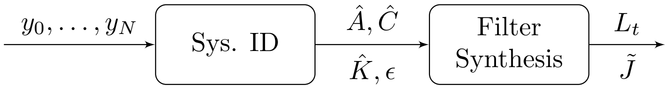


Figure 4.1: The proposed identification and filter synthesis pipeline. Using a single trajectory of N samples $\{y_t\}_{t=0}^N$ generated by system (4.2), a system identification algorithm computes estimates of $(\hat{A}, \hat{C}, \hat{K})$ with corresponding identification error bounds $\epsilon := \max(\epsilon_A, \epsilon_C, \epsilon_K)$. Then, using these estimates, we synthesize a filter defined by dynamic gains $\{L_t\}_{t=1}^\infty$, which has mean square prediction error $\tilde{J}$, defined in (4.3).

$\sum_{t=0}^\infty \|\Phi_t\|_F^2$, where $\|\cdot\|_F$ is the Frobenius norm. The $\mathcal{H}_\infty$ system norm is defined by $\|\mathbf{\Phi}\|_{\mathcal{H}_\infty} \triangleq \sup_{\|z\|=1} \|\mathbf{\Phi}(z)\|_2$, where $\|\cdot\|_2$ is the spectral norm. Let $\frac{1}{z}\mathcal{RH}_\infty$ be the set of real rational stable strictly proper transfer matrices.

4.2 Problem Formulation

Similar to Chapter 3, we consider the steady-state Kalman Filter form of system (4.1):

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Ke_k \\ y_k &= C\hat{x}_k + e_k, \end{aligned} \tag{4.2}$$

where $\hat{x}_k \in \mathbb{R}^n$ is the prediction (state), $y_k \in \mathbb{R}^m$ is the output, $e_k \in \mathbb{R}^n$ is the innovation process, and K is the Kalman filter gain defined in (2.2). The innovations e_k are i.i.d. zero mean Gaussians, with positive definite covariance matrix $\bar{R}$–see (3.4), and the initial predicted state is assumed to be $\hat{x}_0 = 0$. In general, the system (4.1) driven by i.i.d. zero mean Gaussian process and sensor noise is equivalent to system (4.2) for a suitable gain matrix K, as both noise models produce outputs with identical statistical properties–see Remark 1 or (Van Overschee & De Moor, 2012, Chapter 3). Similar to Chapter 3, we assume that the system is observable and minimal.

Assumption 4.1. *Matrices $A, C, K, \bar{R}$ are unknown. The order of the system n is known. Let Assumptions 3.1, 3.2, 3.3 hold. Assume that matrix A has spectral radius less than 1, i.e., $\rho(A) < 1$.*

Assumptions 3.1, 3.2, 3.3 guarantee that the Kalman filter is well-defined and in steady state (see Corollary 2.1). They also guarantee minimality of system (4.2). We note that the filter synthesis procedures we propose can be applied even if $\rho(A) \geq 1$ – however, in this case, we are unable to guarantee bounded estimation error for the resulting CE and robust KFs (see Theorem 4.1 and Lemma 4.2).

Our goal is to provide end-to-end sample complexity bounds for the two step pipeline illustrated in Fig. 4.1. First, we collect a trajectory $\{y_t\}_{t=0}^N$ of length N from system (4.2), and use system identification tools with finite data guarantees to learn the parameters $\hat{A}, \hat{C}, \hat{K}$ and bound the corresponding parameter uncertainties by $(\epsilon_A, \epsilon_C, \epsilon_K)$. Second, we use these approximate parameters to synthesize a filter from the following class:

$$\tilde{x}_k = \hat{A}\tilde{x}_{k-1} + \sum_{t=1}^{k} L_t(y_{k-t} - \hat{C}\tilde{x}_{k-t}), \quad \tilde{J} \triangleq \sqrt{\lim_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T} \|\tilde{x}_k - \hat{x}_k\|_2^2} \tag{4.3}$$

where $\{L_t\}_{t=1}^{\infty}$ are to be designed and $\tilde{J}$ is the filter's mean square prediction error as defined with respect to the optimal KF. Note that the predictor class above includes the CE KF – see Section 4.4 – and that if the the true system parameters are known, i.e., if $\hat{A} = A$, $\hat{C} = C$, $\hat{K} = K$, then the optimal mean squared prediction error $\tilde{J} = 0$ is achieved.

Problem 4.1 (End-to-end Sample Complexity). *Fix a failure probability $\delta > 0$. Given a single trajectory $y_0, \dots, y_N$ of system* (4.2)*, compute system parameter estimates $\hat{A}, \hat{C}, \hat{K}, \hat{R}$, and design a Kalman filter in class* (4.3)*, defined by gains $\{L_t\}_{t=1}^{\infty}$, such that with probability at least $1-\delta$, we have that $\tilde{J} \leq \epsilon_J$, so long as $N \geq \mathrm{poly}(1/\epsilon_J, \log(1/\delta))$.*

To address Problem 4.1, we will: i) leverage the results of the previous chapter regarding the sample complexity of stochastic system identification, ii) provide estimation guarantees for certainty equivalent as well as robust Kalman filter designed using the identified system parameters (see Problem 4.2 below), and (iii) provide end-to-end performance guarantees by integrating steps (i) and (ii) (see Problem 4.1 above).

In Chapter 3, we provided a finite sample analysis for stochastic system identification which provides bounds on the identification error $\epsilon := \max(\epsilon_A, \epsilon_C, \epsilon_K)$. In this chapter, we focus more on solving the Filter Synthesis task described below using both a certainty equivalent Kalman filter as well as a robust Kalman filter.

Problem 4.2 (Near Optimal Kalman Filtering of an Uncertain System)**.** *Consider system (4.2). Let $\hat{A}, \hat{C}, \hat{K}$ be estimates satisfying $\|A - \hat{A}\|_2 \le \epsilon_A$, $\|C - \hat{C}\|_2 \le \epsilon_C$, $\|K - \hat{K}\|_2 \le \epsilon_K$. Design a Kalman filter in class (4.3), defined by gains $\{L_t\}_{t=0}^{\infty}$, with mean square prediction error decaying with the size of the parameter uncertainty, i.e., such that $\tilde{J} \le O(\epsilon_A, \epsilon_C, \epsilon_K)$.*

Estimating the parameters of a partially observed system (4.2) is ill-posed, in that any similarity transformation S can be applied to generate parameters $(S^{-1}AS, CS, S^{-1}K, \bar{R})$ describing the same system. Formally, we should state Problems 4.1, 4.2 up to *some* similarity transformation S; to be more formal, we should define $\tilde{J}$ in terms of $S\tilde{x}_k - \hat{x}_k$ up to such a similarity transformation. For ease of exposition, we omit the transformation S in the problem formulation. However, we include it in the formal end-to-end sample complexity result.

4.3 System Identification Algorithm

In this section, we briefly present the stochastic identification algorithm. Here, we will consider a slightly more general version of the identification algorithm considered in Section 3.3. We borrow the notation of Section 3.3. The past and future horizons are denoted by p, f. The future outputs $Y_k \in \mathbb{R}^{mf}$ and past outputs $Z_k \in \mathbb{R}^{mp}$ at time $k \ge p$ are defined as

$$Y_k \triangleq \begin{bmatrix} y_k \\ \vdots \\ y_{k+f-1} \end{bmatrix}, \quad Z_k \triangleq \begin{bmatrix} y_{k-p} \\ \vdots \\ y_{k-1} \end{bmatrix}, \; k \ge p,$$

where we assume that we are given $N + p + f - 2$ output samples. The (extended) observability matrix $\mathcal{O}_k \in \mathbb{R}^{mk \times n}$ and the reversed (extended) controllability matrix $\mathcal{K}_k \in \mathbb{R}^{n \times mk}$ are defined in (3.8) and (3.9) respectively. The Hankel-like matrix is defined as $G = \mathcal{O}_f \mathcal{K}_p$. The covariance of the future noises Σ_E^+ is defined in (3.13).

Algorithm 1 Stochastic Identification Algorithm

```
Require: $p, f, y_0, \dots, y_{N+p+f-2}$, $W$.
Ensure: Estimates: $\hat{C}$, $\hat{A}$, $\hat{K}$, $\hat{R}$.
Compute $\hat{G} = \sum_{k=p}^{N+p-1} Y_k Z_k^* \left(\sum_{k=p}^{N+p-1} Z_k Z_k^*\right)^{-1}$.
Compute SVD: $\hat{G}W = \begin{bmatrix} \hat{U}_1 & \hat{U}_2 \end{bmatrix} \begin{bmatrix} \hat{\Sigma}_1 & 0 \\ 0 & \hat{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \hat{V}_1^* \\ \hat{V}_2^* \end{bmatrix}$, $\hat{\Sigma}_1 \in \mathbb{R}^{n \times n}$.
Set $\hat{\mathcal{O}}_f = \hat{U}_1 \hat{\Sigma}_1^{1/2}$, $\hat{\mathcal{K}}_f = \hat{\Sigma}_1^{1/2} \hat{V}_1^* W^{-1}$.
Set $\hat{C} = \hat{\mathcal{O}}_f(1:m,:)$, $\hat{K} = \hat{\mathcal{K}}_p(:, m(p-1)+1 : mp)$.
Set $\hat{A} = \hat{\mathcal{O}}_f(:, 1:m(f-1))^{\dagger} \hat{\mathcal{O}}_f(:, m+1 : mf)$.
Set $\hat{\Sigma}_E^+ = 1/N \sum_{k=p}^{N+p-1} (Y_k - \hat{G}Z_k)(Y_k - \hat{G}Z_k)^*$.
Set $\hat{R} = \hat{\Sigma}_E^+(1:m, 1:m)$.
```

Again, we first regress future outputs to past outputs to obtain a Hankel-like matrix. Then, we perform a realization step based on Singular Value Decomposition–see Section 3.3 for more details. The outline can be found in Algorithm 1. The algorithm is slightly more general since we allow the use of a weighting matrix W that post-multiplies $\hat{G}$ before the realization step; carefully choosing the weighting matrix might often improve the performance of subspace identification algorithms Van Overschee & De Moor (1995); Ljung (1999); Van Overschee & De Moor (2012). We also provide an estimate $\hat{R}$ of $\bar{R}$. Note that our finite-sample guarantees in Chapter 3 are for $W = I$ and do not cover the recovery of the innovation covariance $\bar{R}$. It is straightforward to extend the sample-complexity bounds to the case $W = \left(\sum_{k=p}^{N+p-1} Z_k Z_k^*\right)^{1/2}$[3], which is the value we selected in simulations. It is also straightforward to obtain upper bounds for the estimation error $\hat{R} - \bar{R}$ using the tools of Chapter 3. We omit the formal proof here and we leave the detailed analysis for future work.

[3] This is a variation of the MOESP algorithm Qin (2006).

4.4 Prediction Error Guarantees for Certainty Equivalent Kalman Filtering

For the Certainty Equivalent (CE) Kalman Filter, we directly use the estimated state-space parameters from the system identification step. Based on the estimated $\hat{K}, \hat{R}$ we compute the covariance:

$$\begin{bmatrix} \hat{Q} & \hat{S} \\ \hat{S}^* & \hat{R} \end{bmatrix} \triangleq \mathbb{E} \begin{bmatrix} \hat{K} e_k \\ e_k \end{bmatrix} \begin{bmatrix} e_k^* \hat{K}^* & e_k^* \end{bmatrix} = \begin{bmatrix} \hat{K} \hat{R}^{1/2} \\ \hat{R}^{1/2} \end{bmatrix} \begin{bmatrix} \hat{R}^{1/2} \hat{K}^* & \hat{R}^{1/2} \end{bmatrix}.$$

Then, based on standard Kalman Filter theory, we compute the stabilizing solution (see Chapter 2), of the following Riccati equation with correlation terms (Kailath et al., 2000):

$$P = \hat{A} P \hat{A}^* + \hat{Q} - (\hat{A} P \hat{C}^* + \hat{S})(\hat{C} P \hat{C}^* + \hat{R})^{-1}(\hat{C} P \hat{A}^* + \hat{S}^*). \tag{4.4}$$

Then, the CE Kalman Filter gain is static and takes the form

$$L_1 = L_{CE} \triangleq (\hat{A} P \hat{C}^* + \hat{S})(\hat{C} P \hat{C}^* + \hat{R})^{-1},\ L_t = 0,\ \text{for } t = 2, \dots. \tag{4.5}$$

Trivially, if $\rho(\hat{A} - \hat{K}\hat{C}) < 1$, then the stabilizing solution of the Riccati equation is $P = 0$ with $L_{CE} = \hat{K}$; the solution does not depend on $\hat{R}$. Formally this follows from the theory of non-stabilizable Riccati equations Chan et al. (1984). The following result explicitly describes the certainty equivalent gain.

Lemma 4.1. *Consider the assumptions of Problem 4.2. Assume that $(\hat{A}, \hat{C})$ is observable and $\hat{R}$ is positive definite. The CE Kalman Filter gain L_{CE} (4.5) has the following properties:*

- *If $\rho(\hat{A} - \hat{K}\hat{C}) < 1$, then $L_{CE} = \hat{K}$ and $\hat{A} - L_{CE}\hat{C}$ is asymptotically stable.*
- *If $\rho(\hat{A} - \hat{K}\hat{C}) > 1$, and $\hat{A} - \hat{K}\hat{C}$ has no eigenvalues on the unit circle, then $\hat{A} - L_{CE}\hat{C}$ is asymptotically stable.*

- *If $\hat{A}-\hat{K}\hat{C}$ has eigenvalues on the unit circle, then* (4.4) *does not admit a stabilizing solution.*

Proof. After some algebraic manipulations–see also Kailath et al. (2000), the Riccati equation (4.4) can be rewritten as:

$$P=(\hat{A}-\hat{K}\hat{C})P(\hat{A}-\hat{K}\hat{C})^*-(\hat{A}-\hat{K}\hat{C})P\hat{C}^*(\hat{C}P\hat{C}^*+\hat{R})^{-1}\hat{C}P(\hat{A}-\hat{K}\hat{C})^*$$

Notice that there is no Q term in the equivalent algebraic Riccati equation. If $\hat{A}-\hat{K}\hat{C}$ is already stable then the trivial solution $P=0$ is the stabilizing one. If $\hat{A}-\hat{K}\hat{C}$ is not asymptotically stable the results follow from Theorem 3.1 of Chan et al. (1984). □

The next result shows that if the underlying true Kalman Filter is sufficiently robust, as measured by a spectral decay rate, and that estimation parameter errors are sufficiently small, then the CE Kalman Filter achieves near optimal performance.

Theorem 4.1 (Near Optimal Certainty Equivalent Kalman Filtering)**.** *Consider Problem 4.2 and the CE Kalman Filter* (4.5)*. For any $\rho(A-KC)\le\rho<1$, define $\tau(A-KC,\rho)\triangleq\sup_{t\ge0}\left\|(A-KC)^t\right\|_2\rho^{-t}$. If the robustness condition*

$$2\tau(A-KC,\rho)\cdot(\epsilon_A+\epsilon_C(\|K\|_2+\epsilon_K)+\epsilon_K\|C\|_2)\le1-\rho$$

is satisfied, then $L_{CE}=\hat{K}$ and:

$$\tilde{J}\le\sqrt{3}\bar{\mathcal{C}}\epsilon\left\|\begin{bmatrix}\mathfrak{R}_AK\\ I\end{bmatrix}R^{1/2}\right\|_{\mathcal{H}_2}$$

where $\epsilon=\max\{\epsilon_A,\epsilon_C,\epsilon_K\}$, $\bar{\mathcal{C}}=2\frac{\tau(A-KC,\rho)}{1-\rho}(1+\|K\|_2+\epsilon_K)$ and $\mathfrak{R}_A=(zI-A)^{-1}$.

When CE Kalman filtering fails. The transient behavior of the CE Kalman Filter is governed by the closed loop eigenvalues of $\hat{A}-\hat{K}\hat{C}$, with performance degrading as eigenvalues approach the unit circle. This may occur if the estimation errors $(\epsilon_A,\epsilon_C,\epsilon_K)$

are large enough to cause $\rho(\hat{A} - \hat{K}\hat{C}) \approx 1$ even if the true system has spectral radius $\rho(A - KC) < 1$. We show in the next section that this undesirable scenario can be avoided by explicitly constraining the transient response of the resulting Kalman Filter to satisfy certain robustness constraints.

4.5 Prediction Error Guarantees for Robust Kalman Filtering

To address the possible poor performance of the CE Kalman Filter when model uncertainty is large, we propose to search over dynamic filters (4.3) subject to additional robustness constraints on their transient response. Using the System Level Synthesis (SLS) framework (Wang et al., 2019; Anderson et al., 2019) for Kalman filtering (Wang et al., 2015), we parameterize the class of dynamic filters (4.3) subject to additional robustness constraints in a way that leads to convex optimization problems.

4.5.1 SLS preliminaries

For this subsection, we assume that $\hat{A} = A, \hat{C} = C, \hat{K} = K$. The case of model error $A \neq \hat{A}$, $C \neq \hat{C}$, $K \neq \hat{K}$ is studied in the following subsection. Using bold symbols to denote the frequency representation of signals, we can rewrite the original system equation (4.2) and the predictor equation (4.3) as:

$$(zI - A + KC)\hat{\boldsymbol{x}} = K\boldsymbol{y}, \quad (zI - A + \boldsymbol{L}C)\tilde{\boldsymbol{x}} = \boldsymbol{L}\boldsymbol{y}.$$

Subtracting the two equations and using the fact that $\boldsymbol{y} = C\hat{\boldsymbol{x}} + \boldsymbol{e}$, we obtain:

$$\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}} = (zI - A + \boldsymbol{L}C)^{-1}K\boldsymbol{e} - (zI - A + \boldsymbol{L}C)^{-1}\boldsymbol{L}\boldsymbol{e}$$

Define the responses to $K\boldsymbol{e}$ and $\boldsymbol{e}$ by $\boldsymbol{\Phi}_w \triangleq (zI - A + \boldsymbol{L}C)^{-1}$ and $\boldsymbol{\Phi}_v \triangleq -(zI - A + \boldsymbol{L}C)^{-1}\boldsymbol{L}$ respectively. Then the error obtains the linear representation:

$$\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}} = (\boldsymbol{\Phi}_w K + \boldsymbol{\Phi}_v)\boldsymbol{e}$$

In (Wang et al., 2015), it is shown that these responses are in fact the closed loop maps from process and sensor noise $(\boldsymbol{w}, \boldsymbol{v})$ (here $K\boldsymbol{e}$ and $\boldsymbol{e}$) to state estimation error, and that the filter gain achieving the desired behavior can be recovered via $\boldsymbol{L} = -\boldsymbol{\Phi}_w^{-1}\boldsymbol{\Phi}_v$ so long as the responses $(\boldsymbol{\Phi}_w, \boldsymbol{\Phi}_v)$ are constrained to lie in an affine space defined by the system dynamics– see the following proposition. The following result from Wang et al. (2015) parameterizes the set of stable closed-loop transfer matrices $\boldsymbol{L}$.

Proposition 4.1 (Predictor parameterization)**.** *Consider system (4.2). Let $\frac{1}{z}\mathcal{RH}_\infty$ denote the set of real rational stable strictly proper transfer matrices. The closed-loop responses $\boldsymbol{\Phi}_w$, $\boldsymbol{\Phi}_v$ from $K\boldsymbol{e}$ and $\boldsymbol{e}$ to $\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}}$ can be induced by an internally stable predictor $\boldsymbol{L}$ if and only if they belong to the following affine subspace:*

$$\begin{bmatrix} \boldsymbol{\Phi}_w & \boldsymbol{\Phi}_v \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I,\ \boldsymbol{\Phi}_w,\ \boldsymbol{\Phi}_v \in \frac{1}{z}\mathcal{RH}_\infty. \tag{4.6}$$

Given the responses, we can parameterize the prediction gain as $\boldsymbol{L} = -\boldsymbol{\Phi}_w^{-1}\boldsymbol{\Phi}_v$.

Let $\boldsymbol{\Phi}_w = \sum_{t=0}^{\infty} \Phi_{w,t} z^{-t}$ and $\boldsymbol{\Phi}_v = \sum_{t=0}^{\infty} \Phi_{v,t} z^{-t}$. The strictly proper condition enforces the constraint $\Phi_{w,0} = 0, \Phi_{v,0} = 0$. The affine constraints simply imply that the system responses $\boldsymbol{\Phi}_w, \boldsymbol{\Phi}_v$ should satisfy the linear system recursions:

$$\Phi_{w,t+1} = \Phi_{w,t}A + \Phi_{v,t}C,\ t \geq 1, \quad \Phi_{w,1} = I$$

Assuming that the predictor is internally stable, then the mean square error is equal to

$$\tilde{J} = \|(\boldsymbol{\Phi}_w K + \boldsymbol{\Phi}_v)\bar{R}^{1/2}\|_{\mathcal{H}_2},$$

where $\|\cdot\|_{\mathcal{H}_2}$ is the $\mathcal{H}_2$ system norm. Hence, the error-free Kalman Filter synthesis problem could be re-written as:

$$\min_{\mathbf{\Phi}_w, \mathbf{\Phi}_v} \|(\mathbf{\Phi}_w K + \mathbf{\Phi}_v)\bar{R}^{1/2}\|_{\mathcal{H}_2}, \quad \text{s.t. (4.6)}$$

Of course, when the model knowledge is perfect, the solution to this problem is trivially $\boldsymbol{L} = K$, $\mathbf{\Phi}_w = (zI - A + KC)^{-1}$, $\mathbf{\Phi}_v = -(zI - A + KC)^{-1}K$, $\tilde{J} = 0$.

4.5.2 Filter Synthesis and Prediction Error Analysis

For a given dynamic predictor $\boldsymbol{L}(z) = \sum_{t=0}^{\infty} z^{-t} L_{t+1}$, we define the closed loop *system responses*:

$$\mathbf{\Phi}_w(z) \triangleq (zI - \hat{A} + \boldsymbol{L}\hat{C})^{-1}, \; \mathbf{\Phi}_v(z) \triangleq -(zI - \hat{A} + \boldsymbol{L}\hat{C})^{-1}\boldsymbol{L}. \tag{4.7}$$

By expressing the mean squared prediction error of the filters (4.3) in terms of their system responses, we are able to clearly delineate the effects of parametric uncertainty from the cost of deviating from the CE Kalman Filter.

Lemma 4.2 (Error analysis)**.** *Consider system (4.2). Let $\Delta_A \triangleq A - \hat{A}$, $\Delta_C \triangleq C - \hat{C}$, $\Delta_K \triangleq K - \hat{K}$. Any filter (4.3) with parameterization (4.7) has mean squared prediction error given by*

$$\tilde{J} = \left\| \begin{bmatrix} \mathbf{\Phi}_w & \mathbf{\Phi}_v \end{bmatrix} \left\{ \begin{bmatrix} \Delta_A & \Delta_K \\ \Delta_C & 0 \end{bmatrix} \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} + \begin{bmatrix} \hat{K} \\ I \end{bmatrix} \right\} \bar{R}^{1/2} \right\|_{\mathcal{H}_2}$$

Based on the previous lemma, we can upper bound the mean squared prediction error of filters (4.3) by

$$\tilde{J} \leq \underbrace{\sqrt{3}\epsilon \left\| \begin{bmatrix} \mathbf{\Phi}_w & \mathbf{\Phi}_v \end{bmatrix} \right\|_{\mathcal{H}_2} \left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} \bar{R}^{1/2} \right\|_{\mathcal{H}_\infty}}_{\text{parameter uncertainty term}} + \underbrace{\left\| \mathbf{\Phi}_w \hat{K} + \mathbf{\Phi}_v \right\|_{\mathcal{H}_2} \left\| \bar{R}^{1/2} \right\|_2}_{\text{suboptimality term}},$$

where $\epsilon = \max\{\epsilon_A, \epsilon_C, \epsilon_K\}$. This upper bound clearly separates the effects of parameter uncertainty, as captured by the first term, and the performance cost incurred by the filter $\boldsymbol{L}$ due to its deviation from the CE Kalman gain $\hat{K}$, as captured by the second. In order to optimally tradeoff between these two terms, we propose the following robust SLS optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{\Phi}_w, \boldsymbol{\Phi}_v} \left\| \boldsymbol{\Phi}_w \hat{K} + \boldsymbol{\Phi}_v \right\|_{\mathcal{H}_2} \\ &\text{s.t.} \left\| \begin{bmatrix} \boldsymbol{\Phi}_w & \boldsymbol{\Phi}_v \end{bmatrix} \right\|_{\mathcal{H}_2} \leq \mathcal{C}, \boldsymbol{\Phi}_w(zI - \hat{A}) - \boldsymbol{\Phi}_v \hat{C} = I, \, \boldsymbol{\Phi}_w, \, \boldsymbol{\Phi}_v \in \frac{1}{z}\mathcal{RH}_\infty \end{aligned} \tag{4.8}$$

where the constant $\mathcal{C}$ is a regularization parameter, and the affine constraint $\boldsymbol{\Phi}_w(zI - \hat{A}) - \boldsymbol{\Phi}_v \hat{C} = I$, $\boldsymbol{\Phi}_w$, $\boldsymbol{\Phi}_v \in \frac{1}{z}\mathcal{RH}_\infty$ parameterizes all filters of the form (4.3) that have bounded mean squared prediction error (see Wang et al. (2015) for more details). As we formalize in the following theorem, for appropriately selected regularization parameter $\mathcal{C}$ and sufficiently accurate estimation errors (ϵ_A, ϵ_C), the robust KF has near optimal mean square estimation error.

Theorem 4.2 (Robust Kalman Filter)**.** *Consider Problem 4.2 with Kalman Filters from class* (4.3) *synthesized using the robust SLS optimization problem* (4.8)*. If the regularization parameter is chosen such that* $\mathcal{C} \geq 2(1 + \|K\|_2) \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2}$*, and further, the estimation errors* (ϵ_A, ϵ_C) *are such that*

$$(\epsilon_A + \epsilon_C \|K\|_2) \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_\infty} \leq 1/2 \tag{4.9}$$

then the robust SLS optimization problem is feasible, and the synthesized robust Kalman Filter has mean squared prediction error upper-bounded by

$$\tilde{J} \leq \sqrt{3}\mathcal{C}\epsilon \left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} \right\|_{\mathcal{H}_\infty} \|\bar{R}^{1/2}\|_2 + 2\epsilon \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2} \|\bar{R}^{1/2}\|_2, \tag{4.10}$$

where $\epsilon = \max\{\epsilon_A, \epsilon_C, \epsilon_K\}$.

We further note that whenever the system responses induced by the CE Kalman Filter $\tilde{\boldsymbol{\Phi}}_w \triangleq (zI - \hat{A} + \hat{K}\hat{C})^{-1}$, $\tilde{\boldsymbol{\Phi}}_v \triangleq -(zI - \hat{A} + \hat{K}\hat{C})^{-1}\hat{K}$ are a feasible solution to optimization problem (4.8), they are also optimal, resulting in a filter $\boldsymbol{L} = \hat{K}$ with performance identical to the CE setting.

4.6 End-to-End Sample Complexity for the Kalman Filter

Theorems 4.1 and 4.2 provide two different solutions to Problem 4.2. Combining these theorems with the finite data system identification guarantees of Chapter 3, we now derive, to the best of our knowledge, the first end-to-end sample complexity bounds for the Kalman filtering of an unknown system. For both the CE and robust Kalman Filter, we show that the mean squared estimation error defined in (4.3) decreases with rate $O(1/\sqrt{N})$ up to logarithmic terms, where N is the number of samples collected during the system identification step.

Theorem 4.3 (End-to-end guarantees)**.** *Consider the conditions of Theorem 3.1 and suppose Assumption 4.1 holds. Let $p = \beta \log N$, $p \geq f > n$, with β as in (3.24). Consider the definition of S in (3.32) and δ_N in (3.19). Fix a failure probability $\delta \in (0, 1)$. Then, if*

$$N \geq \operatorname{poly}(\log(1/\delta), \beta, \sigma_n(G)),$$

with probability at least $1 - 6\delta - \delta_N$ the identification and filter synthesis pipeline of Fig. 4.1, with system identification performed as in Algorithm 1 with $W = I$ and filter synthesis performed as in Sections 4.4, 4.5, achieves mean squared prediction error satisfying

$$\sqrt{\lim_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t} \|\tilde{x}_k - T^* S^{-1} x_k\|_2} \leq \mathcal{C}_{ID}\mathcal{C}_{KF}\tilde{O}(\sqrt{\frac{\log 1/\delta}{N}}) \tag{4.11}$$

for some orthonormal matrix T. Constant $\mathcal{C}_{KF}$ is defined as:

$$\mathcal{C}_{KF} = \inf_{\rho > \rho(A-KC)} \frac{\tau(A-KC,\rho)}{1-\rho}(1+\|K\|_2)\left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} R^{1/2} \right\|_{\mathcal{H}_2}$$

in the case of CE Kalman filtering and

$$\mathcal{C}_{KF} = \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2}(1+\|K\|_2)\left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} R^{1/2} \right\|_{\mathcal{H}_\infty}$$

in the case of robust KF. Constant $\mathcal{C}_{ID}$ captures the difficulty of identifying system (4.2) and is defined as:

$$\mathcal{C}_{ID} = \sqrt{\frac{\|\Sigma_E^+\|_2}{\sigma_E}} \frac{1}{\sigma_n(\mathcal{O}_{f-1}S)\sqrt{\sigma_n(G)}} \sqrt{fmpn} \tag{4.12}$$

Here, $\tilde{O}$ hides constants, other system parameters, and logarithmic terms.

The proof follows from combining Corollary 3.1 and Theorem 3.4 with Theorems 4.1, 4.2. The notation poly$(\cdot)$ denotes a polynomial function of its arguments. Note that the condition on N follows from the condition $N \geq N_0$, N_1, N_2 in Corollary 3.1 along with (3.25) and the SVD robustness condition (3.33). Of course, the condition on N also depends on other system theoretic quantities as well, e.g. $\|\mathcal{T}_f\|_2$, which we omit for simplicity; they can be inferred from the proof.

The constant $\mathcal{C}_{KF}$ captures how robust the underlying open loop system A and closed loop Kalman Filter $A-KC$ are. We expect $\mathcal{C}_{KF}$ to be small for systems that admit optimal Kalman Filters with favorable robustness and transient performance. For example, if the spectral gaps $1-\rho(A)$ and $1-\rho(A-KC)$ are small (close to instability), then the constant $\mathcal{C}_{KF}$ becomes large. In contrast, the constant $\mathcal{C}_{ID}$ captures how easy it is to identify a system. The intuition is the following. The noise both excites the system and also introduces errors that obstruct identification; this is captured by the square root of the condition number of the covariances $\Sigma_{E,f}, \Sigma_{E,p}$. Moreover, $\sigma_n(\mathcal{O}_{f-1}S)$ quantifies how easy it is to observe system (4.2). A similar interpretation holds for $\sigma_n(G)$. Finally, larger dimensions f, p, m, n require more samples for identification since there are more unknowns in matrix G.

The bound derived in Theorem 4.3 highlights an interesting tension between how easy it is to identify the unknown system, and the robustness of the underlying optimal Kalman Filter. Recent results for the fully observed setting (Simchowitz et al., 2018; Sarkar & Rakhlin, 2018) suggest that systems with *larger* spectral radius are in fact easier to identify, as they provide more "signal" to the identification algorithm. In this way, our upper bound suggests that systems which properly balance between these two properties, robust transient performance and ease of identification, enjoy favorable sample complexity.

Note that the mean squared prediction error in (4.11) is computed with respect to the estimated state-space basis, i.e. up to the similarity transformation ST, where S is defined in (3.32) and T is some orthonormal matrix. In terms of the original state-space basis, the mean squared prediction error (4.11) would be:

$$\sqrt{\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|ST\tilde{x}_k - x_k\|_2} \leq \|S\|_2\,\mathcal{C}_{ID}\mathcal{C}_{KF}\tilde{O}(\sqrt{\frac{\log 1/\delta}{N}})$$

From (3.32), the norms of S, S^{-1} are bounded, so, the bound (4.11) is not vacuous.

We also note that the degradation of our bound with the inverse of the spectral gap $1-\rho(A)$ appears to be a limitation of the proposed offline two step architecture – indeed, Lemma 4.2 suggests that any estimation error in the state-space parameters (A, C) causes an increase in mean squared prediction error as $\|\mathfrak{R}_A\| \propto (1-\rho(A))^{-1}$ increases. As we will see in the next chapter, we can avoid this limitation in the online estimation setting.

4.7 Simulations

We perform Monte Carlo simulations of the proposed pipeline for the system

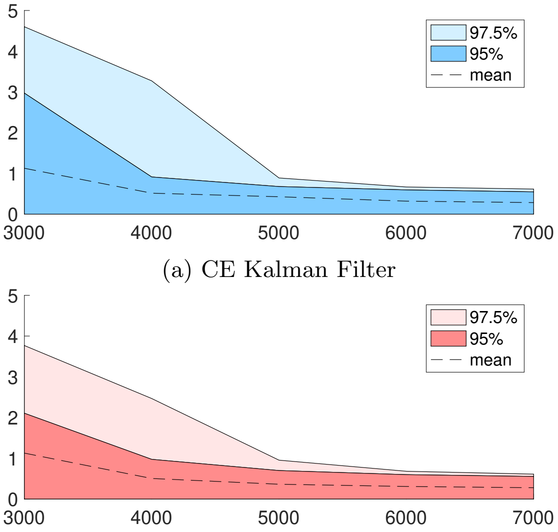
$$A = \begin{bmatrix} 0.8 & 1 & 0 \\ 0 & 0.9 & 1 \\ 0 & 0 & 0.9 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

$$K = \begin{bmatrix} 1.5320 & 0.9401 & 0.1923 \end{bmatrix}^*, \qquad \bar{R} = 10.6414.$$

for varying sample lengths N. We simulate both the CE and robust Kalman filters, and set the regularization parameter to $\mathcal{C} = 10$ in the robust SLS optimization problem (4.8). For each iteration, we first simulate system (4.2) to obtain N output samples. Then, we perform system identification to obtain the system parameters, after which we synthesize both CE and robust Kalman filters. Finally, we compute the mean prediction error of the designed filters.

For the identification scheme, we used $W = (\sum_{k=p}^{N+p-1} Z_k Z_k^*)^{1/2}$, i.e. the variation of the MOESP algorithm Qin (2006), which is more sample efficient in practice compared to the choice $W = I$–see Algorithm 1. The basis of the state-space representation returned by the subspace algorithm is data-dependent and varies with each simulation. For this reason, to compare the performance across different simulations, we compute the mean square error in terms of the original state space basis. Note that the SLS optimization problem (4.8) is semi-infinite since we optimize over the infinite variables $\{\Phi_{w,t}\}_{t=1}^{\infty}$ and $\{\Phi_{v,t}\}_{t=0}^{\infty}$. To deal with this issue, we optimize over a finite horizon T–see for example Dean et al. (2018), which makes the problem finite and tractable. Here, we selected $T = 30$.

Figure 4.2 (a) and (b) show the empirically computed mean squared prediction errors of the CE and Robust Kalman filters, with the mean, 95th, and 97.5th percentiles being shown. Notice that both errors decrease with a rate of $1/\sqrt{N}$, and that while the average behavior of both filters is quite similar, there is a noticeable gap in their tail behaviors. We observe that the most significant gap between the CE and Robust Kalman filters occurs



(a) CE Kalman Filter

(b) Robust Kalman Filter

Figure 4.2: The 95% and 97.5% empirical percentiles for the mean squared prediction error $\tilde{J}$ of the CE and Robust Kalman filters. We run 1000 Monte Carlo simulations for different sample lengths N (x-axis, number of samples).

when the eigenvalues of the CE matrix $\hat{A} - L_{CE}\hat{C}$ are close to the unit circle. Fig. 4.3 shows the empirical distribution of mean squared prediction errors conditioned on the event that $\rho(\hat{A} - L_{CE}\hat{C}) > 0.97$. In this case, the CE filter can exhibit *extremely* poor mean squared prediction error, with the worst observed error (not shown in Fig. 4.3 in the interst of space) approximately equal to 70 – in contrast, the worst error exhibited by the robust Kalman filter was approximately equal to 5. Thus, we were able to achieve a 14x reduction in worst-case mean squared error. For some simulations the robust KF can exhibit worse performance compared to the CE Kalman filter. However, over all simulations, the mean squared error achieved by the robust Kalman filter was at most 1.64x greater than that achieved by CE Kalman filter.

4.8 Conclusions & Future work

In this chapter, we proposed and analyzed a system identification and filter synthesis pipeline. Leveraging contemporary finite data guarantees from system identification (Tsiamis & Pappas, 2019), as well as novel parameterizations of robust Kalman filters (Wang

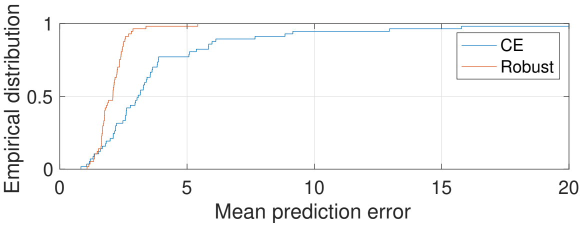


Figure 4.3: Performance improvement for the robust KF conditioned on the event that $\rho(\hat{A} - L_{CE}\hat{C}) > 0.97$.

et al., 2015), we provided, to the best of our knowledge, the first end-to-end sample complexity bounds for the Kalman filtering of an unknown LTI autonomous system, in an offline learning context. Our analysis revealed that, depending on the spectral properties of the CE Kalman filter, that a robust Kalman filter approach may lead to improved performance. In future work, we would like to explore how to improve robustness and performance by further exploiting information about system uncertainty, as well as how to integrate our results into an optimal control framework, such as Linear Quadratic Gaussian control. We would also like to explore more options for the regularization parameter $\mathcal{C}$, for example, let it depend on the data or the number of samples.

4.9 Proofs

Proof of Theorem 4.1

Let $\Delta_{A_{cl}} = (A - KC) - (\hat{A} - \hat{K}\hat{C})$. By adding and subtracting $\hat{K}C$, we obtain the bound:

$$\|\Delta_{A_{cl}}\| \leq \epsilon_A + \|\hat{K}\|_2\epsilon_C + \epsilon_K\|C\|_2 \leq \epsilon_A + (\|K\|_2 + \epsilon_K)\epsilon_C + \epsilon_K \|C\|_2$$

Hence, from the robustness condition of the theorem it follows that

$$2\tau(A - KC, \rho)\|\Delta_{A_{cl}}\|_2 \leq 1 - \rho \tag{4.13}$$

Now, from Lemma 5 in Mania et al. (2019) it follows that:

$$\|(\hat{A}-\hat{K}\hat{C})^k\|_2 = \|(A-KC-\Delta_{A_{cl}})^k\|_2 \le \tau(A-KC,\rho)\left(\tau(A-KC,\rho)\left\|\Delta_{A_{cl}}\right\|_2+\rho\right)^k \tag{4.14}$$

Combining (4.13), (4.14), we finally obtain:

$$\|(\hat{A}-\hat{K}\hat{C})^k\|_2 \le \tau(A-KC,\rho)\left(\frac{1+\rho}{2}\right)^k.$$

Thus, the $\mathcal{H}_\infty$ norm of $\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}$ is upper bounded by

$$\begin{aligned}
\left\|\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}\right\|_{\mathcal{H}_\infty} &\le \sum_{t=0}^{\infty}\|(\hat{A}-\hat{K}\hat{C})^t\|_2 \\
&\le \tau(A-KC,\rho)\sum_{k=0}^{\infty}\left(\frac{1+\rho}{2}\right)^k = \frac{2\tau(A-KC,\rho)}{1-\rho}
\end{aligned}$$

This further implies

$$\begin{aligned}
\left\|\begin{bmatrix}\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}} & -\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}\hat{K}\end{bmatrix}\right\|_{\mathcal{H}_\infty} &\le (1+\|K\|_2+\epsilon_K)\left\|\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}\right\|_{\mathcal{H}_\infty} \\
&\le (1+\|K\|_2+\epsilon_K)\frac{2\tau(A-KC,\rho)}{1-\rho}.
\end{aligned}$$

Now let $\Phi_w = \boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}$ and $\Phi_v = -\boldsymbol{\mathfrak{R}}_{\hat{A}-\hat{K}\hat{C}}\hat{K}$. The proof follows from Lemma 4.2 and the inequality

$$\begin{aligned}
&\left\|\begin{bmatrix}\boldsymbol{\Phi}_w & \boldsymbol{\Phi}_v\end{bmatrix}\begin{bmatrix}\Delta_A & \Delta_K \\ \Delta_C & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathfrak{R}}_A K \\ I\end{bmatrix}\right\|_{\mathcal{H}_2} \\
&\le \left\|\begin{bmatrix}\boldsymbol{\Phi}_w\boldsymbol{\Phi}_v\end{bmatrix}\right\|_{\mathcal{H}_\infty}\left\|\begin{bmatrix}\Delta_A & \Delta_K \\ \Delta_C & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathfrak{R}}_A K \\ I\end{bmatrix}\bar{R}^{1/2}\right\|_{\mathcal{H}_2} \\
&\le \sqrt{3}\epsilon(1+\|K\|_2+\epsilon_K)\frac{2\tau(A-KC,\rho)}{1-\rho}\left\|\begin{bmatrix}\boldsymbol{\mathfrak{R}}_A K \\ I\end{bmatrix}\bar{R}^{1/2}\right\|_{\mathcal{H}_2}
\end{aligned}$$

■

Proof of Lemma 4.2

It is sufficient to show that

$$\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}} = \Big\{ (\boldsymbol{\Phi}_w \Delta_A + \boldsymbol{\Phi}_v \Delta_C)\, \mathfrak{R}_A K + \boldsymbol{\Phi}_w \Delta_K + \boldsymbol{\Phi}_w \hat{K} + \boldsymbol{\Phi}_v \Big\} \boldsymbol{e},$$

then the result follows from the definition of $\mathcal{H}_2$ norm and the fact that $\bar{R}^{-1/2}e$ is white noise with unit variance.

In frequency domain, equations (4.2), (4.3) can be rewritten as

$$(zI - \hat{A})\tilde{\boldsymbol{x}} = \boldsymbol{L}(\boldsymbol{y} - \hat{C}\tilde{\boldsymbol{x}}),\ (zI - A)\hat{\boldsymbol{x}} = K(\boldsymbol{y} - C\hat{\boldsymbol{x}})$$

Subtracting the two equations yields:

$$(zI - \hat{A} + \boldsymbol{L}\hat{C})(\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}}) + (-\Delta_A - \boldsymbol{L}\hat{C} + KC)\hat{\boldsymbol{x}} = (K - \boldsymbol{L})\boldsymbol{y}$$

Using the fact that $\boldsymbol{y} = C\boldsymbol{x} + \boldsymbol{e}$, we obtain:

$$(zI - \hat{A} + \boldsymbol{L}\hat{C})(\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}}) = (\Delta_A - \boldsymbol{L}[C - \hat{C}])\hat{\boldsymbol{x}} + (K - \boldsymbol{L})\boldsymbol{e}.$$

Multiplying from the left by $\boldsymbol{\Phi}_w$ and using the fact that $\boldsymbol{\Phi}_v = -\boldsymbol{\Phi}_w \boldsymbol{L}$

$$\hat{\boldsymbol{x}} - \tilde{\boldsymbol{x}} = (\boldsymbol{\Phi}_w \Delta_A + \boldsymbol{\Phi}_v \Delta_C)\hat{\boldsymbol{x}} + (\boldsymbol{\Phi}_w K + \boldsymbol{\Phi}_v)\boldsymbol{e}$$

The result follows from adding and subtracting $\boldsymbol{\Phi}_w \hat{K}\boldsymbol{e}$ and the fact that $\hat{\boldsymbol{x}} = \mathfrak{R}_A K \boldsymbol{e}$. ■

Proof of Theorem 4.2

Step a: First we prove that when optimization problem (4.8) is feasible, the the mean square error is bounded by:

$$\tilde{J} \leq \sqrt{3}\mathcal{C}\epsilon \left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} \right\|_{\mathcal{H}_\infty} \|\bar{R}^{1/2}\|_2 + \text{opt}(\mathcal{C})\|\bar{R}^{1/2}\|_2. \tag{4.15}$$

Assume that $(\mathbf{\Phi}_w, \mathbf{\Phi}_v)$ is an optimal solution to (4.8). From Lemma 4.2:

$$\begin{aligned}\tilde{J} \leq & \sqrt{3}\epsilon \left\| \begin{bmatrix} \mathbf{\Phi}_w & \mathbf{\Phi}_v \end{bmatrix} \right\|_{\mathcal{H}_2} \left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} \right\|_{\mathcal{H}_\infty} \|\bar{R}^{1/2}\|_2 + \left\| (\mathbf{\Phi}_w \hat{K} + \mathbf{\Phi}_v) \right\|_{\mathcal{H}_2} \|\bar{R}^{1/2}\|_2, \\ \leq & \sqrt{3}\mathcal{C}\epsilon \left\| \begin{bmatrix} \mathfrak{R}_A K \\ I \end{bmatrix} \right\|_{\mathcal{H}_\infty} \|\bar{R}^{1/2}\|_2 + \text{opt}(\mathcal{C})\|\bar{R}^{1/2}\|_2,\end{aligned}$$

where we used $\left\| \begin{bmatrix} \mathbf{\Phi}_w & \mathbf{\Phi}_v \end{bmatrix} \right\|_{\mathcal{H}_2} \leq \mathcal{C}$ and optimality of $(\mathbf{\Phi}_w, \mathbf{\Phi}_v)$.

Step b: We prove that under condition (4.9), the static Kalman gain K is a feasible gain for (4.8); equivalently, the responses $\tilde{\mathbf{\Phi}}_w = \mathfrak{R}_{\hat{A}-K\hat{C}}$, and $\tilde{\mathbf{\Phi}}_v = -\mathfrak{R}_{\hat{A}-K\hat{C}}K$ satisfy the constraints of (4.8). Consider the responses $\mathbf{\Phi}_{w,opt} \triangleq \mathfrak{R}_{A-KC}$ and $\mathbf{\Phi}_{v,opt} \triangleq -\mathfrak{R}_{A-KC}K$, which are optimal for the original unknown system. They satisfy the affine relation for the original system:

$$\begin{bmatrix} \mathbf{\Phi}_{w,opt} & \mathbf{\Phi}_{v,opt} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I$$

Adding and subtracting the estimated matrices, we can show that they also satisfy a perturbed affine relation for the estimated system:

$$\begin{bmatrix} \mathbf{\Phi}_{w,opt} & \mathbf{\Phi}_{v,opt} \end{bmatrix} \begin{bmatrix} zI - \hat{A} \\ -\hat{C} \end{bmatrix} = I + \underbrace{(\mathbf{\Phi}_{w,opt}\delta_A + \mathbf{\Phi}_{v,opt}\delta_C)}_{\mathbf{\Delta}}$$

If the perturbation $(I + \mathbf{\Delta})^{-1}$ is stable, we can multiply both sides from the left, which yields:

$$\begin{bmatrix} \tilde{\mathbf{\Phi}}_w & \tilde{\mathbf{\Phi}}_v \end{bmatrix} \begin{bmatrix} zI - \hat{A} \\ -\hat{C} \end{bmatrix} = I,$$

where we used the fact that:

$$(I + \mathbf{\Delta})^{-1} \mathbf{\Phi}_{w,opt} = \tilde{\mathbf{\Phi}}_w, \quad (I + \mathbf{\Delta})^{-1} \mathbf{\Phi}_{v,opt} = \tilde{\mathbf{\Phi}}_v$$

Under condition (4.9), the perturbation $\mathbf{\Delta}$ has norm bounded by:

$$\|\mathbf{\Delta}\|_{\mathcal{H}_\infty} \leq (\epsilon_A + \epsilon_C \|K\|_2) \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_\infty} \leq 1/2$$

Hence:

$$\left\|(I + \mathbf{\Delta})^{-1}\right\|_{\mathcal{H}_\infty} \leq \sum_{t=0}^{\infty} \|\mathbf{\Delta}\|_{\mathcal{H}_\infty}^t \leq \frac{1}{1 - \|\mathbf{\Delta}\|_{\mathcal{H}_\infty}} = 2$$

which shows that the responses $\tilde{\mathbf{\Phi}}_w, \tilde{\mathbf{\Phi}}_v$ are stable. By construction, they are also strictly proper. What remains to show is that the robustness constraint holds. We have:

$$\begin{aligned} \left\|\begin{bmatrix} \tilde{\mathbf{\Phi}}_w & \tilde{\mathbf{\Phi}}_v \end{bmatrix}\right\|_{\mathcal{H}_2} &\leq \left\|(I + \mathbf{\Delta})^{-1} \begin{bmatrix} \mathbf{\Phi}_w & \mathbf{\Phi}_v \end{bmatrix}\right\|_{\mathcal{H}_2} \\ &\leq \left\|(I + \mathbf{\Delta})^{-1}\right\|_{\mathcal{H}_\infty} (1 + \|K\|_2) \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2} \\ &\leq 2(1 + \|K\|_2) \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2} \leq \mathcal{C} \end{aligned}$$

Step c: Since K is a feasible gain, by suboptimality

$$\begin{aligned} \operatorname{opt}(\mathcal{C}) &\leq \left\|\tilde{\mathbf{\Phi}}_w \hat{K} + \tilde{\mathbf{\Phi}}_v\right\|_{\mathcal{H}_2} \leq \left\|(I + \mathbf{\Delta})^{-1}\right\|_{\mathcal{H}_\infty} \left\|\mathbf{\Phi}_w \hat{K} + \mathbf{\Phi}_v\right\|_{\mathcal{H}_2} \\ &\leq 2 \left\|\mathbf{\Phi}_w \hat{K} + \mathbf{\Phi}_v\right\|_{\mathcal{H}_2} = 2 \left\|\mathbf{\Phi}_w (\hat{K} - K)\right\|_{\mathcal{H}_2} \\ &\leq 2\epsilon \|\mathfrak{R}_{A-KC}\|_{\mathcal{H}_2} \end{aligned}$$

where we used $\mathbf{\Phi}_v = -\mathbf{\Phi}_w K$. ■

Chapter 5

Online Learning of the Kalman Filter

5.1 Introduction

In this chapter, we study the sample complexity of the online architecture for learning the Kalman filter of an unknown LTI system. From a control theoretic perspective, this problem has also been known as the adaptive filtering problem (Ljung, 1978; Moore & Ledwich, 1979; Lai & Ying, 1991; Ding et al., 2006). Adaptive filtering algorithms address the problem of making observation predictions when the system model or the noise statistics are unknown or changing. These adaptive filtering approaches are usually based on variations of extended least squares. The statistical analysis of their behavior has relied on asymptotic tools, which assume that the number of collected data N is infinite. However, our asymptotic tools, e.g. the Central Limit Theorem or Law of Large Numbers, do not always capture all aspects of finite sample performance (Vershynin, 2018, Chapter 2). Moreover, the dependence of prediction performance on various system theoretic parameters has been hidden under the big-O notation.

Here, we consider the problem of predicting observations generated by an unknown, partially observable LTI dynamical system under *finite samples*. The system dynamics and observation map are corrupted by Gaussian noise. For the theoretical analysis, we adopt the notion of regret Cesa-Bianchi & Lugosi (2006), which captures the finite sample performance of online prediction. It measures how far our online predictions are from the optimal Kalman Filter predictions that has access to the full system model. Our goal is to find an online prediction algorithm that has provably small regret. Our technical contributions are:

System theoretic regret: We define a notion of regret that has a natural, system theoretic interpretation. The prediction error of an online prediction algorithm is compared against the prediction error of the Kalman filter that has access to the exact model, which is allowed to be arbitrary. Previous regret definitions Kozdoba et al. (2019) required the model to lie in a finite set.

Logarithmic regret for the Kalman filter: We present the first online prediction algorithm with provable logarithmic regret upper bounds for the classical Kalman filter. In fact, we prove that with high probability the regret of our algorithm is of the order of $\tilde{O}(1)$, where $\tilde{O}$ hides $\operatorname{poly}\log N$ terms, where N is the number of observations collected. Our algorithm is based on subspace system identification techniques Qin (2006). Instead of optimizing over the state-space parameters, which is a non-convex problem, we convexify the problem by establishing an approximate regression between the next observation and past observations. Our analysis is based on the stability properties of the Kalman filter, tools for self-normalized martingales and matrices, high-dimensional statistics Vershynin (2018), and additional results for persistency of excitation developed here. We note that our bounds are qualitative, showing how various system theoretic parameters affect learning performance.

Logarithmic regret for non-explosive systems: Our regret guarantees hold for the class of non-explosive systems, which includes marginally stable linear systems. This settles an open question and concludes that online prediction performance does not depend on the system stability gap ($1/(1-\rho)$, where ρ is the spectral radius of the system). Although it was recently shown that the stability gap does not affect system identification Simchowitz et al. (2019), whether the stability gap affects online prediction under stochastic noise was an open problem.

Regret analysis for other predictors: Our approach directly carries over to various interesting online predictors. For example, our analysis can be directly extended to the case of $f-$step ahead prediction of observations. Another extension focuses on the regret of hidden state predictors when the state-space basis representation is known a priori. The latter situation arises, for example, when the state-space model is known but the noise statistics are unknown. Another case is prediction of stable closed-loop systems. All these predictors enjoy similar logarithmic regret bounds.

Learning gap between LQR and Kalman filter: One of the implications of our bounds is that learning to predict observations like the Kalman filter is provably easier than solving the online Linear Quadratic Regulator (LQR) problem, which in general requires $\Omega(\sqrt{N})$ regret Simchowitz & Foster (2020); Ziemann & Sandberg (2020). This might not be surprising due to the fact that, in the case of exogenous inputs, we need to inject exploratory signals into the system.

5.1.1 Related work

Recently, there have been important results addressing the regret of the adaptive Linear Quadratic Regulator (LQR) problem Abbasi-Yadkori & Szepesvári (2011); Faradonbeh et al. (2020b); Ouyang et al. (2017b); Abeille & Lazaric (2018); Dean et al. (2018); Mania et al. (2019); Cohen et al. (2019). The best regret for LQR is sublinear and of the order of $\tilde{O}(\sqrt{N})$, where N is the numbers of state samples collected; an in-depth survey can be found in Matni et al. (2019). When the system model is *known*, then the Kalman filter is the dual of the Linear Quadratic Regulator, suggesting that this duality can be exploited in deriving the regret of the Kalman filter. However, when the system model is *unknown*, the Linear Quadratic Regular and the Kalman filter are not dual problems Tsiamis et al. (2020). As the state is fully observed in LQR, the system identification in adaptive LQR reduces to a simple least squares problem. In the adaptive Kalman filter, the state is *partially* observed resulting in non-convex system identification problems requiring us to consider a different approach.

A related but different problem focuses on online prediction algorithms for systems without internal states (such as ARMA - autoregressive moving average) Anava et al. (2013). Prediction of observations generated by state space models in the case of exogenous inputs and adversarial noise but with a bounded budget was studied in Hazan et al. (2018). Recently, Kozdoba et al. (2019) introduced regret bounds for the Kalman Filter in the restricted context of scalar and bounded observations. The regret is shown to be of the order of $\sqrt{N}$ along with a small linear term. Here, we improve the state of the art to logarithmic bounds for general observations. Concurrently and independently Ghai et al. (2020) also proved logarithmic regret bounds for the Kalman Filter. Our analysis here is different focusing on persistency of excitation, which can also provide simultaneous parameter estimation guarantees. After our work, regret bounds were extended to the case where the Kalman Filter closed-loop matrix is close to instability Rashidinejad et al. (2020). Finally, Lale et al. (2020b) proved logarithmic regret bounds for the Linear Quadratic Gaussian (LQG) control problem, which is different from the prediction problem studied here.

Our online algorithm is inspired by subspace identification techniques Bauer et al. (1999). The technical approach is based on classical results for the analysis of the least-squares estimator for time series Lai & Wei (1982), as well as modern results for finite sample analysis of system identification in both the fully observed Faradonbeh et al. (2018a); Simchowitz et al. (2018); Sarkar & Rakhlin (2018) and the partially observed case Hardt et al. (2018); Oymak & Ozay (2018); Simchowitz et al. (2019); Sarkar et al. (2019); Tsiamis & Pappas (2019).

Notation. The term universal constant is used for numbers which are independent of the system's (algorithm's) parameters. The operator $\succeq$ denotes comparison in the positive semi-definite cone. With $\|\|_2$ we denote the Euclidean norm for vectors and the spectral norm for matrices. The spectral radius of a matrix A is denoted by $\rho(A)$. The smallest singular value of a matrix A is denoted by $\sigma_{\min}(A)$. By A^* we denote the transpose of A. The big-O notation $O(f(N))$ represents a function that grows at most as fast as $f(N)$. The

$\tilde{O}(\cdot)$ notation hides poly-logarithmic terms. When it is not clear from the context, we write $O_x(\cdot)$ to explicitly denote big-O notation with respect to x. The poly(x) notation means a polynomial function of x.

5.2 Problem Formulation

As we covered in Chapter 2, the Kalman filter considers the problem of predicting observations generated by the following state-space LTI system:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k, \quad w_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, Q) \\ y_k &= Cx_k + v_k, \quad v_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, R) \end{aligned} \tag{5.1}$$

where $x_k \in \mathbb{R}^n$ is the state, $y_k \in \mathbb{R}^m$, $m \leq n$, are the observations (outputs), $A \in \mathbb{R}^{n \times n}$ is the system matrix and $C \in \mathbb{R}^{m \times n}$ is the observation matrix. The time series w_k, v_k represent the process and measurement noise respectively and are modeled as zero mean i.i.d. Gaussian variables, independent of each other, with covariances Q and R respectively. The initial state is zero mean Gaussian with covariance Σ_0 and independent of the noises. Similar to Chapter 3, we make the following assumption.

Assumption 5.1. *The pair (A, C) is observable, $(A, Q^{1/2})$ is stabilizable. Let Assumption 3.2 hold, i.e. the Kalman filter has already reached steady state with $\Sigma_0 = P$, where P is defined in (2.3).*

Then, by Corollary 2.1, the Kalman Filter of system (5.1) is given by:

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Ke_k, \quad \hat{x}_0 = 0 \\ y_k &= C\hat{x}_k + e_k, \end{aligned} \tag{5.2}$$

where $\hat{x}_k \in \mathbb{R}^n$ is the prediction (state), $y_k \in \mathbb{R}^m$ is the output, $e_k \in \mathbb{R}^n$ is the innovation process, and K is the Kalman filter gain defined in (2.2). Recall that the innovation process $e_k \triangleq y_k - C\hat{x}_k$ is i.i.d. Gaussian with covariance matrix $\bar{R}$, defined in (3.4). The following assumption also holds throughout the chapter.

Assumption 5.2. *Matrices $A, C, Q, R, K, \bar{R}$ are unknown. System (5.1) is non-explosive, namely the spectral radius is $\rho(A) \leq 1$. The state matrices are bounded*

$$\|A\|_2, \|C\|_2, \|Q\|_2, \|R\|_2 \leq M,$$

for some $M \geq 0$. Let $A = SJS^{-1}$ be the Jordan form decomposition of A, then the similarity transformation is bounded $\|S\|_2, \|S^{-1}\|_2 \leq M$.

Contrary to the offline architecture, we can extend the prediction guarantees to the case $\rho(A) = 1$, when the system is marginally stable or polynomially unstable; this is possible only under an online learning architecture. The upper bound M captures the size of the state parameters and we expect it to affect the learning complexity. For the economy of the presentation we only use a common bound for all parameters. The next assumption makes sure that system (5.2) is minimal.

Assumption 5.3. *The pair (A, K) is controllable.*

If the pair (A, K) is stabilizable, then the results of the chapter apply to the controllable subspace of the state; the uncontrollable subspace remains zero.

The following assumption is for notational simplicity. It assumes that the largest eigenvalue of $A - KC$ is simple. It also assumes that the responses $C(A - KC)^t K$ have bounded norm. We could remove it at the expense of slightly more complicated bounds in the regret analysis.

Assumption 5.4. *For some all $t \geq 0$, the closed-loop matrix satisfies $\left\|(A - KC)^t\right\|_2 \leq M\rho(A - KC)^t$. The Kalman Filter gain K and the innovation covariance $\bar{R}$ are bounded $\|K\|_2, \|\bar{R}\|_2 \leq M$. The responses are upper bounded $\sum_{t \geq 0} \|C(A - KC)^t K\| \leq M$.*

Let $\mathcal{F}_k \triangleq \sigma(y_0, \ldots, y_k)$ be the filtration generated by the observations $y_0, \ldots, y_k$ with $\mathcal{L}_2^m(\mathcal{F}_k)$ the space of square integrable $\mathcal{F}_k-$measurable random vectors. Given the observations up to time k, the Kalman filter gives the optimal prediction $\hat{y}_{k+1}$ at time $k + 1$ in the minimum mean square error (mmse) sense:

$$\hat{y}_{k+1} \triangleq \arg \min_{z \in \mathcal{L}_2^m(\mathcal{F}_k)} \mathbb{E}\left[\|y_{k+1} - z\|_2^2 \,|\mathcal{F}_k\right].$$

Of course to compute the Kalman filter prediction, knowledge of the system matrices A, C, Q, R is required.

Our goal is to design an online predictor $\tilde{y}_{k+1} \in \mathcal{L}_2^m(\mathcal{F}_k)$ that adapts to the unknown model and competes with the optimal Kalman filter. Since we do not have access to the model, we can only rely on the data, which are revealed sequentially. To quantify the online prediction performance, we define the regret of our online learning algorithm with respect to the Kalman filter (5.2) that has full knowledge of system model (5.1). Our goal is to achieve sublinear regret, as defined in the following problem statement.

Problem 5.1. *Assume that A, C, Q, R in system model (5.1) are unknown. Consider a sequence $y_0, y_1 \dots$ of observations, which are generated sequentially by system (5.1). Let $\tilde{y}_k \in \mathcal{L}_2^m(\mathcal{F}_k)$ be the prediction of an online learning algorithm based on the history $y_{k-1}, \dots, y_0$ and $\hat{y}_k$ be the Kalman filter prediction (5.2) that has full knowledge of model (5.1). Define the regret:*

$$\mathcal{R}_N \triangleq \sum_{k=1}^{N} \|y_k - \tilde{y}_k\|^2 - \sum_{k=1}^{N} \|y_k - \hat{y}_k\|^2. \tag{5.3}$$

Fix a failure probability $\delta > 0$. Our goal is to find a learning algorithm such that with probability at least $1 - \delta$:

$$\mathcal{R}_N \le \text{poly}(\log 1/\delta) o(N),$$

where $o(N)$ does not depend on δ.

The regret captures the average suboptimality of the online predictor. From this point of view, if the regret is sublinear this implies average convergence since $\mathcal{R}_N/N = o(1)$.

Recall that one of the main properties of the Kalman Filter is that the innovation sequence $e_k = y_k - \hat{y}_k$ is orthogonal and, by Gaussianity, also i.i.d. By the law of large numbers, this implies that the ℓ_2 accumulative error $\sum_{k=0}^{N} \|y_k - \hat{y}_k\|_2^2$ will be of the order of $\Omega(N)$ almost surely. Predicting the true observations exactly is impossible in the stochastic noise setting, even if we know the system.

Systems (5.1), (5.2) are statistically equivalent, i.e. they generate observations with the same distribution. This implies that the noise parameterization is not unique Van Overschee & De Moor (2012). Another source of ill-posedness is that the state space parameterization is non-unique. Any similar system $S^{-1}AS$, CS, $S^{-1}QS^{-*}$ generates the same observations. These problems will be addressed later by considering an alternative system representation.

5.3 Online Prediction Algorithm

Our online prediction algorithm is based on a system response representation (Markov parameterization) that has been used in subspace system identification Bauer et al. (1999). Let p be an integer that represents how far we look into the past. We define the vector of past observations at time k:

$$Z_{k,p} \triangleq \begin{bmatrix} y_{k-p}^* & \dots & y_{k-1}^* \end{bmatrix}^*, \, k \geq p. \tag{5.4}$$

Define also the matrix of closed-loop responses:

$$G_p \triangleq \begin{bmatrix} C(A-KC)^{p-1}K & \cdots & CK \end{bmatrix}, \tag{5.5}$$

which are essentially the Markov parameters of the stochastic system. By expanding the Kalman filter (5.2) p-steps into the past, the observation at time k can be rewritten as

$$y_k = G_p Z_{k,p} + \underbrace{C(A-KC)^p \hat{x}_{k-p}}_{\text{bias}} + e_k. \tag{5.6}$$

Instead of optimizing over system parameters A, C, K, which results in a non-convex optimization problem, we optimize over (the higher dimensional) G_p, which makes the problem convex. From an online learning perspective, this technique is also known as improper

Algorithm 2 Online Prediction Algorithm

```
Input: $\beta$, $\lambda$, $T_{\text{init}}$
for k $= 0, \dots, T_{\text{init}} - 1$ do
    Observe $y_k$
end for
for i $= 1, 2, \dots$ do
    $T = 2^{i-1} T_{\text{init}}$
    $p = \beta \log T$
    $\bar{V}_{T-1} = \lambda I + \sum_{t=p}^{T-1} Z_{t,p} Z_{t,p}^*$
    $\tilde{G}_{T-1} = \left( \sum_{t=p}^{T-1} y_t Z_{t,p}^* \right) \bar{V}_{T-1}^{-1}$
    for $k = T, \dots, 2T - 1$ do
        Predict $\tilde{y}_k = \tilde{G}_{k-1} Z_{k,p}$
        Observe $y_k$
        Update $\bar{V}_k = \bar{V}_{k-1} + Z_{k,p} Z_{k,p}^*$
        $\tilde{G}_k = \tilde{G}_{k-1} + (y_k - \tilde{y}_k) Z_{k,p}^* \bar{V}_k^{-1}$
    end for
end for
```

learning. Using this lifting, we can learn a least squares estimate $\tilde{G}_{k,p}$ by regressing outputs y_t to past outputs $Z_{t,p}$ for $t \leq k$:

$$\tilde{G}_{k,p} = \sum_{t=p}^{k} y_t Z_{t,p}^* \left(\lambda I + \sum_{t=p}^{k} Z_{t,p} Z_{t,p}^* \right)^{-1}, \tag{5.7}$$

where $\lambda > 0$ is a regularization parameter to be designed. Then, to predict the next observation, we can compute:

$$\tilde{y}_{k+1} = \tilde{G}_{k,p} Z_{k+1,p}. \tag{5.8}$$

Due to the stability properties of the Kalman filter (Chapter 2), if we consider p past observations, then the bias term in equation (5.6) is of the order of $\rho(A - KC)^p \left\| \hat{x}_{x-p} \right\|_2$. Notice that for non-explosive systems the state $\hat{x}_{k-p}$ can grow polynomially fast in the worst case. Even if $\hat{x}_{k-p}$ remains bounded, keeping the past p constant would lead to a non-vanishing bias error (linear regret). Thus, to make sure that the prediction error decreases, we need to gradually increase the past horizon p. For this reason, inspired by the "doubling trick" Cesa-Bianchi & Lugosi (2006), we divide the learning in epochs, where every epoch is twice longer than the previous one. During every epoch we keep the past horizon constant. Since $\rho(A-KC)^p$ is exponentially decreasing, it is sufficient to slowly increase the past as $p = O(\log T)$, where T is the epoch duration.

The pseudo-code of our online prediction approach can be found in Algorithm 2. Each epoch lasts from time $T_i, \dots, 2T_i - 1$, where $i = 1, \dots,$ is the epoch, $T_i = 2^{i-1}T_{\text{init}}$, and T_{init} is a design parameter (the length of the first epoch). During every epoch, we keep the past $p_i = \beta \log(T_i)$ constant, where β is a design parameter. Initially, from time 0 to $T_{\text{init}} - 1$, we have a warm-up phase where we gather enough observations to start predicting. To make sure that $p_i < T_i$, we tune T_{init} accordingly. Within an epoch, the least squares based predictor (5.8) can be implemented in a recursive way:

$$\begin{aligned}\bar{V}_{k,p_i} &= \bar{V}_{k-1,p_i} + Z_{k,p_i}Z^*_{k,p_i}\\ \tilde{G}_{k,p_i} &= \tilde{G}_{k-1,p_i} + (y_k - \tilde{y}_k)Z^*_{k,p_i}\bar{V}_k^{-1},\end{aligned}$$

which requires polynomial complexity and at most $O(\log T_i)$ memory. In the beginning of an epoch, when p_i is updated, we re-initialize the predictor based on the whole past, which requires polynomial complexity and $O(T_i)$ memory. Hence, in total, after N collected samples, the computational complexity is polynomial and the memory requirement is linear $O(N)$.

No knowledge about the dynamics or even the state dimension n is required. We only need to know an upper bound on the state dimension n in order to tune the past horizon p–see Theorem 5.1. There is a tradeoff between the bias error and statistical efficiency. Increasing the past horizon by selecting larger β leads to smaller bias error, but increases the sample complexity of learning G_p since we have more unknowns.

5.4 Regret Analysis

In this section, we prove that with high probability the prediction regret is not only sublinear, but also of the order of poly $\log N$ (or $\tilde{O}(1)$), where N is the number of observations collected so far. The challenge in the non-explosive regime is that the observations grow unbounded polynomially fast ($\Omega(\sqrt{N})$). Meanwhile, recent work in finite sample analysis of system identification Oymak & Ozay (2018); Simchowitz et al. (2019); Tsiamis & Pappas (2019); Sarkar et al. (2019) shows that the model parameters can be learned at a slower rate ($O(1/\sqrt{N})$). Therefore these system identification results cannot be directly applied to obtain regret bounds for our problem. Nonetheless, we show that our online Algorithm 2 mitigates the effect of unbounded observations. As a result, the logarithmic regret bound of $\tilde{O}(1)$ remains valid even as we approach instability.

We provide two results, one for non-explosive systems ($\rho(A) \leq 1$) and one for stable systems ($\rho(A) < 1$). Before we present the regret results, let us introduce some notion. Let $a(s) = s^d - a_{d-1}s^{d-1} \cdots - a_0$ be the minimal polynomial of matrix A, i.e. the minimum degree polynomial such that $a(A) = 0$. Denote its degree by d. We define the ℓ_1 norm of its coefficients as $\|a\|_1 \triangleq 1 + \sum_{i=0}^{d-1} |a_i|$; the ℓ_2 norm $\|a\|_2$ is defined in a similar way. Let κ be the dimension of the largest Jordan block of A. In general, $\kappa \leq d \leq n$. The minimum singular value of R is denoted by σ_R. For brevity, we omit the first T_{init} terms from the regret. We also do not show the dependence on M and λ in the theorem statements. More precise bounds can be inferred by the proofs, at the cost of more complicated and less intuitive expressions.

Theorem 5.1 (Regret for non-explosive systems)**.** *Consider system* (5.2) *with* $\rho(A) \leq 1$*. Let* $y_0, \dots, y_N$ *be the sequence of system observations. Let* $\tilde{y}_0, \dots, \tilde{y}_N$ *be the predictions of Algorithm* 2 *with*

$$\beta = \Omega\left(\frac{\kappa}{\log(1/\rho(A - KC))}\right) \tag{5.9}$$

and fix a failure probability $\delta > 0$*. There exists a* $N_0 = \text{poly}\left(n, \beta, \log 1/\delta, 1/\sigma_R\right)$*, independent of* N*, such that with probability at least* $1 - \delta$*, if* $N > N_0$ *then:*

$$\mathcal{R}_N \leq \text{poly}(d^\kappa, n, \beta, \|a\|_2, \log 1/\delta, 1/\sigma_R)\tilde{O}(1). \tag{5.10}$$

Theorem 5.1 provides the first logarithmic regret upper bounds for the general problem of Kalman filter prediction. The burn-in time N_0 is related to persistency of excitation conditions, i.e. initially we need enough samples to guarantee that the smallest singular value of the Gram matrix $\bar{V}_k$ increases linearly with k. Our bounds do not depend on the stability gap $1/(1-\rho(A))$ and they do not degrade as we approach instability. However, they suggest, via β, that the stability radius $\rho(A-KC)$ of the Kalman filter closed-loop matrix affects the learning difficulty.

The upper bound also depends on the quantities d^{κ} and $\|a\|$, both of which can be exponential in the dimension of the system state n in the worst case. This can happen, for example, if $\kappa = n$, i.e. the system is an $n-$th order integrator. Dependence of learning performance on the coefficients of the characteristic or minimal polynomial has been found in related work Hardt et al. (2018). This dependence can be improved–see for example Hazan et al. (2018), Simchowitz et al. (2019), Ghai et al. (2020). However, it is an open question whether it is possible to avoid the exponential dependence on κ. It might be possible that systems with long chain structure, e.g. integrators, are harder to learn; such systems are difficult to observe even in the known model case. In open-loop system identification Simchowitz et al. (2019), such a dependence also appears. It might be an inherent limitation of the problem, since fundamental quantities of the system, for example matrix A^i or the observability matrix $\mathcal{O}_i$ scale with i^{κ}–see Lemma 5.5.

The above issues can be avoided in the case of stable systems ($\rho(A) < 1$), by exploiting stationarity. Let $\Gamma_k \triangleq \mathbb{E}\hat{x}_k\hat{x}_k^*$ be the covariance of the state with $\Gamma_\infty = \lim \Gamma_k$ the steady-state covariance. Let the *mixing time* τ_{mix} be the minimum time such that the system is close to steady-state:

$$\|\Gamma_\infty^{-1/2}\Gamma_k\Gamma_\infty^{-1/2}\|_2 \leq 1/2, \text{ for all } k \geq \tau_{\text{mix}}. \tag{5.11}$$

Based on Lemma 5.7, the mixing time scales with:

$$\tau_{\text{mix}} \leq \text{poly}((\log 1/\rho(A))^{-1}, \kappa, \log \tfrac{\sigma_{\max}(\Gamma_\infty)}{\sigma_{\min}(\Gamma_\infty)}).$$

We obtain the following guarantees.

Theorem 5.2 (Regret for stable systems)**.** *Consider system* (5.2) *with* $\rho(A) < 1$*. Let* $y_0, \dots, y_N$ *be the sequence of system observations. Let* $\tilde{y}_0, \dots, \tilde{y}_N$ *be the predictions of Algorithm* 2 *with* β *as in* (5.9)*. Fix a failure probability* $\delta > 0$*. There exists*

$$N_0^{\mathrm{s}} = \mathrm{poly}\left(n, \beta, \log 1/\delta, \tau_{\mathrm{mix}}, 1/\sigma_R\right)$$

such that with probability at least $1-\delta$*, if* $N > N_0^{\mathrm{s}}$ *then:*

$$\mathcal{R}_N \leq \mathrm{poly}(n, \beta, \log 1/\delta)\tilde{O}(1). \tag{5.12}$$

Notice that for stable systems we no longer have quantities that depend exponentially on the dimension n. The main bound (5.12) does not depend on the stability gap $1/(1-\rho(A))$. However, via τ_{mix} in N_0^{s}, the guarantees depend logarithmically on the inverse radius $\log 1/\rho(A)$. This quantity affects the mixing time needed for a stable system to approach stationarity.

The proofs of Theorem 5.1 and Theorem 5.2 can be found in Sections 5.7 and 5.8 respectively. In the next subsection, we provide an overview of the regret analysis and explain why the quantities d^{κ} and $\|a\|_2$ appear in the bound in Theorem 5.1. We also explain what changes in the case of stable systems addressed by Theorem 5.2.

5.4.1 Regret analysis overview

Recall the definition of the innovation error $e_k = y_k - \hat{y}_k$. For brevity, we also define the error $\tilde{e}_k \triangleq \tilde{y}_k - \hat{y}_k$ between the online prediction of Algorithm 2 and the Kalman Filter prediction. Adding and subtracting $\hat{y}_k$ in the first term:

$$\begin{aligned}\mathcal{R}_N &= \sum_{k=T_{\mathrm{init}}}^{N} \|e_k + \hat{y}_k - \tilde{y}_k\|_2^2 - \|e_k\|_2^2 \\ &= \underbrace{\sum_{k=T_{\mathrm{init}}}^{N} \|\hat{y}_k - \tilde{y}_k\|_2^2}_{\mathcal{L}_N} + 2 \underbrace{\sum_{k=T_{\mathrm{init}}}^{N} e_k^* \left(\hat{y}_k - \tilde{y}_k\right)}_{\text{martingale term}}\end{aligned}$$

It is sufficient to prove that the square loss ℓ_2:

$$\mathcal{L}_N \triangleq \sum_{k=T_{\text{init}}}^{N} \|\hat{y}_k - \tilde{y}_k\|_2^2 \tag{5.13}$$

is logarithmic in N. Because the innovations are i.i.d., we have a martingale structure for the second term since $e_k \in \mathcal{F}_k$, while $\tilde{e}_k \in \mathcal{F}_{k-1}$. The martingale term is small and can be bounded in terms of the square loss $\mathcal{L}_N$. In particular, the quantity

$$(\mathcal{L}_N + 1)^{-1/2} \sum_{k=T_{\text{init}}}^{N} e_k^* \left(\hat{y}_k - \tilde{y}_k\right)$$

is a self-normalized martingale. In Chapter 3, we analyzed self-normalized martingales using tools developed in Abbasi-Yadkori et al. (2011), Sarkar & Rakhlin (2018). In particular, we utilized Theorem 3.3, which extends Theorem 1 in Abbasi-Yadkori et al. (2011) and Proposition 8.2 in Sarkar & Rakhlin (2018). We repeat the statement of Theorem 3.3 below for convenience.

Theorem 5.3 (Self-normalized martingales Theorem 3.3)**.** *Let $\{\mathcal{F}_t\}_{t=0}^{\infty}$ be a filtration. Let $\eta_t \in \mathbb{R}^m$, $t \geq 0$ be $\mathcal{F}_t$-measurable, independent of $\mathcal{F}_{t-1}$. Suppose also that $\eta_t \sim \mathcal{N}(0, I)$ is isotropic Gaussian. Let $X_t \in \mathbb{R}^d$, $t \geq 0$ be $\mathcal{F}_{t-1}-$measurable. Assume that V is a $d \times d$ positive definite matrix. For any $t \geq 0$, define:*

$$\bar{V}_t = V + \sum_{s=1}^{t} X_s X_s^*, \qquad S_t = \sum_{s=1}^{t} X_s H_s^*$$
$$H_s^* = \begin{bmatrix} \eta_s^* & \dots & \eta_{s+r-1}^* \end{bmatrix} \in \mathbb{R}^{rm},$$

for some integer r. Then, for any $\delta > 0$, with probability at least $1 - \delta$, for all $t \geq 0$

$$\left\|\bar{V}_t^{-1/2} S_t\right\|_2^2 \leq 8r \left(\log \frac{r5^m}{\delta} + \log\det \bar{V}_t V^{-1}\right)$$

An application of the above theorem implies that

$$\sum_{k=T_{\text{init}}}^{N} e_k^* \left(\hat{y}_k - \tilde{y}_k\right) = \tilde{O}(\sqrt{\mathcal{L}_N}),$$

with high probability. Hence, we focus on bounding $\mathcal{L}_N$.

For the remaining section, we will assume that we are within one epoch i so that the past horizon $p = p_i$ and $T = 2^{i-1}T_{init}$ are kept constant. For brevity, we omit the subscript p from all variables and write $G, \tilde{G}_k, Z_k$ instead of $G_p, \tilde{G}_{k,p}, Z_{k,p}$.

Define $S_{k-1} \triangleq \sum_{i=p}^{k-1} e_i Z_i^*$ and $\bar{V}_{k-1} \triangleq \lambda I + \sum_{i=p}^{k-1} Z_i Z_i^*$. Then, the error between our online prediction and the Kalman filter prediction can be written as:

$$\begin{aligned}\tilde{e}_k &= (\tilde{G}_{k-1} - G)Z_k - C(A-KC)^p \hat{x}_{k-p} \\ &= \underbrace{S_{k-1}\bar{V}_{k-1}^{-1}Z_k}_{\text{regression}} + \underbrace{\lambda G \bar{V}_{k-1}^{-1} Z_k}_{\text{regularization}} \\ &\quad \underbrace{+\, C(A-KC)^p \left(\sum_{i=T}^{k-1} \hat{x}_{i-p} Z_i^* \bar{V}_{k-1}^{-1} Z_k - \hat{x}_{k-p} \right)}_{\text{truncation bias}}.\end{aligned} \tag{5.14}$$

The regression term is due to the noise e_k perturbing the system. The truncation bias is due to using only p past observations and not all of them. The key ingredients to analyze the cumulative error $\mathcal{L}_N$ are i) the stability properties of the closed-loop matrix $A - KC$; ii) self-normalization properties of predictor (5.8); and iii) persistency of excitation for the past observations with high probability. By persistency of excitation we mean that the least singular value of the Gram matrix $\bar{V}_k$ is increasing as fast as $O(k)$ with high probability.

Regression term. We can rewrite the regression term as a product of two separate terms:

$$S_{k-1}\bar{V}_{k-1}^{-1}Z_k = (S_{k-1}\bar{V}_{k-1}^{-1/2})(\bar{V}_{k-1}^{-1/2}Z_k).$$

The first term, $S_{k-1}\bar{V}_{k-1}^{-1/2}$ is a self-normalized martingale and can be bounded based on Theorem 5.3 and Lemma 5.28. Thus, the term $\|S_{k-1}\bar{V}_{k-1}^{-1/2}\|_2$, grows at most logarithmically with k. The martingale property comes from the fact that the innovation process e_k of the Kalman Filter is i.i.d.–see Chapter 2.

The second term, $\bar{V}_{k-1}^{-1/2} Z_k$, is almost self-normalized since $\bar{V}_{k-1}$ is the Gram matrix of $Z_{k-1}, \dots, Z_p$. It could be bounded using the following lemma which is inspired by Lai & Wei (1982).

Lemma 5.1. *Let $\bar{V}_{k-1} = \lambda I + \sum_{i=p}^{k-1} Z_i Z_i^*$. Then:*

$$\sum_{k=T+1}^{2T} Z_{k-1} \bar{V}_{k-1}^{-1} Z_{k-1} \leq \log\det(\bar{V}_{2T-1} \bar{V}_{T-1}^{-1})$$

The intuition is that $Z_{k-1} Z_{k-1}^*$ appears in $\bar{V}_{k-1}$ and, hence, it cancels out the effect of Z_{k-1}. Unfortunately, we cannot directly use the above inequality for $\bar{V}_{k-1}^{-1/2} Z_k$ since $Z_k Z_k^*$ is not explicitly contained in $\bar{V}_{k-1}$. However, we can exploit the fact that Z_k does not change too fast compared to the most recent past Z_{k-1}, ..., Z_{k-n}.

Lemma 5.2 (ARMA-like representation)**.** *Let $y_0, y_1 \dots$ be observations generated by system (5.1). Fix a past horizon p and let a be the minimal polynomial of A with degree d. Then, the past observations satisfy the following recursion*

$$Z_k = a_{d-1} Z_{k-1} + \cdots + a_0 Z_{k-d} + \delta_k, \tag{5.15}$$

where $\delta_k \in \mathcal{F}_{k-1}$ with

$$\|\delta_k\|_2 \leq \Delta \sup_{i \leq k-1} \|e_i\|_2 \,, \tag{5.16}$$

and $\Delta = O(d^{\kappa-1} \|a\|_1 \sqrt{p})$

The proof is in Section 5.7.2. Intuitively, the unbounded components of Z_k are captured by the recent history Z_{k-1}, ..., Z_{k-d} and the residual δ_k is bounded. Replacing Z_k with (5.15) we obtain by two Cauchy-Schwarz inequality applications:

$$\|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2 \leq 2 \|a\|_2^2 \sum_{i=0}^{d-1} \|\bar{V}_{k-1}^{-1/2} Z_{k-d+i}\|_2^2 + 2\|\bar{V}_{k-1}^{-1/2} \delta_k\|_2^2.$$

The terms $\bar{V}_{k-1}^{-1/2} Z_{k-d+i}$ in the sum are now indeed normalized and can be bounded using Lemma 5.1. For $\bar{V}_{k-1}^{-1/2}\delta_k$ we exploit a new persistency of excitation result.

Lemma 5.3 (Uniform Persistency of Excitation)**.** *Consider the conditions of Theorem 5.1. Select a failure probability $\delta > 0$. Let $T = 2^{i-1}T_{init}$ for some fixed epoch i with $p = \beta \log T$ the corresponding past horizon. There exists a $N_0 = \mathrm{poly}(n, \beta, \log 1/\delta)$ such that if $T \geq N_0$, then with probability at least $1-\delta$:*

$$\sum_{j=p}^{k} Z_j Z_j^* \succeq \frac{k}{4}\sigma_R I, \tag{5.17}$$

uniformly for all $T \leq k \leq 2T-1$, where $\sigma_R = \sigma_{\min}(R)$.

The above persistency of excitation condition holds uniformly over all times k as long as $k \geq N_0$. This is why the burn-in time N_0 appears in Theorem 5.1; if k is very small, then matrix $\sum_{j=p}^{k} Z_j Z_j^*$ might not be invertible.

Regularization and Truncation terms. For the regularization term we follow the same steps as with the regression one. Since matrix $A-KC$ is stable, the truncation term decreases exponentially fast with p. System quantity κ governs how fast the observations grow polynomially. Parameter β should be large enough to cancel out this polynomial rate. This explains why κ affects the choice of β in (5.9).

Stable Systems. If $\rho(A) < 1$, then we can exploit the fact that Z_k converges exponentially fast to a stationary distribution. Hence the term $\bar{V}_{k-1}^{-1/2} Z_k$ will effectively be self-normalized, without the need to express Z_k as a function of the past observations. In particular, for stable systems we prove a new and stronger persistency of excitation result. Similar to the definition of the state covariance, we define the past outputs' covariance $\Gamma_{Z,k} \triangleq \mathbb{E} Z_k Z_k^*$. We have the following.

Lemma 5.4 (Uniform Persistency of Excitation: Stable case)**.** *Consider the conditions of Theorem 5.2. Select a failure probability $\delta > 0$. Let $T = 2^{i-1}T_{init}$ for some fixed epoch i with $p = \beta \log T$ the corresponding past horizon. There exists a $N_0^{\mathrm{s}} = \mathrm{poly}(n, \beta, \tau_{\mathrm{mix}}, \log 1/\delta)$*

such that if $T \geq N_0^{\mathrm{s}}$*, with probability at least* $1-\delta$*:*

$$\sum_{j=p}^{k-1} Z_j Z_j^* \succeq \frac{k}{8}\Gamma_{Z,\infty}, \tag{5.18}$$

uniformly for all $T \leq k \leq 2T-1$*.*

Hence, the term $\bar{V}_{k-1}^{-1/2} Z_k$ can be bounded by:

$$\left\| \bar{V}_{k-1}^{-1/2} Z_k \right\|_2 \leq O\left(\frac{1}{\sqrt{k}}\right) \left\| \Gamma_{Z,k}^{-1/2} Z_k \right\|_2$$

where now the normalized term $\Gamma_{Z,k}^{-1/2} Z_k$ behaves like a standard isotropic Gaussian variable.

5.4.2 Numerical example

For a sanity check, we perform simulations of our algorithm for a simple example. Let

$$A = \begin{bmatrix} 0.9 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, Q = I, R = 1.$$

The system is non-explosive with $\kappa = 2$. We select $T_{\text{init}} = 20$, $\beta = 2$, $\lambda = 1$, $N = 500$. We ignore the errors for the initial epoch until T_{init}. We perform 1000 Monte Carlo simulations. For every Monte Carlo simulation, we compute the regret trajectory $\mathcal{R}_k$, $k \leq N$. In Fig. 5.1, show the mean, and the empirical 80% and 90% percentiles of the regret trajectories $\mathcal{R}_k$, $k \leq N$. The simulations show that the regret is indeed logarithmic with high probability.

5.5 Discussion and Extensions

In this section, we discuss implications and generalizations of Algorithm 2 and the regret analysis.

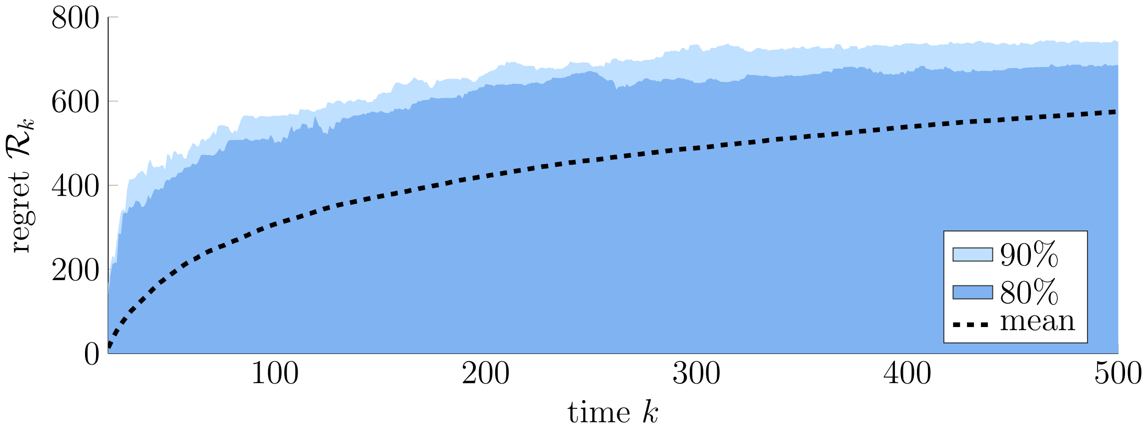


Figure 5.1: The 80% and 90% empirical percentiles for the regret trajectories $\mathcal{R}_k$, $k \leq N$, for $N = 500$. We run 1000 Monte Carlo simulation in total. The regret is indeed logarithmic with the number of samples.

Comparison with online LQR Our bounds show that the problem of learning the Kalman filter online is provably easier than the online LQR, in the case of unknown systems. The latter requires in general regret of the order of $\Omega(\sqrt{N})$ Simchowitz & Foster (2020); Ziemann & Sandberg (2020). This is another reason why the problems are not dual in the unknown model case Tsiamis et al. (2020). This gap might be expected since in the case of control, there is a need for exogenous exploratory signals, which increase the LQR control cost.

Parameter estimation Our regret analysis relied on proving persistency excitation in finite time of the past outputs. A byproduct of this result is that we can also obtain parameter estimation gurantees for the system responses $C(A-KC)^iK$. In fact, following a similar approach as in Sarkar & Rakhlin (2018); Tsiamis & Pappas (2019), we can show that the estimation error $\|\tilde{G}_{k,p} - G_p\|_2$ will decay with a rate of $\tilde{O}(1/\sqrt{k})$ with high probability. This follows from the fact that if we ignore all other error components:

$$\tilde{G}_{k,p} - G_p \approx S_{k-1}\bar{V}_{k-1}^{-1} = (S_{k-1}\bar{V}_{k-1}^{-1/2})\bar{V}_{k-1}^{-1/2},$$

where we used the notation of the previous section. We already showed previously that the first term is a self-normalized matringale and can grow at most logarithmically. Meanwhile, the second term is decaying as fast as $O(1/\sqrt{k})$ thanks to persistency of excitation. Based on the estimate of the responses we can obtain an estimate of the state space matrices $\tilde{A}, \tilde{C}, \tilde{K}$ (up to a similarity transformation) using any realization algorithm, e.g. the Ho-Kalman algorithm Oymak & Ozay (2018).

f-steps ahead predictor An immediate generalization of Algorithm 2 is to consider the $f-$steps ahead predictor, where f is some future horizon. Instead of predicting only the next observation, we predict the sequence $y_k, \ldots, y_{k+f-1}$. Denote the future observations and noises by:

$$\begin{aligned} Y_k &= \begin{bmatrix} y_k^* & \cdots & y_{k+f-1}^* \end{bmatrix}^* \\ E_k^+ &= \begin{bmatrix} e_k^* & \cdots & e_{k+f-1}^* \end{bmatrix}^*. \end{aligned}$$

Similar to (5.4), we can establish a regression:

$$Y_k = \mathcal{O}_f \mathcal{K}_p Z_k + \mathcal{O}_f (A - KC)^p \hat{x}_{k-p} + \mathcal{T}_f E_k^+$$

where $\mathcal{K}_p \triangleq \begin{bmatrix} (A-KC)^{p-1}K & \cdots & K \end{bmatrix}$, and $\mathcal{T}_f$ is a lower triangular block Toeplitz matrix generated by $I, CK, \ldots, CA^{f-2}K$. The optimal Kalman filter predictor is:

$$\hat{Y}_k = \mathcal{O}_f \mathcal{K}_p Z_k + \mathcal{O}_f (A - KC)^p \hat{x}_{k-p}$$

Hence, the regret can be defined as in (5.3), with the lowercase y replaced with uppercase Y. The online predictor (5.8) can be adapted here:

$$\tilde{Y}_k = \tilde{G}_{k,f,p} Z_k,$$

where $\tilde{G}_{k,f,p}$ is obtained similar to (5.7) by regressing future observations Y_t to past observations Z_t from time p up to $k-f$. The logarithmic regret guarantees of $\tilde{O}(1)$ also hold with the final bound depending polynomially on f and $\|\mathcal{T}_f\|_2$.

State prediction If we have some knowledge about the state, e.g. the state space basis and the state space dimension n, then we can use the $f-$step ahead predictor to predict the hidden state $\hat{x}_k$. Notice that the Kalman filter state prediction $\hat{x}_k$ can be rewritten as:

$$\hat{x}_k = \mathcal{K}_p Z_k + (A - KC)^p \hat{x}_{k-p} = \mathcal{O}_f^\dagger \hat{Y}_k$$

If we know $\mathcal{O}_f$ and the future horizon is large enough $f \geq n$ we can compute the state prediction:

$$\tilde{x}_k = \mathcal{O}_f^\dagger \tilde{Y}_k,$$

where $\tilde{Y}_k$ is our $f-$step ahead prediction and $\dagger$ denotes the pseudo-inverse. In this case the regret:

$$\mathcal{R}_{x,N} \triangleq \sum_{k=1}^{N} \|x_k - \tilde{x}_k\|^2 - \sum_{k=1}^{N} \|x_k - \hat{x}_k\|^2 \tag{5.19}$$

will enjoy the same logarithmic guarantees. Hence, our algorithm can be used to solve the adaptive Kalman filter problem posed in Mehra (1970); Anderson & Moore (2005), where the dynamics A, C are known but the noise statistics Q, R are unknown, with logarithmic regret.

Closed-loop prediction We can apply the same algorithm to closed-loop control systems Lee & Lamperski (2020):

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Bu_k + Ke_k & \qquad s_{k+1} &= A_c s_k + B_c y_k + \eta_k \\ y_k &= C\hat{x}_k + e_k & u_k &= C_c s_k + D_c y_k + \zeta_k, \end{aligned}$$

where s_k is the internal state of the feedback controller. Extend the vector of past outputs $Z_{k,p}^{\mathrm{cl}}$ to also include all past internal states $s_{k-1}, \dots, s_{k-p}$ and all inputs $\zeta_{k-1}, \dots, \zeta_{k-p}$. Then similar to (5.6) we have an approximate linear relation:

$$y_k = G_p^{\mathrm{cl}} Z_{k,p}^{\mathrm{cl}} + e_k + C(A - KC)^p \hat{x}_{k-p},$$

for some appropriate matrix of closed-loop responses G_p^{cl}. The online predictor will be of the form:

$$\tilde{y}_k = \tilde{G}_{k,p}^{\mathrm{cl}} Z_{k,p}^{\mathrm{cl}},$$

where $\tilde{G}_{k,p}^{\mathrm{cl}}$ is obtained by regressing observations to past observations and past controller states. If the control law is internally stable, then we can apply Theorem 5.2 to obtain logarithmic regret guarantees.

5.6 Conclusion and Future Work

We provided the first logarithmic regret upper bounds for learning the classical Kalman filter of an unknown system with unknown stochastic noise. Our regret analysis holds for non-explosive systems and our bounds do not degrade with the system stability gap. Going forward, our work opens up several research directions. Our current bounds are mainly qualitative and data-independent, focusing on how various system theoretic properties affect learning. It is an open problem to develop sharper data-dependent bounds that are more suitable for control applications, possibly at the cost of losing some interpretability. Another interesting direction is to study how the learning performance is affected by system theoretic properties, such as the exponential quantity d^κ in the case of systems with long chain structure, e.g. κ-order integrators. Analyzing the regret of other online algorithms, e.g. extended least squares, is also an open problem. Another challenging problem for both prediction and system identification is the case of explosive systems. Although in the fully observed case, this problem has been studied Faradonbeh et al. (2018a); Sarkar & Rakhlin (2018), it remains open in the case of partially observable systems.

Table 5.1: Notation table for fixed past horizon p

Z_t	$\triangleq \begin{bmatrix} y_{t-p}^* & \cdots & y_{t-1}^* \end{bmatrix}^*$	Past outputs at time t
E_t	$\triangleq \begin{bmatrix} e_{t-p}^* & \cdots & e_{t-1}^* \end{bmatrix}^*$	Past noises at time t
$\bar{Z}_k$	$\triangleq \begin{bmatrix} Z_p & \cdots & Z_k \end{bmatrix}$	Batch past outputs up to time k
$\bar{E}_k$	$\triangleq \begin{bmatrix} E_p & \cdots & E_k \end{bmatrix}$	Batch past noises up to time k
$\bar{X}_k$	$\triangleq \begin{bmatrix} \hat{x}_0 & \cdots & \hat{x}_{k-p} \end{bmatrix}$	Batch past states up to time k
S_k	$\triangleq \sum_{t=p}^{k} e_t Z_t^*$	Correlation of current noise with past outputs
V_k	$\triangleq \bar{Z}_k \bar{Z}_k^* = \sum_{t=p}^{k} Z_t Z_t^*$	Gram matrix of past outputs
$\bar{V}_k$	$\triangleq \lambda I + V_k$	Regularized Gram matrix of past outputs
$\bar{R}$	$\triangleq \mathbb{E} e_k e_k^*$	Covariance of innovations
$\mathcal{T}_k$	$\triangleq \text{Toep}\left(I, CK, \dots, CA^{k-2}K\right)$	Toeplitz matrix of responses CA^tK (5.23)
Σ_E	$\triangleq \mathbb{E}\mathcal{T}_p E_t E_t^* T_p^*$	Covariance of weighted past noises
σ_R	$\triangleq \sigma_{\min}(R)$	Smallest singular value of R
Γ_t	$\triangleq \mathbb{E}\hat{x}_t \hat{x}_t^*$	Covariance of Kalman filter state prediction
$\Gamma_{Z,t}$	$\triangleq \mathbb{E} Z_t Z_t^*$	Covariance of past outputs
G	$\triangleq \begin{bmatrix} C(A-KC)^{p-1}K & \dots & CK \end{bmatrix}$	Kalman filter responses

5.7 Regret analysis for non-explosive systems

5.7.1 Notation and organization

Structure In Sections 5.7.2, 5.7.3 we review fundamental results from system theory and statistics. These provide the main tools for proving Theorems 5.1, 5.2. In Section 5.7.4, we provide finite sample complexity bounds and persistency of excitation (PE) results for a fixed time k and fixed past horizon p. In Section 5.7.5, we generalize those results from pointwise to uniform over all times k in one epoch. In Section 5.7.6 we prove Lemma 5.1. By combining the uniform bounds and Lemma 5.1, we prove in Section 5.7.7 that the square loss suffered within one epoch is logarithmic with respect the length of the epoch. Hence, we can now prove Theorem 5.1–see Section 5.7.8. In Section 5.8, we analyze the case of stable systems and prove Theorem 5.2. Section 5.9 includes some technical results about logarithmic inequalities, which are used to show that the burn-in time N_0 depends polynomially on the various system parameters.

Notation A summary of the notation can be found in Table 5.1. We will analyze the performance of Algorithm 2 based mainly on a fixed epoch i. Since the past horizon p is kept constant during an epoch, we drop the index p from $Z_{k,p}$, G_p, $\tilde{G}_{k,p}$, $\bar{V}_{k,p}$ and write Z_k, G, $\tilde{G}_k$, $\bar{V}_k$ instead. Similar to the past outputs Z_k, we also define the past noises:

$$E_k \triangleq \begin{bmatrix} e^*_{t-p} & \cdots & e^*_{t-1} \end{bmatrix}^* \tag{5.20}$$

The batch past outputs, batch past noises, and batch past Kalman filter states are defined as:

$$\begin{aligned} &\bar{Z}_k \triangleq \begin{bmatrix} Z_p & \cdots & Z_k \end{bmatrix}, \bar{E}_k \triangleq \begin{bmatrix} E_p & \cdots & E_k \end{bmatrix}, \\ &\bar{X}_k \triangleq \begin{bmatrix} \hat{x}_0 & \cdots & \hat{x}_{k-p} \end{bmatrix} \end{aligned} \tag{5.21}$$

Recall the definition of the correlations between the current innovation and the past outputs $S_k \triangleq \sum_{t=p}^{k} e_t Z_t^*$ and the regularized autocorrelations of past outputs $\bar{V}_k \triangleq \lambda I + \bar{Z}_k \bar{Z}_k^*$. The innovation sequence e_k is i.i.d. zero mean Gaussian. Its covariance has a closed-form expression and is defined as:

$$\bar{R} \triangleq \mathbb{E} e_k e_k^* = CPC^* + R, \tag{5.22}$$

where P is the solution to the Riccati equation (2.3). Next we define the Toeplitz matrix $\mathcal{T}_k$, for some $k \geq 1$:

$$\mathcal{T}_k \triangleq \begin{bmatrix} I_m & 0 & & 0 \\ CK & I_m & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ CA^{k-2}K & CA^{k-3}K & \cdots & I_m \end{bmatrix}. \tag{5.23}$$

The past outputs can be written as:

$$Z_t = \mathcal{O}_p \hat{x}_{t-p} + \mathcal{T}_p E_t \tag{5.24}$$

The covariance of $\mathcal{T}_p E_t$ is denoted by:

$$\Sigma_E \triangleq \mathbb{E}\mathcal{T}_p E_t E_t^* \mathcal{T}_p^* = \mathcal{T}_p \operatorname{diag}(\bar{R}, \dots, \bar{R}) \mathcal{T}_p^*. \tag{5.25}$$

We define the covariance of the state predictions:

$$\Gamma_k \triangleq \mathbb{E}\hat{x}_k \hat{x}_k^* \tag{5.26}$$

and the covariance of the past outputs:

$$\Gamma_{Z,k} \triangleq \mathbb{E}Z_k Z_k^* = \mathcal{O}_p \Gamma_{k-p} \mathcal{O}_p^* + \Sigma_E. \tag{5.27}$$

5.7.2 System Theoretic Bounds

Bounds for system and covariance matrices

Lemma 5.5. *Consider matrix A with Jordan form SJS^{-1}. Let κ be the largest Jordan block of A. Then:*

$$\left\|A^i\right\|_2 \le M^2 \rho(A)^{i-\kappa+1} \left(\frac{ei}{\kappa - 1}\right)^{\kappa-1}. \tag{5.28}$$

As a result the following bounds hold:

$$\begin{aligned} &\left\|A^t\right\|_2 \le M^2 O\left(t^{\kappa-1}\right), && \|\mathcal{O}_t\|_2 \le M^3 O\left(t^{\kappa}\right), \\ &\|\mathcal{T}_t\|_2 \le M^3 O\left(t^{\kappa}\right), && \|\Gamma_t\|_2 \le M^7 O\left(t^{2\kappa-1}\right). \end{aligned}$$

Proof. Recall the Jordan form of $A = SJS^{-1}$. For simplicity, assume that J is equal to a

$n \times n$ Jordan block with eigenvalue $\lambda = \rho(A)$. Then we have that:

$$J^i = \begin{bmatrix} \lambda^i & \binom{i}{1}\lambda^{i-1} & \dots & \binom{i}{n-2}\lambda^{i-n+2} & \binom{i}{n-1}\lambda^{i-n+1} \\ 0 & \lambda^i & \dots & \binom{i}{n-3}\lambda^{i-n+3} & \binom{i}{n-2}\lambda^{i-n+2} \\ & & \ddots & & \\ 0 & 0 & \dots & \lambda^i & \binom{i}{1}\lambda^{i-1} \\ 0 & 0 & \dots & 0 & \lambda^i \end{bmatrix}$$

By Lemma 3.7 and Assumption 5.2, we obtain:

$$\left\|A^i\right\|_2 \leq M^2 \lambda^{i-n+1} \sum_{k=0}^{n-1} \binom{i}{k} \leq M^2 \lambda^{i-n+1} \left(\frac{ei}{n-1}\right)^{n-1}$$

where the second inequality is standard (Vershynin, 2018, Exercise 0.0.5). This completes the proof of (5.28). The proof for the other cases is similar since the Jordan matrix is block diagonal.

The bounds on $\mathcal{O}_t, \mathcal{T}_t, \Gamma_t$ follow from (5.28). □

Lemma 5.6 (Monotonicity Tsiamis & Pappas (2019))**.** *Consider system* (5.2)*, with* $\Gamma_k \triangleq \mathbb{E}\hat{x}_k\hat{x}_k^*$*. The sequence* Γ_k *is increasing in the positive semi-definite cone.*

Lemma 5.7 (Mixing time)**.** *Consider system* (5.2)*, with* $\Gamma_k \triangleq \mathbb{E}\hat{x}_k\hat{x}_k^*$*. Assume that the system is stable with* $\rho(A) < 1$*. Then, the sequence* Γ_k *converges to* $\Gamma_\infty \succ 0$*, the unique positive definite solution of the Lyapunov equation:*

$$\Gamma_\infty = A\Gamma_\infty A^* + K\bar{R}K^*.$$

Let the mixing time τ_{mix} *be defined as in* (5.11)*. Then:*

$$\tau_{\text{mix}} \leq \frac{1}{\log 1/\rho(A)} \tilde{O}(\max\{\log \operatorname{cond}(\Gamma_\infty), \kappa_{\max}\}),$$

with $\operatorname{cond}(\Gamma_\infty) = \frac{\sigma_{\max}(\Gamma_\infty)}{\sigma_{\min}(\Gamma_\infty)}$.

Proof. Since A is stable, $\Gamma_\infty = \sum_{k=0}^{\infty} A^k K \bar{R} K^* (A^*)^k$ is well defined and solves the Lyapunov equation. Since (A, K) is controllable Γ_∞ is strictly positive definite: $\Gamma_\infty \succeq \mathcal{K}_n \mathcal{K}_n^* \succ 0$, where the following controllability matrix has full rank

$$\mathcal{K}_n \triangleq \left[\begin{array}{ccc} KR^{1/2} & AKR^{1/2} & \dots A^{n-1}KR^{1/2} \end{array} \right].$$

It is unique since the operator $\mathcal{L}(M) = M - AMA^*$ is invertible; its eigenvalues are bounded below by $1 - \rho^2(A)$.

Notice that $\Gamma_0 = 0 \preceq \Gamma_\infty$ and by induction, we can show that $\Gamma_k \preceq \Gamma_\infty$. Since Γ_k is also monotone, it converges to the unique Γ_∞. Now after some algebra we obtain

$$\Gamma_k - \Gamma_\infty = -A^k \Gamma_\infty (A^*)^k.$$

It is sufficient to find a τ_{mix} such that:

$$\|A^{\tau_{\text{mix}}}\|^2 \, \sigma_{\max}(\Gamma_\infty) \leq \frac{\sigma_{\min}(\Gamma_\infty)}{2}$$

Since the norm of matrix A grows as fast as $\|A^{\tau_{\text{mix}}}\|_2 = O(\rho(A)^{\tau_{\text{mix}} - \kappa + 1} \tau^\kappa)$, it is sufficient to pick:

$$\tau_{\text{mix}} \geq \frac{\kappa \log \tau_{\text{mix}}}{\log \frac{1}{\rho(A)}} + \frac{\log \operatorname{cond}(\Gamma_\infty)/2}{\log \frac{1}{\rho(A)}} + \kappa - 1.$$

By Lemma 5.17, the order of τ_{mix} is

$$\tau_{\text{mix}} = \frac{1}{\log 1/\rho(A)} \tilde{O}(\max\{\log \operatorname{cond}(\Gamma_\infty), \kappa\}) \qquad \square$$

Proof of Lemma 5.2

For the plain observations y_k:

$$y_{k-t} = CA^{d-t}\hat{x}_{k-d} + \sum_{s=t+1}^{d} CA^{s-t-1}Ke_{k-s} + e_{k-t},$$

for $t = 0, \ldots, d$. By the definition of the minimal polynomial:

$$CA^d \hat{x}_k = a_{d-1} CA^{d-1} \hat{x}_{k-1} + \cdots + a_0 C \hat{x}_{k-d}$$

which leads to:

$$y_k = a_{d-1} y_{k-1} + \cdots + a_0 y_{k-d} + \sum_{s=0}^{d} L_s e_{k-s},$$

with $L_0 = I$ and

$$L_s = -a_{d-s} I_m - \sum_{t=1}^{s-1} a_{d-s+t} CA^{t-1} K + CA^{s-1} K$$

The norm of the above matrices is upper bounded by

$$\left\| L_s \right\|_2 \leq \left\| a \right\|_1 \left\| C \right\|_2 \left\| K \right\|_2 \max_{0 \leq i \leq d} \left\| A^{i-1} \right\|_2, \tag{5.29}$$

where $\left\| a \right\|_1$ denotes the ℓ_1 norm of the polynomial coefficients. The same will now hold for the past outputs:

$$Z_k = a_{d-1} Z_{k-1} + \cdots + a_0 Z_{k-p} + \sum_{s=0}^{d} \operatorname{diag}(L_s, \ldots, L_s) E_{k-s}$$

where E_k is the vector of past noises. We can bound the residual

$$\delta_k \triangleq \sum_{s=0}^{d} \operatorname{diag}(L_s, \ldots, L_s) E_{k-s}$$

by:

$$\left\| \delta_k \right\|_2 \leq \Delta \left\| e_s \right\|_2, \text{ where}$$

$$\Delta \triangleq (d+1) \max_{0 \leq s \leq d} \left\| L_s \right\|_2 \sqrt{p}. \tag{5.30}$$

From (5.29), (5.28) it follows that $\Delta = O(d^{\kappa-1} \left\| a \right\|_1 \sqrt{p})$. □

5.7.3 Statistical Toolbox

Least singular value of Toeplitz matrix

Let $u_t \in \mathbb{R}^m$, $t = 0, \dots$ be an i.i.d. sequence, where $u_k \sim \mathcal{N}(0, I)$ are isotropic Gaussians. The following result, which is adapted from Lee (2020), shows that the Toeplitz matrix of u_t is well conditioned with high probability. Similar results appeared in Sarkar et al. (2019); Oymak & Ozay (2018).

Lemma 5.8 (Toeplitz Isometry Lee (2020))**.** *Let $u_t \in \mathbb{R}^m$, $t = 0, \dots,$ be an i.i.d. sequence of Gaussian variables with unit covariance matrix. Consider the Toeplitz matrix*

$$U = \begin{bmatrix} u_{k-p} & u_{k-p-1} & \dots & u_0 \\ u_{k-p+1} & u_{k-p} & \dots & u_1 \\ \vdots & & & \\ u_{k-1} & u_{k-2} & \dots & u_{p-1} \end{bmatrix}.$$

Fix a failure probability $0 < \delta < 1/2$. There exists a universal constant C such that if

$$k \geq Cpm \log(pm/\delta),$$

then with probability at least $1 - \delta$:

$$\tfrac{1}{2}kI \preceq UU^* \preceq \tfrac{3}{2}kI. \tag{5.31}$$

Proof. Define $\bar{U} = \begin{bmatrix} U & U_0 \end{bmatrix}$, where

$$U_0 = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ u_0 & 0 & \dots & 0 & 0 \\ \vdots & & & & \\ u_{p-2} & u_{p-3} & \dots & u_0 & 0 \end{bmatrix}.$$

Notice that we have:

$$\|UU^* - kI\|_2 \leq \|\bar{U}\bar{U}^* - kI\|_2 + \|U_0U_0^*\|_2$$

By (Lee, 2020, Theorem A.2) it follows that with prob. at least $1-\delta/2$:

$$\left\|\bar{U}\bar{U}^* - kI\right\|_2 \leq C'(pm\log(k/\delta) + \sqrt{kpm\log(k/\delta)}),$$

for some universal constant C'. To bound $U_0U_0^*$, we follow the same steps as in the proof of (Lee, 2020, Theorem A.2) (see bounds on $U_2U_2^*$ there). With probability at least $1-\delta/2$:

$$\|U_0U_0^*\|_2 \leq C''pm\log(pm/\delta),$$

for some other universal constant C''. By a union bound and if we select

$$k \geq Cpm\log(pm/\delta),$$

with C large enough, we get that with probability at least $1-\delta$:

$$\|UU^* - kI\|_2 \leq k/2.$$

The condition on k makes use of Lemma 5.17. □

Gaussian suprema

Lemma 5.9. *Consider $v_t \in \mathbb{R}^d \sim \mathcal{N}(0, I)$ i.i.d., for $t = 1, \dots, k$. Let $X_k \in \mathbb{R}^q$ be a linear combination:*

$$X_k \triangleq \sum_{t=1}^{k} M_{k,t}v_t, \text{ for } k = 1, \dots, T$$

where $M_{t,k} \in \mathbb{R}^{q\times d}$. For some $\mu > 0$ define:

$$\Sigma_k \triangleq \mu I + \mathbb{E}X_kX_k^*$$

Fix a failure probability $\delta > 0$. With probability at least $1-\delta$:

$$\sup_{k=1,\ldots,T} \|\Sigma_k^{-1/2} X_k\|_2 \le \sqrt{q} + \sqrt{2\log\frac{T}{\delta}} \tag{5.32}$$

If $\mathbb{E}X_kX_k^*$, $k \le T$ are invertible, the result holds for $\mu = 0$.

Proof. Fix a k. An application of Jensen's inequality gives:

$$\mathbb{E}\|\Sigma_k^{-1/2} X_k\|_2 \le \sqrt{\mathbb{E}X_k^*\Sigma_k^{-1}X_k} = \sqrt{\operatorname{tr}\Sigma_k^{-1}\mathbb{E}X_kX_k^*} \le \sqrt{q}$$

Observe that we have $\mathbb{E}X_kX_k^* = \sum_{t=1}^k M_tM_t^*$. As a result,

$$\left\|\Sigma_k^{-1/2}\begin{bmatrix} M_1 & \cdots & M_k \end{bmatrix}\right\|_2 \le 1$$

since by definition $\Sigma_k \succeq \mathbb{E}X_kX_k^*$. Hence, the function $\|\Sigma_k^{-1/2}X_k\|_2$ is $1-$Lipschitz with respect to $v_{t,i}$, for $t = 1,\ldots,k$, $i = 1,\ldots,d$. By concentration of Lipschitz functions of independent Gaussian variables (Boucheron et al., 2013, Theorem 5.6):

$$P(\|\Sigma_k^{-1/2}X_k\|_2 > \sqrt{q} + t) \le e^{-t^2/2}$$

Now select $t = \sqrt{2\log\frac{T}{\delta}}$ and take a union bound over k. □

5.7.4 Finite sample bounds for fixed-time, fixed-past

In this subsection, we include results for persistence of excitation and for identification of the system parameters for a fixed time instance k and a fixed past horizon p.

Theorem 5.4 (Finite-sample bounds for identification)**.** *Consider system* (5.2) *with observations $y_0,\ldots,y_k$. Fix a past horizon p and recall the notation of Table 5.1 and the* *universal constant C in Lemma 5.8. Define*

$$k_1(p,\delta) \triangleq Cpm\log(pm/\delta) \tag{5.33}$$
$$k_2(k,p,\delta) \triangleq \frac{512pn}{\min\{4,\sigma_R\}}\log\left(\frac{5p}{\delta}\frac{n\,\|\mathcal{O}_p\|_2^2\,\|\Gamma_{k-p}\|_2+\delta}{\delta}\right)$$

With probability at least $1-5\delta$ the following events hold:

a) ***Persistency of excitation***

$$\mathcal{E}_{PE} \triangleq \begin{Bmatrix} \mathcal{T}_p\bar{E}_k\bar{E}_k^*\mathcal{T}_p^* \succeq \frac{k}{2}\Sigma_E \succeq \frac{k}{2}\sigma_R I, \\ \bar{Z}_k\bar{Z}_k^* \succeq \frac{1}{2}\mathcal{O}_p\bar{X}_k\bar{X}_k^*\mathcal{O}_p^* + \frac{1}{2}\mathcal{T}_p\bar{E}_k\bar{E}_k^*\mathcal{T}_p^*, \end{Bmatrix} \tag{5.34}$$

if k satisfies the following perstistency of excitation requirement

$$k \geq k_3(k,p,\delta) \triangleq \max\{k_1(p,\delta), k_2(k,p,\delta)\}. \tag{5.35}$$

b) ***Bounded outputs***

$$\mathcal{E}_{\bar{Z}} \triangleq \left\{\bar{Z}_k\bar{Z}_k^* \preceq k\frac{mp}{\delta}\Gamma_{Z,k}\right\} \tag{5.36}$$

c) ***Bounded correlations***

$$\mathcal{E}_{\text{cross}} \triangleq \left\{\|S_k\bar{V}_k^{-1/2}\|_2^2 \leq \varepsilon\,(k,p,\delta)\right\},\ \textit{where} \tag{5.37}$$
$$\varepsilon\,(k,p,\delta) \triangleq 16\|\bar{R}\|mp\log\frac{5mpk\left(\|\Gamma_{Z,k}\|_2\lambda^{-1}+1\right)}{\delta} \tag{5.38}$$

Proof. Let $\bar{W}_t, \bar{V}_t$ be the Gram matrices:

$$\bar{W}_k \triangleq \bar{X}_k\bar{X}_k^* + W, \qquad W \triangleq \frac{k}{\|\mathcal{O}_p\|_2^2}I \tag{5.39}$$
$$\bar{V}_k \triangleq \bar{Z}_k\bar{Z}_k^* + V, \qquad V \triangleq \lambda I. \tag{5.40}$$

Define the base events:

$$\begin{aligned}&\mathcal{E}_{\bar{X}}\triangleq\left\{\bar{X}_k\bar{X}_k^*\preceq k\frac{n}{\delta}\Gamma_{k-p}\right\},\quad \mathcal{E}_E\triangleq\left\{\mathcal{T}_p\bar{E}_k\bar{E}_k^*\mathcal{T}_p^*\succeq\frac{k}{2}\Sigma_E\right\}\\ &\mathcal{E}_{XE}\triangleq\left\{\left\|\bar{W}_k^{-1/2}\bar{X}_k\bar{E}_k^*\mathcal{T}_p^*\Sigma_E^{-1/2}\right\|_2^2\leq 8p\log\frac{p5^m\det(\bar{W}_kW^{-1})}{\delta}\right\},\\ &\mathcal{E}_{EZ}\triangleq\left\{\left\|\bar{R}^{-1/2}S_k\bar{V}_k^{-1/2}\right\|_2^2\leq 8\log\frac{5^m\det(\bar{V}_kV^{-1})}{\delta}\right\}\end{aligned}$$

We will show that $\mathcal{E}_{\bar{Z}}$ and all of the base events occur with probability at least $1-\delta$ each. Moreover

$$\mathcal{E}_{PE}\cap\mathcal{E}_{\text{cross}}\supseteq\mathcal{E}_{\bar{X}}\cap\mathcal{E}_{\bar{Z}}\cap\mathcal{E}_E\cap\mathcal{E}_{XE}\cap\mathcal{E}_{ZE}.$$

Hence, by a union bound: $\mathbb{P}(\mathcal{E}_{PE}\cap\mathcal{E}_{\bar{Z}}\cap\mathcal{E}_{\text{cross}})\geq 1-5\delta$.

a) Base events: The fact that $\mathbb{P}\left(\mathcal{E}_{\bar{X}}\right)\geq 1-\delta$, $\mathbb{P}\left(\mathcal{E}_{\bar{Z}}\right)\geq 1-\delta$ follows by a Markov inequality argument–see (Simchowitz et al., 2018, Section 3). The fact that $\mathbb{P}(\mathcal{E}_E)\geq 1-\delta$ follows from Lemma 5.10. For each of the remaining events, we apply Theorem 3.3; note that $\bar{R}^{-1/2}e_k$ and $\Sigma_E^{-1/2}\mathcal{T}_pE_k$ are isotropic.

b) Event $\mathcal{E}_{PE}$: From Lemma 5.11 below, we have that $\mathcal{E}_{PE}\supseteq\mathcal{E}_{\bar{X}}\cap\mathcal{E}_E\cap\mathcal{E}_{XE}$ if k satisfies (5.35).

c) Event $\mathcal{E}_{\text{cross}}$: We show that $\mathcal{E}_{\text{cross}}\supseteq\mathcal{E}_{\bar{\mathrm{Z}}}\cap\mathcal{E}_{\mathrm{ZE}}$. Conditioned on $\mathcal{E}_{\bar{\mathrm{Z}}}\cap\mathcal{E}_{\mathrm{ZE}}$, we have

$$\begin{aligned}\|S_kV_k^{-1/2}\|_2^2&\leq 8\left\|\bar{R}\right\|_2\left(\log\frac{5^m}{\delta}+\log\det\bar{V}_kV^{-1}\right)\\ &\leq 8\left\|\bar{R}\right\|_2\left(\log\frac{5^m}{\delta}+mp\log\left(k\frac{mp}{\delta}\|\Gamma_{Z,k}\|_2\lambda^{-1}+1\right)\right),\end{aligned}$$

where the second inequality follows from $|\det L|\leq\|L\|_2^{mp}$ for any matrix $L\in\mathbb{R}^{mp\times mp}$. The final bound is simplified using $\frac{\delta}{mpk}<1$ and $\log\frac{5^m}{\delta}\leq mp\log\frac{5mp}{\delta}$. □

Next, we prove the persistency of excitation results that are required in the proof of the above theorem.

Lemma 5.10 (Noise PE)**.** *Consider the conditions of Theorem 5.4 and the definition of*

$k_1(p,\delta)$. If

$$k \geq k_1(p,\delta) \tag{5.41}$$

then with probability at least $1-\delta$

$$\frac{k}{2}\sigma_R I \preceq \frac{k}{2}\Sigma_E \preceq \mathcal{T}_p \bar{E}_k \bar{E}_k^* \mathcal{T}_p^* \preceq \frac{3k}{2}\Sigma_E.$$

Proof. Matrices $U_k \triangleq \Sigma_E^{-1/2}\mathcal{T}_p E_k$ satisfy the conditions of Lemma 5.8. Hence, under (5.41), with probability at least $1-\delta$:

$$\frac{k}{2} I \preceq \sum_{t=p}^{k} U_t U_t^* \preceq \frac{3k}{2} I.$$

Multiplying from both sides with $\Sigma_E^{1/2}$ gives

$$\frac{k}{2}\Sigma_E \preceq \mathcal{T}_p \bar{E}_k \bar{E}_k^* \mathcal{T}_p^* \preceq \frac{3k}{2}\Sigma_E$$

Finally, from (Tsiamis & Pappas, 2019, Lemma 2), we have $\Sigma_E \succeq \sigma_R I$. □

Next, we prove persistency of excitation for the past outputs.

Lemma 5.11 (Output PE)**.** *Consider the conditions of Theorem 5.4 and the definition of $k_2(k,p,\delta)$, $\mathcal{E}_E$, $\mathcal{E}_{\bar{X}}$, $\mathcal{E}_{XE}$. If:*

$$k \geq k_2(k,p,\delta) \tag{5.42}$$

then the following output PE condition holds:

$$\left\{ \bar{Z}_k \bar{Z}_k^* \succeq \frac{1}{2}\mathcal{O}_p \bar{X}_k \bar{X}_k^* \mathcal{O}_p + \frac{1}{2}\mathcal{T}_p \bar{E}_k \bar{E}_k \mathcal{T}_p \right\} \supseteq \mathcal{E}_E \cap \mathcal{E}_{\bar{X}} \cap \mathcal{E}_{XE}$$

Proof. The batch past outputs can be written as:

$$\bar{Z}_k = \mathcal{O}_p \bar{X}_k + \mathcal{T}_p \bar{E}_k.$$

As a result, the sample-covariance matrix $\bar{Z}_k \bar{Z}_k^*$ will be:

$$O_p \bar{X}_k \bar{X}_k^* \mathcal{O}_p^* + \mathcal{T}_p \bar{E}_k \bar{E}_k^* \mathcal{T}_p^* + \mathcal{O}_p \bar{X}_k \bar{E}_k^* \mathcal{T}_p^* + \mathcal{T}_p \bar{E}_k \bar{X}_k^* \mathcal{O}_p^*$$

The proof proceeds in two steps. First, we bound the cross-terms based on the events $\mathcal{E}_{\bar{X}}, \mathcal{E}_{XE}$. Second, we show that if k is large enough, then the cross-terms are dominated by the auto-correlation terms:

$$\mathcal{O}_p \bar{X}_k \bar{E}_k^* \mathcal{T}_p^* + \mathcal{T}_p \bar{E}_k \bar{X}_k^* \mathcal{O}_p^* \preceq \frac{1}{2}\left(O_p \bar{X}_k \bar{X}_k^* \mathcal{O}_p^* + \mathcal{T}_p \bar{E}_k \bar{E}_k^* \mathcal{T}_p^*\right).$$

Cross-term bounds: For simplicity, we rewrite $\Sigma_E^{-1/2} \mathcal{T}_p \bar{E}_k = \bar{U}_k$, where $\bar{U}_k$ is defined similarly to $\bar{E}_k$ but has unit variance components. Conditioned on $\mathcal{E}_{\bar{X}}$

$$\log\det \bar{W}_k W^{-1} \leq n \log\left(\frac{n \left\|\mathcal{O}_p\right\|_2^2 \left\|\Gamma_{k-p}\right\|_2}{\delta} + 1\right),$$

where we used the property $|\det L| \leq \|L\|_2^n$, for any matrix $L \in \mathbb{R}^n$. Conditioned also on $\mathcal{E}_{XE}$:

$$\left\|\bar{W}_k^{-1/2} \bar{X}_k \bar{U}_k^*\right\|_2^2 \leq \mathcal{C}_{XE}^2$$

where we define

$$\mathcal{C}_{XE} \triangleq \sqrt{8p\left(\log\frac{p5^m}{\delta} + n\log\left(\frac{n\left\|\mathcal{O}_p\right\|_2^2 \left\|\Gamma_{k-p}\right\|_2}{\delta} + 1\right)\right)}.$$

Let now $u \in \mathbb{R}^{mp}$, $\|u\|_2 = 1$ be an arbitrary unit vector. Then, consider the quadratic form

$$q(u) \triangleq \frac{1}{k}\left(u^* \mathcal{O}_p \bar{X}_k \bar{U}_k^* \Sigma_E^{1/2} u + u^* \Sigma_E^{1/2} \bar{U}_k \bar{X}_k^* \mathcal{O}_p^* u\right).$$

Conditioned on $\left\{\left\|\bar{W}_k^{-1/2} \bar{X}_k \bar{U}_k^*\right\|_2^2 \leq \mathcal{C}_{XE}\right\} \cap \mathcal{E}_E \cap \mathcal{E}_X$ and using $I = \bar{W}_k \bar{W}_k^{-1}$ we can bound the quadratic form by:

$$\begin{aligned}|q(u)| &\leq \frac{2}{k}\left\|u^*\mathcal{O}_p\bar{W}_k^{1/2}\right\|_2\left\|\bar{W}_k^{-1/2}\bar{X}_k\bar{U}_k^*\right\|_2\left\|\Sigma_E^{1/2}u\right\|_2 \\ &\leq 2\sqrt{\frac{1}{k}u^*\mathcal{O}_p\bar{X}_k\bar{X}_k^*\mathcal{O}_p^*u+1}\frac{\mathcal{C}_{XE}}{\sqrt{k}}\left\|\Sigma_E^{1/2}u\right\|_2\end{aligned}$$

Cross-terms are dominated: Define variables:

$$a \triangleq \frac{1}{k}u^*\mathcal{O}_p\bar{X}_k\bar{X}_k^*\mathcal{O}_p^*u, \quad b \triangleq u^*\Sigma_E u$$

To complete the proof, it is sufficient to show that for any unit vector u if (5.42) holds then:

$$2\sqrt{a+1}\frac{\mathcal{C}_{XE}}{\sqrt{k}}\sqrt{b} \leq \frac{1}{2}\left(a+\frac{1}{k}u^*\Sigma_E^{1/2}\bar{U}_k\bar{U}_k^*\Sigma_E^{1/2}u\right).$$

But on $\mathcal{E}_E$ we have $\Sigma_E^{1/2}\bar{U}_k\bar{U}_k^*\Sigma_E^{1/2} \succeq k/2\Sigma_E$. Thus, it is sufficient to show

$$2\sqrt{a+1}\frac{\mathcal{C}_{XE}}{\sqrt{k}}\sqrt{b} \leq \frac{a}{2}+\frac{b}{4}.$$

To guarantee the inequality, we apply the following Lemma 5.12, where we exploit the fact that $b \geq \sigma_{\min}(\Sigma_E) \geq \sigma_R$. It follows that it is sufficient to have:

$$\mathcal{C}_{XE}/\sqrt{k} \leq \min\left\{2, \sqrt{\sigma_R}\right\}/8 \qquad \square$$

To obtain the final expression for $k_2(k,p,\delta)$ we use the simplification $8p\log(p5^m/\delta) \leq 8pn\log(5p/\delta)$, since $m \leq n$.

Lemma 5.12. *Let $a \geq 0$ and $b \geq \sigma_R > 0$. Then if*

$$\gamma \leq \frac{\min\left\{2, \sqrt{\sigma_R}\right\}}{8}$$

then $f(a,b) \triangleq \frac{a}{2}+\frac{b}{4}-2\sqrt{a+1}\sqrt{b}\gamma \geq 0$.

Proof. By minimizing over a, we obtain:

$$\min_{0\le a} f(a,b) = \left.\begin{matrix} b/4 - 2\sqrt{b}\gamma, & \text{if } 2\gamma\sqrt{b} \le 1 \\ b\left(1/4 - 2\gamma^2\right) - 1/2, & \text{if } 2\gamma\sqrt{b} > 1 \end{matrix}\right\}.$$

If $2\gamma\sqrt{b} \le 1$ then we have

$$\min_{0\le a,\sigma_R\le b} f(a,b) = \frac{b}{4} - 2\sqrt{b}\gamma \ge \frac{b-\sqrt{b\sigma_R}}{4} \ge 0.$$

Since $\gamma \le 1/4$, the case $2\gamma\sqrt{b} > 1$ can occur only if $b > 4$. But then, for $b > 4$

$$b\left(\frac{1}{4} - 2\gamma^2\right) - \frac{1}{2} \ge b\left(\frac{1}{4} - \frac{1}{8}\right) - \frac{1}{2} = \frac{b-4}{8} \ge 0$$

□

5.7.5 Proof of Lemma 5.3

We prove a more general result.

Lemma 5.13 (Uniform bounds)**.** *Consider the conditions of Theorem 5.1. Select a failure probability $\delta > 0$. Let $T = 2^{i-1}T_{init}$ for some fixed epoch i with $p = \beta \log T$ the corresponding past horizon. Consider the definition of $\varepsilon(k,p,\delta)$ in (5.38). There exists a $N_0 = \mathrm{poly}(n,\beta,\log 1/\delta, 1/\sigma_R)$ such that with probability at least $1 - 5\sum_{k=T}^{2T-1}\frac{1}{k^2}\delta$ the following events hold:*

$$\mathcal{E}_{\text{unif}} \triangleq \bigcup_{k=T}^{2T-1} \left\{ \begin{matrix} \bar{Z}_k\bar{Z}_k^* \preceq \dfrac{k^3 mp}{\delta}\Gamma_{Z,k} \\ \left\| S_k \bar{V}_k^{-1/2} \right\|_2 \le \varepsilon(k,p,\delta/k^2) \end{matrix} \right\} \tag{5.43}$$

$$\mathcal{E}_{\text{unif}}^{\text{PE}} \triangleq \bigcup_{k=\max\{N_0,T\}}^{2T-1} \left\{ \bar{Z}_k\bar{Z}_k^* \succeq \frac{k}{4}\sigma_R I \right\} \tag{5.44}$$

Proof. Recall $k_3(k,p,\delta)$ defined in (5.35) and let:

$$N_0 \triangleq \min\left\{t:\ k \ge k_3(k,p,\delta/k^2),\ \text{for all } k \ge t\right\}. \tag{5.45}$$

Now, fix a k such that $T \leq k \leq 2T-1$ apply Theorem 5.4 for δ replaced with δ/k^2. Taking the union bound over all $T \leq k \leq 2T-1$, from (5.34), (5.36), (5.37) we obtain that:

$$\mathbb{P}(\mathcal{E}_{\text{unif}} \cap \mathcal{E}^{\text{PE}}_{\text{unif}}) \geq 1 - 5\sum_{k=T}^{2T-1} \frac{1}{k^2}\delta$$

What remains to show is that N_0 depends polynomially on $\beta, n, \log 1/\delta, 1/\sigma_R$. Now, by Lemma 5.5, the covariance matrix Γ_k increases at most as fast as $M^4 k^{2\kappa-1}$ (a similar result holds for the observability matrix). Hence, the dominant term in k_3 is of the order of:

$$k_3(k, \beta \log k, \delta/k^2) \leq \beta n \kappa \tfrac{M}{\sigma_R} \log(1/\delta)\, O(\log^2 k).$$

For simplicity define $N_{\text{diff}} = \beta n \kappa \frac{M}{\sigma_R} \log(1/\delta)$. By Lemmas 5.18, it follows that N_0 is at most of the order of:

$$N_0 \leq C' N_{\text{diff}} \log^2 N_{\text{diff}},$$

where C' is some universal constant. □

5.7.6 Proof of Lemma 5.1

We prove a slightly more general version.

Lemma 5.14. *Fix a p and consider the notation of Table 5.1. Consider an $i \geq 0$. The following inequality is true:*

$$\sum_{k=T}^{2T-1} Z^*_{k-i} \bar{V}_k^{-1} Z_{k-i} \leq \log\det(\bar{V}_{2T-i-1}\bar{V}^{-1}_{T-i-1}) \tag{5.46}$$

Proof. Since $\bar{V}_k$ is increasing in the positive semidefinite cone:

$$\sum_{k=T}^{2T-1} Z^*_{k-i} \bar{V}_k^{-1} Z_{k-i} \leq \sum_{k=T}^{2T-1} Z^*_{k-i} \bar{V}_{k-i}^{-1} Z_{k-i}$$

Hence, it is sufficient to prove the inequality for $i = 0$. Recall that $\bar{V}_{k-1} = \bar{V}_k - Z_k Z_k^*$.

Using the identity $\det(I + FB) = \det(I + BF)$ we obtain:

$$\begin{aligned}\det \bar{V}_{k-1} &= \det \bar{V}_k \det \left(I - \bar{V}_k^{-1/2} Z_k Z_k^* \bar{V}_k^{-1/2} \right) \\ &= \det \bar{V}_k \left(1 - Z_k^* \bar{V}_k^{-1} Z_k \right).\end{aligned}$$

Rearranging the terms gives:

$$Z_k^* \bar{V}_k^{-1} Z_k = 1 - \frac{\det \bar{V}_{k-1}}{\det \bar{V}_k} \leq \log\det \bar{V}_k - \log\det \bar{V}_{k-1},$$

where the inequality follows from the fact that the sequence $\bar{V}_k \succeq \bar{V}_{k-1}$ is increasing and the elementary inequality:

$$1 - x \leq \log 1/x, \text{ for } x \leq 1.$$

Since the upper bound telescopes, we finally get

$$\sum_{k=T}^{2T-1} Z_k^* \bar{V}_k^{-1} Z_k \leq \log\det \bar{V}_{2T-1} - \log\det \bar{V}_{T-1} \qquad \square$$

5.7.7 Analysis within one epoch

We will analyze the ℓ_2 square loss for the duration of one epoch, from time T up to time $2T - 1$ with fixed past horizon $p = \beta \log T$. We have two cases: i) persistency of excitation is established $T \geq N_0$, where N_0 is defined in (5.45); ii) persistency of excitation is not established $T < N_0$.

Consider the ℓ_2 loss within the epoch:

$$\mathcal{L}_T^{2T-1} \triangleq \sum_{k=T}^{2T-1} \|\hat{y}_k - \tilde{y}_k\|_2^2. \tag{5.47}$$

Based on the notation of Table 5.1, the error between the Kalman filter prediction and our online algorithm is:

$$\begin{aligned}\tilde{y}_k - \hat{y}_k = S_{k-1}\bar{V}_{k-1}^{-1} Z_k + \lambda G \bar{V}_{k-1}^{-1} Z_k \\ + C(A-KC)^p \left(\bar{X}_{k-1}\bar{Z}_{k-1}\bar{V}_{k-1}^{-1} Z_k - \hat{x}_{k-p}\right).\end{aligned}$$

By Cauchy-Schwarz, the submultiplicative property of norm and by the fact that

$$\|\bar{Z}_{k-1}\bar{V}_{k-1}^{-1/2}\|_2 \leq 1$$

is normalized:

$$\begin{aligned}&\|\tilde{y}_k - \hat{y}_k\|_2^2 \leq 4\Big(\|S_{k-1}\bar{V}_{k-1}^{-1/2}\|_2^2 + \|\lambda G \bar{V}_k^{-1/2}\|_2^2 \\ &+ \|C(A-KC)^p\|_2^2 \, \|\bar{X}_{k-1}\|_2^2\Big) \, \|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2 \\ &+ 4\, \|C(A-KC)^p\|_2^2 \, \|\hat{x}_{k-p}\|_2^2 . \end{aligned} \tag{5.48}$$

To obtain a bound on the square loss, it is sufficient to bound three terms: i) the supremum over $T \leq t \leq 2T-1$ of:

$$\|S_{t-1}\bar{V}_{t-1}^{-1/2}\|_2^2 + \|\lambda G \bar{V}_t^{-1/2}\|_2^2 + \|C(A-KC)^p\|_2^2 \left\|\bar{X}_{t-1}\right\|_2^2$$

ii) the sum $\|C(A-KC)^p\|_2^2 \sum_{k=T}^{2T-1} \|\hat{x}_{k-p}\|_2^2$.
iii) the sum $\sum_{k=T}^{2T-1} \left\|\bar{V}_{k-1}^{-1/2} Z_k\right\|_2^2$,

Theorem 5.5 (Square loss within epoch)**.** *Consider the conditions of Theorem 5.1. Let a be the minimal polynomial of A with degree d, Δ defined as in (5.30). Fix a failure probability $\delta > 0$ and consider N_0 defined as in (5.45). Let $T = 2^{i-1}T_{init}$ for some fixed epoch $i \geq 1$ with $p = \beta \log T$ the corresponding past horizon. Let $\mathcal{C}_{\text{diff}} = d^\kappa n \beta \|a\|_2 \log 1/\delta$. Then, with probability at least $1 - 7\sum_{k=T}^{2T-1} \frac{1}{k^2}\delta$:*

$$\mathcal{L}_T^{2T-1} \leq \text{poly}(\mathcal{C}_{\text{diff}}, \lambda^{-1}) \left(\tilde{O}(T) + \tilde{O}(\rho(A-KC)^p T^{2\kappa+1})\right).$$

If we have persistency of excitation, $T \geq N_0$, then:

$$\mathcal{L}_T^{2T-1} \leq \mathrm{poly}(\mathcal{C}_{\mathrm{diff}}, \sigma_R^{-1}) \left(\tilde{O}(1) + \tilde{O}(\rho(A-KC)^p T^{2\kappa}) \right).$$

Proof. Consider the uniform events $\mathcal{E}_{\mathrm{unif}}$ and $\mathcal{E}_{\mathrm{unif}}^{\mathrm{PE}}$ defined in (5.43), (5.44) and define the events:

$$\mathcal{E}_x \triangleq \left\{ \sup_{k \leq 2T-1} \|\Gamma_k^{-1/2} \hat{x}_k\|_2 \leq \sqrt{n} + \sqrt{2 \log \frac{4T}{\delta_1}} \right\}$$
$$\mathcal{E}_e = \left\{ \sup_{k \leq 2T-1} \|\bar{R}^{-1/2} e_k\|_2 \leq \sqrt{m} + \sqrt{2 \log \frac{2T}{\delta_1}} \right\},$$

where $\delta_1 = \sum_{k=T}^{2T-1} \delta/k^2$. Based on Lemma 5.13, Lemma 5.9, and a union bound the events $\mathcal{E}_{\mathrm{unif}} \cap \mathcal{E}_{\mathrm{unif}}^{\mathrm{PE}} \cap \mathcal{E}_x \cap \mathcal{E}_e$ occur with probability at least $1 - 7\sum_{k=T}^{2T-1} \frac{1}{k^2}\delta$. We will bound all terms of the square loss based on the above events.

i-a) For the term $S_{k-1}\bar{V}_{k-1}^{-1/2}$, based on event $\mathcal{E}_{\mathrm{unif}}$:

$$\|S_{k-1}\bar{V}_{k-1}^{-1/2}\|_2^2 \leq \varepsilon^2(2T, p, \delta/(2T)^2) = \mathrm{poly}(\mathcal{C}_{\mathrm{diff}})\tilde{O}(1)$$

i-b) Regularization term: $\|\lambda G \bar{V}_t^{-1/2}\|_2^2 \leq \lambda \|G\|_2^2$

i-c) To bound the term $\|C(A-KC)^p\|_2^2 \left\|\bar{X}_{t-1}\right\|_2^2$, we use the inequality

$$\left\|\bar{X}_{t-1}\right\|_2^2 \leq 2T \sup_{k \leq 2T-1} \|\hat{x}_k\|_2^2.$$

Hence it is sufficient to upper-bound the norm of the states. Since the covariances Γ_k are increasing:

$$\sup_{k \leq 2T-1} \|\hat{x}_k\|_2^2 \leq \|\Gamma_{2T-1}\|_2 \sup_{k \leq 2T-1} \left\|\Gamma_k^{-1/2}\hat{x}_k\right\|_2^2.$$

Hence, conditioned on $\mathcal{E}_x$ and since $\|\Gamma_{2T-1}\|_2 = O(T^{2\kappa-1})$:

$$\|C(A-KC)^p\|_2^2 \left\|\bar{X}_{t-1}\right\|_2^2 \leq \mathrm{poly}(\mathcal{C}_{\mathrm{diff}})\tilde{O}(\rho(A-KC)^p T^{2\kappa})$$

ii) To bound the sum $\|C(A-KC)^p\|_2^2 \sum_{k=T}^{2T-1} \|\hat{x}_{k-p}\|_2^2$, we use the exact same steps as above since:

$$\sum_{k=T}^{2T-1} \|\hat{x}_{k-p}\|_2^2 \leq T \sup_{k \leq 2T-1} \|\hat{x}_k\|_2^2$$

iii) To bound the sum of $\|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2$, we exploit the ARMA-like representation of the past outputs (Lemma 5.2) along with Lemma 5.14. First, replace $Z_k = a_{d-1}Z_{k-1} + \cdots + a_0 Z_{k-d} + \delta_k$. Then, by two applications of Cauchy-Schwarz:

$$\|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2 \leq 2\,\|a\|_2^2 \sum_{i=0}^{d-1} Z_{k-i}^* \bar{V}_{k-1}^{-1} Z_{k-i} + 2\|\bar{V}_{k-1}^{-1/2}\delta_k\|_2^2.$$

By Lemma 5.14, we can bound the first summand by:

$$\|a\|_2^2 \sum_{k=T}^{2T-1} \sum_{i=0}^{d-1} Z_{k-i}^* \bar{V}_{k-1}^{-1} Z_{k-i} \leq d\,\|a\|_2^2 \log\det(\bar{V}_{2T-1}\lambda^{-1})$$

The second summand can be bounded by:

$$\sum_{k=T}^{2T-1} \|\bar{V}_{k-1}^{-1/2}\delta_k\|_2^2 \leq \Delta \sup_{k \leq 2T} \|e_k\|_2^2 \sum_{k=T}^{2T-1} \|\bar{V}_{k-1}^{-1/2}\|_2^2$$

He have two cases depending on persistency of excitation:

$$\begin{aligned} &\sum_{k=T}^{2T-1} \|\bar{V}_{k-1}^{-1/2}\|_2^2 \leq \frac{4}{\sigma_R} \sum_{T}^{2T-1} \frac{1}{k} \leq \frac{4}{\sigma_R} \log\frac{2T-1}{T-1}, \text{ if } T \geq N_0 \\ &\sum_{k=T}^{2T-1} \|\bar{V}_{k-1}^{-1/2}\|_2^2 \leq \frac{T}{\lambda}, \text{ if } T < N_0. \end{aligned}$$

The terms $\log\det(\bar{V}_{2T}\lambda^{-1})$, $\sup_{k \leq 2T} \|e_k\|_2^2$ can be bounded by $\text{poly}(\mathcal{C}_{\text{diff}})\tilde{O}(1)$ based on $\mathcal{E}_{\text{unif}}, \mathcal{E}_e$. □

5.7.8 Proof of Theorem 5.1

Recall that the regret can be decomposed into two terms:

$$\mathcal{R}_N = \mathcal{L}_N + 2\sum_{k=T_{\text{init}}}^{N} e_k^* \left(\hat{y}_k - \tilde{y}_k\right)$$

where $\mathcal{L}_N$ is the square loss and the other term is a martingale.

Square loss bound. Without loss of generality assume that $N = 2T_i - 1 = T_{\text{init}}2^i$ is the end of an epoch, where i is the total number of epochs. The number of epochs i depends logarithmically on N. Then the square loss $\mathcal{L}_N$ is written as:

$$\mathcal{L}_N = \sum_{j=1}^{i-1} \mathcal{L}_{T_j}^{2T_j-1}.$$

Let N_0 be defined as in (5.45). Select

$$\beta \geq 4\frac{\kappa}{\log 1/\rho(A-KC)}. \tag{5.49}$$

Then by Theorem 5.5 and a union bound over all epochs, with probability at least $1-7\frac{\pi^2}{6}\delta$:

$$\mathcal{L}_N = \text{poly}(\mathcal{C}_{\text{diff}}, \lambda^{-1})\tilde{O}_{N_0}(N_0) + \text{poly}(\mathcal{C}_{\text{diff}}, \sigma_R^{-1})\tilde{O}(1).$$

Martingale term bound. Denote $u_k \triangleq \bar{R}^{-1/2}e_k$ and $z_k \triangleq R^{1/2}\left(\hat{y}_k - \tilde{y}_k\right)$. Then $\sum_{t=1}^{N} e_t^* \left(\hat{y}_t - \tilde{y}_t\right) = \sum_{t=1}^{N} u_t^* z_t = \sum_{t=1}^{N}\sum_{i=1}^{m} u_{t,i}z_{t,i}$. To apply Theorem 3.3 we need to slightly modify the definition of the filtration. Let $\mathcal{F}_{t,i} \triangleq \sigma(\mathcal{F}_t \cup \{u_{t+1,1},\dots,u_{t+1,i}\})$, with $\mathcal{F}_{t+1} \equiv \mathcal{F}_{t,m}$ and define:

$$\tilde{\mathcal{F}}_0 = \mathcal{F}_0 \tag{5.50}$$

$$\tilde{\mathcal{F}}_s = \mathcal{F}_{t,s-tm}, \text{ if } tm+1 \leq s \leq (t+1)m \tag{5.51}$$

By applying Theorem 3.3 with $\tilde{\mathcal{F}}_s$ we can bound the sum in terms of the square loss

$\mathcal{L}_N$. With probability at least $1-\delta$:

$$\left(\sum_{t=1}^{N} z_t^* z_t + 1\right)^{-1/2} \sum_{t=1}^{N} u_t^* z_t \le 8\log\left(\frac{5}{\delta}\left(\sum_{t=1}^{N} z_t^* z_t + 1\right)\right)$$

or using the fact that $z_k^* z_k \le \left\|\bar{R}\right\|_2 \|\hat{y}_k - \tilde{y}_k\|_2^2$:

$$\sum_{t=1}^{N} u_t^* z_t \le \left(\left\|\bar{R}\right\|_2 \mathcal{L}_N + 1\right)^{1/2} 8\log\left(\frac{5}{\delta}\left(\left\|\bar{R}\right\|_2 \mathcal{L}_N + 1\right)\right).$$

By a union bound, with probability at least $1-(7\frac{\pi^2}{6}+1)\delta$:

$$\mathcal{R}_N = \text{poly}(\mathcal{C}_{\text{diff}}, \lambda^{-1})\tilde{O}_{N_0}(N_0) + \text{poly}(\mathcal{C}_{\text{diff}}, \sigma_R^{-1})\tilde{O}(1).$$

To complete the proof, we re-scale δ and we re-write the bounds with respect $\tilde{\delta} = (7\frac{\pi^2}{6} + 1)^{-1}\delta$. □

5.8 Regret analysis for stable systems

In the case of stable systems, stationarity allows us to prove stronger persistency of excitation results.

Lemma 5.15 (Stable: Output PE)**.** *Consider system* (5.2) *with observations* $y_0, \dots, y_k$*. Let* $\tau = \tau_{\text{mix}} + p$*, where* τ_{mix} *is the mixing time defined in* (5.11)*. Recall the universal constant* C *in Lemma* 5.8 *and the notation of Table* 5.1*. Define:*

$$\begin{aligned} k_4(p,\delta) &\triangleq C\tau m \log(\tau m/\delta) && (5.52)\\ k_5(k,p,\delta) &\triangleq \frac{512pn}{\min\{4,\sigma_R\}}\log\left(\frac{5p}{\delta}\frac{n\,\|\mathcal{O}_\tau\|_2^2\,\|\Gamma_{k-\tau}\|_2}{\delta}+1\right). \end{aligned}$$

With probability at least $1-3\delta$*, if*

$$k \ge k_6(k,p,\delta) \triangleq \max\{k_4(p,\delta), k_5(k,p,\delta)\}, \qquad (5.53)$$

then the following output PE condition holds:

$$\bar{Z}_k \bar{Z}_k^* \succeq \frac{k}{4}\Gamma_{Z,\tau} \succeq \frac{k}{8}\Gamma_{Z,\infty}. \tag{5.54}$$

Proof. Define the controllability matrix:

$$\mathcal{C}_t \triangleq \begin{bmatrix} A^{t-1}K & \dots & AK & K \end{bmatrix}, \, t \geq 1.$$

The state covariance at any time can be conveniently expressed in terms of the controllability matrix:

$$\Gamma_t = \mathcal{C}_t \operatorname{diag}(\bar{R}, \dots, \bar{R})\mathcal{C}_t^*.$$

Combining the above equality with (5.27), we obtain that

$$\Gamma_{Z,\tau}{=}\begin{bmatrix} O_p\mathcal{C}_{\tau_{\text{mix}}} & \mathcal{T}_p \end{bmatrix} \operatorname{diag}(\bar{R}, \dots, \bar{R}) \begin{bmatrix} O_p\mathcal{C}_{\tau_{\text{mix}}} & \mathcal{T}_p \end{bmatrix}^*.$$

Hence, by the definition of mixing time:

$$\Gamma_{Z,\tau} \succeq \Gamma_{Z,\infty}/2. \tag{5.55}$$

What remains is to show the first inequality in (5.54).

The proof is similar to the one of Lemma 5.11. We only need an additional step, to further unroll $\hat{x}_{x-p}$ for τ_{mix} time-steps into the past in (5.24). Extend the definition of the past noises:

$$E_t^{\tau} \triangleq \begin{bmatrix} e_{t-\tau}^* & \cdots & e_{t-1}^* \end{bmatrix}^*.$$

Rolling out the state equations in (5.24), we obtain:

$$Z_t = \mathcal{O}_p A^{\tau_{\text{mix}}}\hat{x}_{t-\tau} + \begin{bmatrix} O_p\mathcal{C}_{\tau_{\text{mix}}} & \mathcal{T}_p \end{bmatrix} E_t^{\tau}. \tag{5.56}$$

Define the isotropic variables:

$$U_t \triangleq \Gamma_{Z,\tau}^{-1/2} \left[\begin{array}{cc} O_p \mathcal{C}_{\tau_{\mathrm{mix}}} & \mathcal{T}_p \end{array} \right] E_t^{\tau},$$

which are well-defined since $\Gamma_{Z,\tau} \succeq \Sigma_E \succeq \sigma_R I$. This enables us to rewrite (5.56) as

$$Z_t = \mathcal{O}_p A^{\tau_{\mathrm{mix}}} \hat{x}_{t-\tau} + \Gamma_{Z,\tau}^{1/2} U_t. \tag{5.57}$$

Expanding the correlations gives:

$$\begin{aligned} Z_t Z_t^* &= \mathcal{O}_p A^{\tau_{\mathrm{mix}}} \hat{x}_{t-\tau} \hat{x}_{t-\tau}^* (\mathcal{O}_p A^{\tau_{\mathrm{mix}}})^* + \Gamma_{Z,\tau}^{1/2} U_t U_t^* \Gamma_{Z,\tau}^{1/2} \\ &+ \mathcal{O}_p A^{\tau_{\mathrm{mix}}} \hat{x}_{t-\tau} U_t^* \Gamma_{Z,\tau}^{1/2} + \Gamma_{Z,\tau}^{1/2} U_t \hat{x}_{t-\tau}^* (\mathcal{O}_p A^{\tau_{\mathrm{mix}}})^* \end{aligned}$$

The remaining proof is now identical to Lemma 5.11 and is, thus, omitted. To simplify the notation for the expression of k_5 we used $\|\mathcal{O}_p A^{\tau_{\mathrm{mix}}}\|_2 \leq \|\mathcal{O}_\tau\|_2$ ($\mathcal{O}_p A^{\tau_{\mathrm{mix}}}$ is a sub-matrix of $\mathcal{O}_\tau$). We also used $\sigma_{\min}(\Gamma_{Z,\infty}) \geq \sigma_{\min}(\Sigma_E) \geq \sigma_R$ in the step where we apply the technical Lemma 5.12. □

5.8.1 Proof of Lemma 5.4

We prove a more general version.

Lemma 5.16 (Stable case: Uniform PAC bounds)**.** *Consider the conditions of Theorem 5.1 with $\rho(A) < 1$. Select a failure probability $\delta > 0$. Let $T = 2^{i-1} T_{init}$ for some fixed epoch i with $p = \beta \log T$ the corresponding past horizon. Consider also the definition of $\varepsilon(k, p, \delta)$ in (5.38). There exists a $N_0^{\mathrm{s}} = \mathrm{poly}(n, \beta, \log 1/\delta, \tau_{\mathrm{mix}}, 1/\sigma_R)$ such that with probability at* *least* $1-5\sum_{k=T}^{2T-1}\frac{1}{k^2}\delta$ *the following events hold:*

$$\mathcal{E}_{\mathrm{unif}} \triangleq \bigcup_{k=T}^{2T-1}\left\{\begin{array}{c} \bar{Z}_k\bar{Z}_k^* \preceq \dfrac{k^3mp}{\delta}\Gamma_{Z,k} \\ \left\|S_k\bar{V}_k^{-1/2}\right\|_2 \leq \varepsilon(k,p,\delta/k^2) \end{array}\right\} \tag{5.58}$$

$$\mathcal{E}_{\mathrm{unif}}^{\mathrm{PE,s}} \triangleq \bigcup_{k=\max\left\{N_0^{\mathrm{s}},T\right\}}^{2T-1}\left\{\bar{Z}_k\bar{Z}_k^* \succeq \frac{k}{8}\Gamma_{Z,\infty}\right\}. \tag{5.59}$$

Proof. The proof is identical to Lemma 5.13. The only difference is the definition of N_0^{s}:

$$N_0^{\mathrm{s}} \triangleq \min\left\{t:\ k \geq k_6(k,p,\delta/k^2),\ \text{for all } k \geq t\right\}, \tag{5.60}$$

where k_6 is defined in (5.53). As a result N_0^{s} is of the order of:

$$N_0^{\mathrm{s}} \leq C' N_{\mathrm{diff}}^{\mathrm{s}} \log^2 N_{\mathrm{diff}}^{\mathrm{s}},$$

where $N_{\mathrm{diff}}^{\mathrm{s}} = \tau_{\mathrm{mix}}\beta n\kappa\frac{M}{\sigma_R}\log(1/\delta)$, C' is some constant. □

5.8.2 Proof of Theorem 5.2

Similar to the non-explosive case, we analyze the square loss for a single epoch. We additionally exploit the sharper persistency of excitation condition.

Theorem 5.6 (Square loss within epoch)**.** *Consider the conditions of Theorem 5.2. Fix a failure probabilities $\delta > 0$ and consider N_0^{s} defined as in (5.60). Let $T = 2^{i-1}T_{init}$ for some fixed epoch $i \geq 1$ with $p = \beta\log T$ the corresponding past horizon. Let β satisfy (5.9). Let $\mathcal{C}_{\mathrm{diff}}^{\mathrm{s}} = n\beta\log 1/\delta$. Then, with probability at least $1-7\sum_{k=T}^{2T-1}\frac{1}{k^2}\delta$:*

$$\mathcal{L}_T^{2T-1} \leq \|\Gamma_\infty\|_2 \mathrm{poly}(\mathcal{C}_{\mathrm{diff}}^{\mathrm{s}},\lambda^{-1})\tilde{O}(T).$$

If moreover $T \geq N_0^{\mathrm{s}}$ then also:

$$\mathcal{L}_T^{2T-1} \leq \mathrm{poly}(\mathcal{C}_{\mathrm{diff}}^{s})\tilde{O}(1).$$

Proof. Consider the uniform events $\mathcal{E}_{\text{unif}}$ and $\mathcal{E}_{\text{unif}}^{\text{PE,s}}$ defined in (5.58), (5.59). Define also

$$\mathcal{E}_x = \left\{ \sup_{k \leq 2T-1} \|\Gamma_k^{-1/2} \hat{x}_k\|_2 \leq \sqrt{n} + \sqrt{2 \log \frac{4T}{\delta_1}} \right\}$$
$$\mathcal{E}_z = \left\{ \sup_{k \leq 2T-1} \|\Gamma_{Z,k}^{-1/2} Z_k\|_2 \leq \sqrt{pm} + \sqrt{2 \log \frac{2T}{\delta_1}} \right\},$$

where $\delta_1 = \sum_{k=T}^{2T-1} \delta/k^2$. Based on Lemma 5.16, Lemma 5.9, and a union bound the all events $\mathcal{E}_{\text{unif}} \cap \mathcal{E}_{\text{unif}}^{\text{PE,s}} \cap \mathcal{E}_x \cap \mathcal{E}_z$ occur with probability at least $1 - 7\sum_{k=T}^{2T-1} \frac{1}{k^2}\delta$. Now we proceed as in the proof of Theorem 5.5. The only different step is the analysis of the sum of $\|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2$, hence we omit the proof for the other terms.

The sum $\|\bar{V}_{k-1}^{-1/2} Z_k\|_2^2$ is upper bounded by:

$$\left(\sum_{k=T}^{2T-1} \left\| \bar{V}_{k-1}^{-1/2} \Gamma_{Z,k}^{1/2} \right\|_2^2 \right) \sup_{k \leq 2T-1} \left\| \Gamma_{Z,k}^{-1/2} Z_k \right\|_2^2$$

There are two cases:

$$\sum_{k=T}^{2T-1} \left\| \bar{V}_{k-1}^{-1/2} \Gamma_{Z,k}^{1/2} \right\|_2^2 \leq 8 \frac{2T}{T-1}, \text{ if } T \geq N_0$$
$$\sum_{k=T}^{2T-1} \left\| \bar{V}_{k-1}^{-1/2} \Gamma_{Z,k}^{1/2} \right\|_2^2 \leq \frac{T}{\lambda} \left\| \Gamma_{Z,2T-1} \right\|_2, \text{ if } T < N_0.$$

Meanwhile, we upper-bound $\|\Gamma_{Z,k}^{-1/2} Z_k\|_2^2$ based on $\mathcal{E}_z$. □

The remaining steps are the same as in the proof of Theorem 5.1.

5.9 Technical lemmas

Lemma 5.17. *Let $c > 0$. The inequality:*

$$k \geq c \log k$$

is true if $k \geq \max\{2c \log 2c, 1\}$.

Proof. If $c \leq e$, then the inequality is satisfied for all $k > 0$. To see why this holds consider $f(k) = k - e\log k$. The minimum is attained at $f(e) = e - e\log e = 0$. Hence, $k \geq e\log k \geq c\log k$. If $c > e$, then the function $k - c\log k$ is increasing for $k \geq c$. Moreover, $2c\log 2c \geq c$. As a result if $k \geq 2c\log 2c$ then also:

$$k - c\log k \geq 2c\log 2c - c\log(2c\log 2c) \geq (c - \frac{c}{e})\log 2c \geq 0$$

where we used Lemma 5.19. □

Lemma 5.18. *Let* $c \geq 0$. *The inequality:*

$$k \geq c\log^2 k$$

is true if $k \geq \max\left\{4c\log^2 4c, 4c\log 4c, 1\right\}$.

Proof. If $c \leq 1$, then the inequality is satisfied for $k \geq 1$. To see why this holds define $f(k) = k - \log^2 k$. Its derivative $f'(k) = 1 - 2\frac{\log k}{k}$ is always positive for $k \geq 1$ since from the proof of Lemma 5.17 $k \geq 2\log k$. Hence $f(k) \geq f(1) = 1$.

Consider now the case $c > 1$ and define $g(k) = k - c\log^2 k$. Its derivative is $g'(k) = 1 - 2c\frac{\log k}{k}$. From Lemma 5.17 $g'(k) \geq 0$, for $k \geq \max\{4c\log 4c, 1\}$. Now, pick $k_1 = 4c\log^2 4c$ and observe that $k_1 \geq 4c\log 4c$ since $4c > e$ and $\log 4c > 1$. Since g is increasing for $k \geq k_1$, it is sufficient to prove that $g(k_1) > 0$. Invoking Lemma 5.19, we obtain the inequality

$$c\log^2(k_1) \leq c\left(\log 4c + \frac{1}{e}\log 4c\right)^2 \leq 4c\log^2 4c = k_1,$$

where (i) follows from Lemma 5.19 below. □

Lemma 5.19. *Let* $c \geq e$, *then the following inequality holds:*

$$\log\log c \leq \frac{1}{e}\log c$$

Proof. Consider function $f(c) = \frac{1}{e}\log c - \log\log c$ and compute the derivative:

$$f'(c) = \frac{1}{ec} - \frac{1}{c\log c}$$

The minimum is attained at e^e. Hence

$$f(c) \geq f(e^e) = 0$$

for all $c \geq e$. □

Part II

Statistical Difficulty of Learning Linear Systems

Chapter 6

Difficulty of System Identification

6.1 Introduction

In this chapter, we study the statistical difficulty of system identification for fully-observed linear systems of form:

$$x_{k+1} = Ax_k + Bu_k + Hw_k, \tag{6.1}$$

where x_k represents the state, u_k represents the control signal, and w_k is the process noise. The statistical analysis of system identification algorithms has a long history Ljung (1999). Until recently, the main focus was providing guarantees for the convergence of system identification in the *asymptotic regime* Deistler et al. (1995); Bauer et al. (1999); Chiuso & Picci (2004), when the number of collected samples N tends to infinity. Under sufficient persistency of excitation Bai & Sastry (1985), system identification algorithms converge and the asymptotic bounds capture very well how the identification error decays with N qualitatively.

However, our standard asymptotic tools (e.g. the Central Limit Theorem), do not always capture all finite-sample phenomena (Vershynin, 2018, Ch 2). Moreover, the identification error depends on various system theoretic constants, like the state space dimension n, which might be hidden under the big-O notation in the asymptotic bounds. As a result, system identification limitations, like the curse of dimensionality, although known to practitioners, are not always reflected in the theoretical asymptotic bounds.

With the advances in high-dimensional statistics Vershynin (2018), there has been a recent shift from asymptotic analysis with infinite data to statistical analysis of system identification with finite samples. Over the past two years there have been significant advances in understanding finite sample system identification for both fully-observed systems Campi & Weyer (2002); Dean et al. (2017); Simchowitz et al. (2018); Faradonbeh et al. (2018a); Sarkar & Rakhlin (2018); Fattahi et al. (2019); Jedra & Proutiere (2019); Wagenmaker & Jamieson (2020) as well as partially-observed systems Oymak & Ozay (2018); Sarkar et al. (2019); Simchowitz et al. (2019); Tsiamis & Pappas (2019); Lee & Lamperski (2020); Zheng & Li (2020); Lee (2020); Lale et al. (2020b); Kozdoba et al. (2019); Tsiamis & Pappas (2020). A tutorial can be found in Matni & Tu (2019). The above approaches offer mainly *data-independent* bounds which reveal how the state dimension n and other system theoretic parameters affect the sample complexity of system identification *qualitatively*. This is different from finite sample data-dependent bounds-see for example bootstrapping Dean et al. (2017) or Carè et al. (2018), which might be more tight and more suitable for applications but do not necessarily reveal this dependence.

Despite these advances, we still do not fully understand the fundamental limits of when identification is easy or hard. In this chapter, we define as statistically easy, classes of systems whose finite-sample complexity is polynomial with the system dimension. Most prior research in the finite-sample analysis of fully observed systems falls in this category by assuming system (6.1) is fully excited by the process noise w_k. We define as statistically hard, classes of linear systems whose worst-case sample complexity is at least exponential with the system dimension, regardless of the learning algorithm. Using recent tools from minimax theory Jedra & Proutiere (2019), we show that classes of linear systems which are statistically hard to learn do indeed exist. Such system classes include, for example, under-actuated systems with weak state coupling. The fact that linear systems may contain exponentially hard classes has implications for broader classes of systems, such as nonlinear systems, as well as control algorithms, such as the linear quadratic regulator Recht (2019) and reinforcement learning Du et al. (2019); Jiang et al. (2017).

By examining classes of linear systems that are statistically easy or hard, we quickly arrive at the conclusion that system theoretic properties, such as controllability, fundamentally affect the hardness of identification. In fact, as we show in this chapter, structural properties like the controllability index can crucially affect learnability, determining whether a problem is hard or not. In summary, our contributions are the following:

–**Learnability of dynamical systems.** We define two novel notions of learnability for classes of dynamical systems. A class of systems is easy to learn if it exhibits polynomial sample complexity with respect the state dimension n. It is hard to learn if for any possible learning algorithm it has exponential worst-case complexity.

–**Exponential sample complexity is possible.** We identify classes of under-actuated linear systems whose worst-case sample complexity increases exponentially with the state dimension n regardless of learning algorithm. These hardness results hold even for robustly controllable systems.

–**Controllability index affects sample complexity.** We prove that under the least squares algorithm, the sample complexity is upper-bounded by an exponential function of the system's controllability index. This implies that if the controllability index is small $O(1)$ (with respect to the dimension n), the sample complexity is guaranteed to be polynomial generalizing previous cases. If, however, the index grows linearly $\Omega(n)$, then there exist non-trivial linear systems which are exponentially hard to identify.

–**New controllability Gramian bound** Our sample complexity upper bound is a consequence of a new result that is of independent, system theoretic interest. We prove that for robustly controllable systems, the least singular value of the controllability Gramian can grow at most exponentially with the controllability index. Although it has been observed empirically that the Gramian might be affected by the curse of dimensionality Baggio et al. (2019), to the best of our knowledge this theoretical bound is new and has implications beyond system identification.

Notation: The transpose operation is denoted by $(\cdot)'$ and the complex conjugate (Hermitian transpose) by $*$. By $e_i \in \mathbb{R}^n$ we denote the $i-$th canonical vector. By $\sigma_{\min}$ we denote the least singular value. $\succeq$ denotes comparison in the positive semidefinite cone. The identity matrix of dimension n is denoted by I_n. The spectral norm of a matrix A is denoted by $\|A\|_2$. The notion of controllability and other related concepts are reviewed in Sections 6.8, 6.9.

6.2 Learnability of System Classes

Consider system (6.1), where $x_k \in \mathbb{R}^n$ is the state and $u_k \in \mathbb{R}^p$ is the input. By $w_k \in \mathbb{R}^r$ we denote the process noise which is assumed to be Gaussian, i.i.d. with covariance I_r. Without loss of generality the initial state is assumed to be zero $x_0 = 0$.

Assumption 6.1. *All state parameters are bounded: $\|A\|_2, \|B\|_2, \|H\|_2 \leq M$, for some positive constant $M > 0$. The noise has unknown dimension r and can be degenerate $r \leq n$. All parameters A, B, H, r are considered unknown. Matrices B, H have full column rank $\mathrm{rank}(B) = p \leq n$, $\mathrm{rank}(H) = r \leq n$. We also assume that the system is non-explosive $\rho(A) \leq 1$. Finally, we assume that the control inputs have bounded energy $\mathbb{E}u_t' u_t \leq M$.*

This setting is rich enough to provide insights about the difficulty of the general learning problem. To simplify the setting we assume that the system is non-explosive. The analysis of unstable systems is left for future research.

A system identification (SI) algorithm $\mathcal{A}$ receives a finite number N of input-state data $(x_0, u_0), \dots, (x_N, u_N)$ generated by system (6.1), and returns an estimate of the unknown system's parameters $\hat{A}_N, \hat{B}_N, \hat{H}_N$. We denote by N the number of collected input-state samples, which are generated during a single roll-out of the system, that is a single trajectory of length N. For simplicity, we focus only on the estimation of matrix A here.

Our goal is to study when the problem of system identification is fundamentally easy or hard. The difficulty is captured by the sample complexity, i.e. how many data N do we need to achieve small identification error with high probability. Formally, let $\epsilon > 0$, $0 < \delta < 1$ be the accuracy and confidence parameters respectively. Then, the sample complexity is the smallest possible number of samples N such that with probability at least $1-\delta$ we can estimate A with small error $\|A-\hat{A}_N\|\leq\epsilon$. Naturally, the sample complexity increases as the accuracy/confidence parameters ϵ,δ decrease. The sample complexity also increases in general with the state-space dimension n and the bound M on the state space parameters.

Ideally, the sample complexity should grow slowly with $n, M, \epsilon^{-1}, \delta^{-1}$. Inspired by Provably Approximately Correct (PAC) learning Shalev-Shwartz & Ben-David (2014); Dann et al. (2017), we classify an identification problem as easy when the sample complexity depends polynomially on $n, M, \epsilon^{-1}, \delta^{-1}$. For brevity we will use the symbol S to denote the tuple $S=(A,B,H)$. Let $\mathbb{P}_S$ denote the probability distribution of the input-state data when the true parameters of the system are equal to S and we apply a control law $u_t\in\mathcal{F}_t$, where $\mathcal{F}_t\triangleq\sigma(x_0,u_0,\dots,u_{t-1},x_t)$ is the sigma algebra generated by the previous outputs and inputs. By $\mathcal{C}_n$ we will denote a class of systems with dimension n.

Definition 1 (poly-learnable classes)**.** *Let $\mathcal{C}_n$ be a class of systems. Consider a trajectory of input-state data $(x_0,u_0),\dots,(x_N,u_N)$, which are generated by a system S in $\mathcal{C}_n$ under some control law $u_t\in\mathcal{F}_t$, $t\leq N$. We call the class $\mathcal{C}_n$* poly$(n)-$*learnable if there exists an identification algorithm such that the sample complexity is polynomial: for any confidence $0\leq\delta<1$ and any tolerance $\epsilon>0$:*

$$\begin{aligned}&\sup_{S\in\mathcal{C}_n}\mathbb{P}_S(\|A-\hat{A}_N\|\geq\epsilon)\leq\delta, \qquad (6.2)\\ &\textit{for } N\geq\mathrm{poly}(n,1/\epsilon,\log 1/\delta,M),\end{aligned}$$

where poly$(\cdot)$ *is some polynomial function.*

Definition 1 provides an intuitive definition for a class $\mathcal{C}_n$ of linear systems whose system identification problem is easy. To prove that a class of systems $\mathcal{C}_n$ is easy, it suffices to provide one algorithm that performs well for any system $S\in\mathcal{C}_n$ in the sense that it requires **at most** a polynomial number of samples. This means that we should obtain sample complexity **upper bounds** across all $S\in\mathcal{C}_n$ which is what the the supremum over $S\in\mathcal{C}_n$ achieves in (6.2). Otherwise, we can construct trivial algorithms that perform well only on

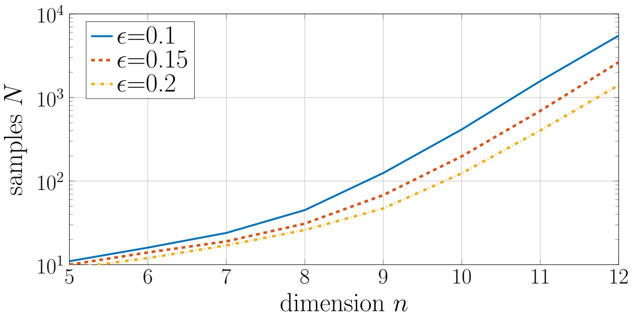


Figure 6.1: The minimum number of samples N such that the (empirical) average error $\mathbb{E}\|A-\hat{A}_N\|_2$, for identifying (6.3), is less than ϵ. The sample complexity appears to be increasing exponentially with the dimension n under the least squares algorithm.

one system and fail to identify the other.

In recent work Simchowitz et al. (2018); Sarkar & Rakhlin (2018); Fattahi et al. (2019), it was shown that under the least squares algorithm, the sample complexity of learning linear systems is polynomial. As we review in Section III, these results hold for classes of linear systems where the noise is isotropic and hence directly exciting all states.

However, if we relax the last assumption it turns out that the sample complexity might degrade dramatically. To raise this issue, consider the following example. Let $J_n(1)$ be a Jordan block of size n with eigenvalue 1 and let e_n be the $n-$th canonical vector. We simulate the performance of least squares identification for the system

$$x_{k+1} = 0.5J_n(1)x_k + e_n(u_k + w_k) \tag{6.3}$$

Note that in system (6.3) the process noise is no longer isotropic. Figure 6.1 shows the minimum number of samples N required to achieve (empirical) average error $\mathbb{E}\|A-\hat{A}_N\| \leq \epsilon$ (the details of the simulation can be found in Section 6.6). It seems that the sample complexity increases exponentially rather than polynomially. Are the results in Figure 6.1 due to the choice of the algorithm or is there a fundamental limitation for all system identification algorithms? We pose the following fundamental problem.

Problem 6.1. *Do there exist classes of linear systems which are hard to learn, meaning* not poly-*learnable by any system identification algorithm? Furthermore, can the sample* *complexity for a class of linear systems be exponential with state dimension n?*

A class of linear systems $\mathcal{C}_n$ that is not poly-learnable will be viewed as hard. By negating Definition 1, this notion of hardness means that given *any* system identification algorithm, there exist instances $S \in \mathcal{C}_n$ that cannot have polynomial sample complexity. In other words, a system class $\mathcal{C}_n$ is classified as hard when its impossible to find any system identification algorithm that achieve polynomial sample complexity for all $S \in \mathcal{C}_n$. This can be viewed as a fundamental statistical limitation for the chosen class of systems $\mathcal{C}_n$.

Motivated by Figure 6.1, we define an important subclass of hard problems, namely linear system classes that have worst-case sample complexity that grows exponentially with the dimension n regardless of identification algorithm choice.

Definition 2 (exp-hard classes)**.** *Let $\mathcal{C}_n$ be a class of systems of dimension n. Consider a trajectory of input-output data $(x_0, u_0), \dots, (x_N, u_N)$, which are generated by a system S in $\mathcal{C}_n$ under some control law $u_t \in \mathcal{F}_t$, $t \le N$. We call a class $\mathcal{C}_n$ of systems* $\exp(n)$*-hard if the sample complexity is at least exponential with the dimension n: there exist confidence $0 \le \delta < 1$ and tolerance ϵ parameters such that for any identification algorithm:*

$$\sup_{S \in \mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \ge \epsilon) \le \delta,$$

$$\textit{only if } N \ge \exp(n),$$

where $\exp(n)$ *denotes an exponential function of n.*

System classes $\mathcal{C}_n$ that are exp-hard are an important subset of hard system classes as they are clearly not poly-learnable. However, not all classes that are not poly-learnable are exp-hard.

In order to show that a class of systems $\mathcal{C}_n$ is exp-hard, one must show that for **any** system identification algorithm the worst-case sample complexity is at least exponential in state dimension n. Contrary to poly-learnable problems, for exponential hardness we should establish sample complexity **lower bounds**.

In this chapter, we first address Question 6.1 and show that exp-hard classes of linear systems do indeed exist. While this can be viewed as a fundamental statistical limitation for all system identification algorithms, our results open a new direction of research that classifies when linear systems are easy to learn and when they are hard to learn. This leads to the following important question.

Problem 6.2. *When is a class of linear systems $\mathcal{C}_n$ guaranteed to be* poly*-learnable?*

Based on prior work, we already have partial answers to Question 6.2 as we know that linear systems with isotropic noise are poly-learnable. In Section 6.5, we seek to broaden the classes of poly-learnable systems and discover their relation to fundamental system theoretic properties such as controllability.

While Definitions 1, 2 are inspired by PAC learning, they have a different flavor. One of the differences is that the guarantees in Definitions 1, 2 are stated in terms of recovering the state-space parameters, while in PAC learning, they would be stated in terms of the prediction error of the learned model or informally $\sum_{k=0}^{N-1} \mathbb{E}\|x_k - \hat{A}x_{k-1} - \hat{B}u_{k-1}\|^2$.

6.3 Directly-excited systems are poly-learnable

In this section, we revisit state-of-the-art results in finite-sample complexity for fully-observed linear systems and re-establish that they all lead to polynomial sample complexity. In prior work Simchowitz et al. (2018); Sarkar & Rakhlin (2018); Fattahi et al. (2019), the class of linear systems considered assumes that the stochastic process noise is isotropic, i.e. $HH' = \sigma_w^2 I_n$. Since all states are directly excited by the process noise, all modes of the system are captured sufficiently in the data. To obtain polynomial complexity, it suffices to use the least squares identification algorithm

$$\begin{bmatrix} \hat{A}_N & \hat{B}_N \end{bmatrix} = \arg\min_{\{F,G\}} \sum_{t=0}^{N-1} \|x_{t+1} - Fx_t - Gu_t\|_2^2 \tag{6.4}$$

with white noise inputs $u_t \sim \mathcal{N}(0, \sigma_u^2 I)$. Based on the algorithm analysis from Simchowitz et al. (2018), let k be a fixed time index which is much smaller than the horizon N (see

Theorem 2.1 in Simchowitz et al. (2018) for details). Let $0 < \delta < 1$ and ϵ be the confidence and accuracy parameters respectively. Then, with probability at least $1 - \delta$, the error is $\|A - \hat{A}_N\|_2 \leq \epsilon$ if:

$$N \geq \frac{c\sigma_w^2}{\sigma_{\min}(\Gamma_k)} \frac{1}{\epsilon^2} \left(n \log \frac{n}{\delta} + \log \det(\Gamma_N \Gamma_k^{-1}) \right),$$

where c is a universal constant, and $\Gamma_k = \sigma_u^2 \Gamma_k(A, B) + \sigma_w^2 \Gamma_k(A, I_n)$ is the (combined) controllability Gramian. Uunder the isotropic noise assumption, the least singular value of the Gramian Γ_k is bounded away from zero, $\sigma_{\min}(\Gamma_k) \geq \sigma_w^2$.

In a slight departure from Simchowitz et al. (2018); Sarkar & Rakhlin (2018); Fattahi et al. (2019), we can show that the determinant of the Gramian $\det(\Gamma_N)$ can only increase at most polynomially with the number of samples N and exponentially with state dimension n. This is a direct consequence of the following lemma, which is a new result.

Lemma 6.1. *Let $A \in \mathbb{R}^{n \times n}$ have all eigenvalues inside or on the unit circle, with $\|A\|_2 \leq M$. Then, the powers of matrix A are bounded by:*

$$\left\| A^k \right\|_2 \leq (ek)^{n-1} \max\{M^n, 1\} \tag{6.5}$$

Lemma 6.1 enables us to eliminate the dependence on the condition number of the Jordan form's similarity transformation, which exists in prior bounds and can be arbitrarily large. We avoid this dependence by using the Schur form of A Horn & Johnson (2012). While this does not alter the already known sample complexity results, it allows us to have sample complexity bounds that are uniform across all systems that satisfy Assumption 6.1.

As a result of Lemma 6.1, we obtain that the system identification problem for linear systems with isotropic noise has polynomial sample complexity. The result can be broadened to the more general case of direct excitation, where the covariance is lower bounded by $HH' + BB' \succeq \sigma_w^2 I_n$, for some $\sigma_w > 0$, as the following theorem states.

Theorem 6.1 (Directly-excited). *Consider the class $\mathcal{C}_n$ of directly-excited systems $S = (A, B, H) \in \mathbb{R}^{n \times (n+p+r)}$ such that Assumption 6.1 is satisfied with covariance $HH' + BB' \succeq$*

$\sigma_w^2 I_n$*, for some* $\sigma_w > 0$*. The class* $\mathcal{C}_n$ *is* poly$-$*learnable under the least squares system identification algorithm with white noise input signals* $u_k \sim \mathcal{N}(0, I_p)$*.*

Proof. It follows as a special case of Theorem 6.4 for controllability index $\kappa = 1$. □

Directly excited systems includes fully-actuated systems (number of inputs equal to the number of states $p = n$), or systems with isotropic noise as special cases. However, having direct excitation might not always be the case. The combined noise and input matrices might be rank-deficient. For example, we might have actuation noise as in:

$$x_{t+1} = Ax_t + B(u_t + w_t).$$

In general, the noise might be ill-conditioned (zero across certain directions), while it might be physically impossible to actuate every state of the system. We call such systems under-actuated or under-excited. It might still be possible to identify underactuated systems, e.g. if the pair $(A, \begin{bmatrix} H & B \end{bmatrix})$ is controllable. However, as we prove in the next section, the identification difficulty might increase dramatically.

6.4 Exp-hard system classes

In this section, we show that there exist common classes of linear systems which are impossible or hard to identify with a finite amount of samples. As we will see, this can happen when systems are under-actuated and under-excited. When only a limited number of system states is directly driven by inputs (or excited by noise) and the remaining states are only indirectly excited, then identification can be inhibited.

6.4.1 Controllable systems with infinite sample complexity

For presentation simplicity, let us assume that there are no exogenous inputs $B = 0$. Similar results also hold when $B \neq 0$–see Remark 7. To fully identify the unknown matrix A, it is necessary that the pair (A, H) is controllable. Furthermore, let's assume that the noise is meaningful, that is $\sigma_{\min}(H) \geq \sigma$ for some $\sigma > 0$. However, controllability of (A, H) and $\sigma_{\min}(H) \geq \sigma$ are not sufficient to ensure system identification from a finite numer of samples. The following, perhaps unsurprising theorem, shows that for this class of linear systems, the worst-case sample complexity is infinite.

Theorem 6.2 (Controllability is not sufficient for finite sample complexity)**.** *Consider the class $\mathcal{C}_n$ of systems $S = (A, H) \in \mathbb{R}^{n\times(n+r)}$ such that Assumption 6.1 is satisfied with (A, H) controllable, and $\sigma_{\min}(H) \geq \sigma$ for some $\sigma > 0$. For any system identification algorithm the sample complexity is infinite: there exist a failure probability $0 \leq \delta < 1$ and a tolerance $\epsilon > 0$ such that we cannot achieve*

$$\sup_{S\in\mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta$$

with a finite number of samples N.

Theorem 6.2 clearly shows that we may need stronger notions of controllability, as done in Section 6.4.2, in order to find classes of systems whose sample complexity is finite. The proof of Theorem 6.2 uses tools from minimax theory Jedra & Proutiere (2019). Adapting these tools in our setting results in the following.

Lemma 6.2 (Minimax bounds)**.** *Let $\mathcal{C}_n$ be a class of systems. Consider a confidence $0 < \delta < 1$ and an accuracy parameter $\epsilon > 0$. Denote by $S_1, S_2 \in \mathcal{C}_n$ any pair of two systems with A_1, H_1, A_2, H_2 the respective unknown matrices, such that $\|A_1 - A_2\| \geq 2\epsilon$. Let $\mathrm{KL}(\mathbb{P}_{S_1}, \mathbb{P}_{S_2})$ be the Kullback-Leibler divergence between the probability distributions of the data when generated under S_1, S_2 respectively. Then for any identification algorithm*

$$\sup_{S\in\mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta$$

holds only if

$$\mathrm{KL}(\mathbb{P}_{S_1}, \mathbb{P}_{S_2}) \geq \log \frac{1}{3\delta}, \tag{6.6}$$

for all such pairs $S_1, S_2 \in \mathcal{C}_n$.

Proof. Let S_1, S_2 be any pair satisfying the conditions. We trivially have that:

$$\sup_{S\in\mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta$$

only if

$$\sup_{S\in\{S_1,S_2\}} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta.$$

The remaining proof is identical to (Jedra & Proutiere, 2019, Proposition 2), where we replaced constant 2.4 with 3 for simplicity and we did not expand the expression for $\mathrm{KL}(\mathbb{P}_{S_1}, \mathbb{P}_{S_2})$ explicitly (term $\mathbb{E}_A(L_t)$ in Jedra & Proutiere (2019)). □

Intuitively, to find difficult learning instances we construct systems which are sufficiently separated (2ϵ away). Meanwhile, the systems should be similar enough to generate data with as indistinguishable distributions as possible (small KL divergence). If the system is hard to excite, then the distributions of the states will look similar under many different matrices A, leading to smaller KL-divergence. Unless we bound the pair (A, H) away from uncontrollability, it might be impossible to satisfy (6.6) for all pairs of systems with a finite number of samples. For example consider:

$$A = \begin{bmatrix} 0 & \alpha & 0 \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

It requires an arbitrarily large number of samples to learn α if the coupling β between $x_{t,2}$ and $x_{t,3}$ is arbitrarily small. The distribution of $x_{t,1}$ remains virtually the same as we perturb α, since the state $x_{t,2}$ is under-excited for small β.

6.4.2 Robustly controllable systems can be exp-hard

Theorem 6.2 implies that we need to bound the system away from uncontrollability in order to obtain non-trivial sample complexity bounds. In order to formulate this, we review the notion of distance from uncontrollability, which is the norm of the smallest perturbation that makes (A, H) uncontrollable.

Definition 3 (Distance from uncontrollability Eising (1984))**.** *Let $(A, H) \in \mathbb{R}^{n\times(n+r)}$ be controllable. Then, the distance from uncontrollability is given by:*

$$d(A,H) \triangleq \inf \left\{ \| \begin{bmatrix} \Delta A & \Delta H \end{bmatrix} \|_2 : \right. \\ \left. (A+\Delta A, H+\Delta H) \text{ uncontrollable} \right\}, \tag{6.7}$$

where perturbations $(\Delta A, \Delta H) \in \mathbb{C}^{n\times(n+r)}$ are complex.

Let us now consider linear systems that are robustly controllable. That is, classes of controllable linear systems whose distance from uncontrollability is lower bounded. The lower bound is allowed to degrade gracefully (polynomially) with the system dimension n.

Assumption 6.2 (Robust Controllability)**.** *Assume that system (A, H) is robustly controllable, that is $(A, H) \in \mathbb{R}^{n\times(n+m)}$ is μ-away from uncontrollability:*

$$d(A, H) \geq \mu, \tag{6.8}$$

for some positive $\mu \geq 0$, with $\mu^{-1} \leq \mathrm{poly}(n)$.

Assumption 6.2 is not restrictive as long as we allow the bound to degrade with the dimension. Common systems like the $n-$th order integrator have distance that degrades linearly with n–see Lemmas 6.3, 6.4 in Section 6.9. However, even for system classes that satisfy Assumption 6.2, the next theorem shows that system identification can be exp-hard.

Theorem 6.3 (Exp(n)-hard classes)**.** *Consider the set $\mathcal{C}_n$ of systems $S = (A, H)$ such that Assumptions 6.1, 6.2 are satisfied with $d(A, H) \geq \mu = 8(n + 1)^{-1}$. Then, for any system* *identification algorithm $\mathcal{A}$ the sample complexity is exponential in the state dimension n. There exist a confidence $0 \leq \delta < 1$ and a tolerance $\epsilon > 0$ such that*

$$\sup_{S \in \mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta$$

is satisfied only if

$$N \geq \frac{4^{n-3}}{3\epsilon^2} \log \frac{1}{\delta}.$$

Theorem 6.3 shows that even for robustly controllable classes of linear systems satisfying Assumptions 6.1, 6.2, any system identification algorithm will have worst-case sample complexity that depends exponentially on the system dimension n. The proof of Theorem 6.3 is based once more on minimax theory used in Lemma 6.2.

The reason for this learning difficulty is due to the need for indirect excitation. Consider, for example, chained systems, where every state indirectly excites the next one. If the states are weakly-coupled, then the exploratory signal (noise or input) attenuates exponentially fast along the chain. As a concrete example, consider the following system for $\rho < 0.5$:

$$A = \begin{bmatrix} \rho & \rho & 0 & \cdots & 0 & 0 \\ 0 & \rho & \rho & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & \rho & \rho \\ 0 & 0 & 0 & \cdots & 0 & \rho \end{bmatrix}, \; H = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 0 & \rho \end{bmatrix} \tag{6.9}$$

which satisfies Assumptions 6.1, 6.2. Matrix A has a chained structure with weak coupling between the states. Noise can only excite states $x_{t,1}, x_{t,n}$ directly. Until the exploratory noise signal reaches $x_{t,2}$ it decreases exponentially fast with the dimension n. As a result, it is difficult to learn A_{12} due to lack of excitation. In terms of Lemma 6.2, the distribution of $x_{t,1}$ will remain virtually the same if we perturb A_{12} since $x_{t,2}$ is under-excited.

Remark 7 (Exogenous inputs)**.** When $B \neq 0$ similar results hold but with an additional interpretation. Consider system (6.9) but with $H = e_1$, $B = \rho e_n$. Then, if we apply white-noise input signals we have two possibilities: i) the control inputs have bounded energy per Assumption 6.1 but we suffer from exponential sample complexity or ii) we obtain polynomial sample complexity but we allow the energy of the inputs to increase exponentially with the dimension. From this alternative viewpoint a system is hard to learn if it requires exponentially large control inputs.

Remark 8. The constant 8 in $8(n+1)^{-1}$ in the statement of Theorem 6.3 is not important in our analysis. We could modify Theorem 6.3 so that 8 can be replaced by any smaller constant. In particular, we can decrease 8 by considering systems with smaller chains, which still have exponential sample complexity. Instead of system (6.9), we can consider for example the following. Let $J_{\lfloor n/m \rfloor}(1)$ be the Jordan block of size $\lfloor n/m \rfloor$, for some m, and eigenvalue 1 and define

$$A = \left[\begin{array}{c|c} \rho J_{\lfloor n/m \rfloor}(1) & 0 \\ \hline 0 & I_{n-\lfloor n/m \rfloor} \end{array}\right],\ H = \left[\begin{array}{cc|ccc} e_1 & \rho e_{\lfloor n/m \rfloor} & e_{\lfloor n/m \rfloor+1} & \cdots & e_n \end{array}\right].$$

Notice that we reduced the size of the chain by $1/m$ and we added $n-\lfloor n/m \rfloor$ directly excited states. By increasing m, we can achieve a larger distance to uncontrollability (constant smaller than 8). However, we will still have exponential sample complexity of the order of at least $\lfloor n/m \rfloor$, based on the length of the chain.

6.5 Controllability index affects learnability

Structural system properties of an underactuated system, such as the chained structure in the dynamics, can be critical in making system identification easy or hard. This poses novel questions about understanding how system theoretic properties affect system learnability as defined in Definitions 1 and 2. We begin a new line of inquiry by characterizing how the controllability index κ, a critical structural system property, affects the statistical properties of system identification.

A brief review of the concept of controllability index can be found in Section 6.8. It can be viewed as a structural measure of whether a system is directly actuated or underactuated resulting in long chains. In this section, we consider systems which have controllability index κ. To avoid pathological examples, we require that the controllability index is equal to κ in a robust way. Inspired by Definition 3, we define a notion of distance from losing controllability index.

Definition 4 (Distance from losing index)**.** *Let $(A,H)\in\mathbb{R}^{n\times(n+r)}$ be controllable and let $n\geq\tau\geq 0$ be some integer smaller than n. The distance from losing index τ is given by:*

$$d_\tau(A,H)\triangleq\inf_{(\Delta A,\Delta H)\in\mathbb{C}^{n\times(n+r)}}\left\{\|\begin{bmatrix}\Delta A & \Delta H\end{bmatrix}\|_2:\kappa(A+\Delta A,H+\Delta H)>\tau\right\},\tag{6.10}$$

where $\kappa(F,G)$ denotes the controllability index of a pair (F,G) with $\kappa(F,G)=\infty$ when (F,G) is uncontrollable.

In other words, the distance $d_\tau(A,H)$ is the smallest perturbation such that the controllability index is increased by at least $\tau+1$. Note that when $\tau=n$, then the above definition coincides with the distance to uncontrollability. To avoid pathological systems which are arbitrarily close to losing controllability index, we make the following assumption.

Assumption 6.3 (Robust Controllability Index)**.** *Assume that system $(A,H)\in\mathbb{R}^{n\times(n+r)}$ has μ-robust controllability index κ, that is $\kappa(A,H)=\kappa$ and*

$$d_\kappa(A,H)\geq\mu,\tag{6.11}$$

for some positive $\mu\geq 0$, with $\mu^{-1}\leq\mathrm{poly}(n)$.

The following theorem, is the first result connecting the controllability index with sample complexity bounds.

Theorem 6.4 (Controllability index-dependent upper bounds)**.** *Consider the set $\mathcal{C}_n$ of systems $S=(A,B,H)$ such that Assumption 6.1 is satisfied. Assume that the controllability index of all pairs $(A,\begin{bmatrix}H & B\end{bmatrix})$ in the class is upper bounded by κ. Let Assumption 6.3 be* *satisfied for the all pairs* $(A, \begin{bmatrix} H & B \end{bmatrix})$. *Then, under the least squares system identification algorithm and white noise inputs* $u_k \sim \mathcal{N}(0, I_p)$, *we obtain that*

$$\sup_{S \in \mathcal{C}_n} \mathbb{P}_S(\|A - \hat{A}_N\| \geq \epsilon) \leq \delta$$

is satisfied for

$$N \geq \mathrm{poly}^{\kappa}(n, M)\mathrm{poly}(\epsilon^{-1}, \log 1/\delta).$$

Theorem 6.4 formalizes our intuition since the controllability index is the length of the chain from input excitation towards the most distant state in the chain. Hence, systems with a large number of inputs (or noise) and small controllability index ($\kappa << n$) are easy to identify. The directly excited case with isotropic noise, presented in Theorem 6.1, is a special case corresponding to a controllability index $\kappa = 1$, recovering prior polynomial bounds.

The implications of Theorems 6.3, 6.4 illustrate the impact controllability properties have on system learnability–see Figure 6.4. Classes of systems with small controllability index $O(1)$ have polynomial sample complexity. Classes where the index grows linearly $\Omega(n)$ can be exponentially hard in the worst case in general. There might still be subclasses of systems with large controllability indexes which nonetheless can be identified with a polynomial number of samples. However, we cannot provide any guarantees without further assumptions.

The proof of Theorem 6.4 crucially depends on the following system theoretic result that bounds the least singular value of the controllability Gramian (a quantitative measure of controllability) with the controllability index (a structural measure of controllability).

Theorem 6.5 (Controllability gramian bound)**.** *Consider a system* (A, H) *that satisfies Assumptions 6.1, 6.3. Let* κ *be its controllability index. Then, the least singular value of the gramian* Γ_κ *is lower bounded by:*

$$\sigma_{\min}^{-1}(\Gamma_\kappa) \leq \mathrm{poly}^{\kappa}(M/\mu).$$

The above theorem is of independent interest, since it states that the controllability index rather than the dimension n controls how fast the controllability Gramian degrades. While the above bound may be loose in general, it gives us qualitative insights about how system structure affects the hardness of input excitation and system identification. Our proof exploits the so-called "staircase" (or Hessenberg) canonical representation (6.13) of state space systems Dooren (2003)–see Section 6.8. The main idea is that if a system is robustly controllable then the coupling between the states is bounded away from zero. Hence, we can avoid the essentially uncontrollable systems of Theorem 6.2 which lead to infinite sample complexity.

6.6 Simulations

We study three simulation scenarios to illustrate the qualitative implications of our results. In the first two cases, we verify that the sample complexity of the least squares algorithm can indeed grow exponentially with the dimension. In the third case, we investigate how the controllability index affects the sample complexity. In all cases, we perform Monte Carlo simulations to compute the empirical mean error $\|A - \hat{A}_N\|_2$ and we count the number of samples required to have error less than ϵ, for some $\epsilon > 0$. For numerical stability in the least squares estimator (6.4) we used a regularization term (ridge regression) with coefficient 0.001.

In the first example in Section 6.2, Figure 6.1, we used 1000 Monte Carlo iterations to approximate the empirical average. We modeled the noise as gaussian with $w_k \sim \mathcal{N}(0, 0.5)$ and used white noise inputs $u_k \sim \mathcal{N}(0, 10)$. The sample complexity of the least squares algorithm seems to be exponential with the dimension. In Section 6.4, we showed that such systems exhibit exponential sample complexity due to the weak coupling between the states.

In the second example, we study the behavior of Jordan blocks actuated from the last state. Let $J_n(\lambda)$ be a Jordan block of dimension n and eigenvalues all λ. We consider the system $A = J_n(\lambda)$, $H = 0.1e_n$, $B = 5e_n$, which means we excite directly only state

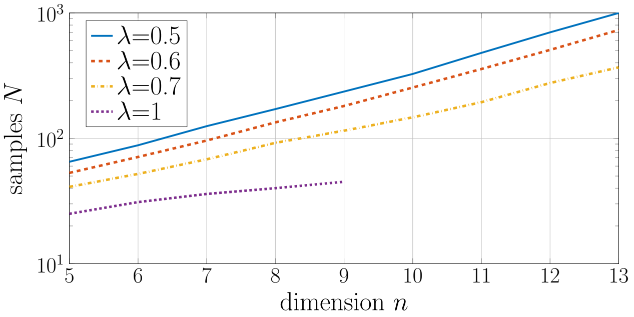


Figure 6.2: Sample complexity of identifying the Jordan block of size n and eigenvalues all λ, actuated from the last state. The figure shows the minimum number of samples N such that the (empirical) average error $\mathbb{E}\|A - \hat{A}_N\|_2$ is less than 0.005. The sample complexity appears to be increasing exponentially with the dimension n for $\lambda < 1$. For $\lambda = 1$, Matlab returns inaccurate results for $n \geq 10$ since the condition number of the data is very large. However, in the regime $5 \leq n \leq 9$, the complexity seems to be polynomial, increasing in 5 sample increments.

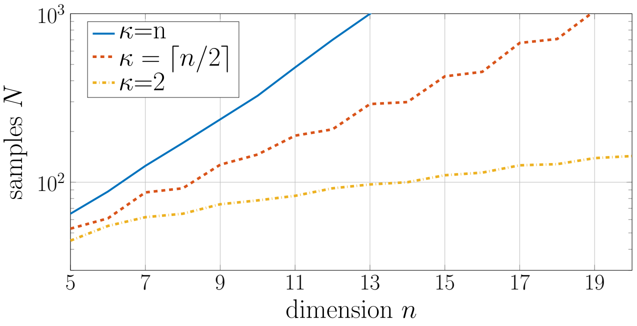


Figure 6.3: Sample complexity of identifying the Jordan block $J_n(0.5)$ of size n and eigenvalues all 0.5, for different values of the controllability index. The figure shows the minimum number of samples N such that the (empirical) average error $\mathbb{E}\|A - \hat{A}_N\|_2$ is less than 0.005. The sample complexity appears to be increasing exponentially with the dimension n for $\kappa = \Theta(n)$. For $\kappa = 2$, the sample complexity is much smaller and increases polynomially.

$x_{t,n}$. We repeat the same experiment as before for 1000 Monte Carlo simulations with $w_k, u_k \sim \mathcal{N}(0,1)$ and for $\epsilon = 0.005$. In Figure 6.2, it seems that the complexity of the least squares algorithm is also exponential when $0 < \lambda < 1$. In this case the coupling between the states is not weak. However, certain subspaces might still be hard to excite. As λ approaches the unit circle eigenvalue 1 the complexity improves. For $\lambda = 1$, after $n = 9$ Matlab returned inaccurate results as the condition number of the data becomes very large. Hence, we do not report any results beyond $n = 9$. However, based on simulations for small n it might be possible that the system can be learned by only a polynomial number of samples. The intuition might be that in this case instability helps with excitation Simchowitz et al. (2018).

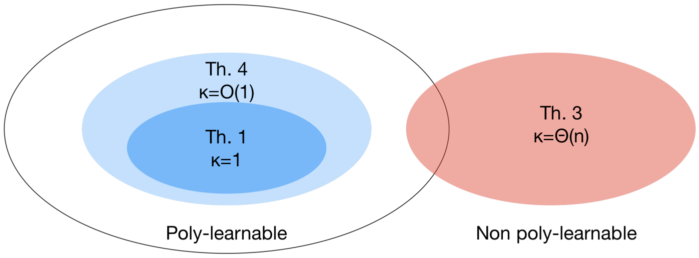


Figure 6.4: Sample complexity classes for linear systems. according to their controllability index.

It is an open problem to prove or disprove exponential lower bounds for the Jordan block when $0 < \lambda < 1$. Similarly, we leave it as an open problem to prove or disprove polynomial upper bounds for the Jordan block when $\lambda = 1$.

In the third example, we consider the Jordan block $A = J_n(0.5)$ with noise $H = 0.1e_n$. We start from $B = 5e_n$ and we gradually add more exogenous inputs to decrease the controllability index: we try $B = 5\begin{bmatrix} e_n & e_{\lceil n/2 \rceil} \end{bmatrix}$ and $B = 5\begin{bmatrix} e_n & e_{n-2} & \dots \end{bmatrix}$ which correspond to indices $\kappa = \lceil n/2 \rceil$ and $\kappa = 2$ respectively. We repeat the same experiment as before for 1000 Monte Carlo simulations with $w_k, u_k \sim \mathcal{N}(0,1)$ and for $\epsilon = 0.005$. In Figure 6.3, it seems that the sample complexity remains exponential when $\kappa = \lceil n/2 \rceil$. However, when $\kappa = 2$ there is a phase transition and the sample complexity becomes polynomial with the dimension.

6.7 Conclusion

The results of this chapter paint a broader and more diverse landscape about the statistical complexity of learning linear systems, summarized in Figure 6.4 according to the controllability index κ of the considered system class. While statistically easy cases that were previously known are captured by Theorem 6.1, we also showed that hard system classes exist (Theorem 6.3). By exploiting structural system theoretic properties, such as the controllability index, we broadened the class of easy to learn linear systems (Theorem 6.4).

Our results pose numerous future questions for exploiting other system properties (e.g. observability) for efficiently learning classes of partially-observed linear systems or nonlinear systems. It remains an open problem to prove whether or not the $n-$th order integrator is poly-learnable as discussed in Section 6.6. Similarly, it is an open problem to prove whether or not the Jordan block of size n and eigenvalues all $0 < \lambda < 1$ has exponential complexity. Finally, this chapter focuses only on the problem of system identification. In Chapter 7, we will see that similar results hold for the problem of learning to control.

6.8 Controllability-related concepts

We briefly review the concept of controllability and other related concepts. We consider the pair (A, H), but the same definitions hold also for (A, B). The controllability matrix of (A, H) is defined as

$$\mathcal{C}_k(A, H) \triangleq \begin{bmatrix} H & AH & \cdots & A^{k-1}H \end{bmatrix}, \, k \geq 1.$$

The pair (A, H) is *controllable* when the controllability matrix $\mathcal{C}_n(A, H)$ has full column rank n. The *controllability Gramian* at time k is defined as :

$$\Gamma_k(A, H) \triangleq \mathcal{C}_k(A, H)\mathcal{C}_k'(A, H) = \sum_{i=0}^{k-1} A^i H H'(A')^i.$$

If H is not a column matrix, the full column rank condition might be satisfied earlier for some $k \leq n$. The minimum time that we achieve controllability is the *controllability index*:

$$\kappa(A, H) \triangleq \min \left\{ k \geq 1 : \operatorname{rank}(\mathcal{C}_k(A, H)) = n \right\}. \tag{6.12}$$

It is the lag between the time the disturbance w_t is applied and the time $t + \kappa$ by which we see the effect of that disturbance in all states. This lag is non-trivial if the number of disturbances $r < n$ is smaller than the number of states; in this case we call the system underactuated.

Based on the fact that the rank of the controllability matrix at time κ is n, we can show that the pair (A, H) admits the following canonical representation, under a unitary similarity transformation Dooren (2003).

Proposition 6.1 (Staircase form)**.** *Consider a controllable pair (A, H) with controllability index κ and controllability matrix $\mathcal{C}_k$, $k \geq 0$. There exists a unitary similarity transformation U such that $U'U = UU' = I$ and:*

$$\begin{aligned} U'H &= \begin{bmatrix} H_1' & 0 & \cdots & 0 \end{bmatrix}' \\ U'AU &= \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,\kappa-1} & A_{1,\kappa} \\ A_{2,1} & A_{2,2} & \cdots & A_{3,\kappa-1} & A_{2,\kappa} \\ 0 & A_{3,2} & \cdots & A_{3,\kappa-1} & A_{3,\kappa} \\ 0 & 0 & \cdots & A_{4,\kappa-1} & A_{4,\kappa} \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & A_{\kappa,\kappa-1} & A_{\kappa,\kappa} \end{bmatrix}, \end{aligned} \tag{6.13}$$

where $A_{i,j} \in \mathbb{R}^{r_i \times r_j}$ are block matrices, with $r_i = \operatorname{rank}(\mathcal{C}_i) - \operatorname{rank}(\mathcal{C}_{i-1})$, $r_1 = r$, $H_1 \in \mathbb{R}^{r \times r}$. Moreover, the matrices $A_{i+1,i}$ have full row rank $\operatorname{rank}(A_{i+1,i}) = r_{i+1}$ *and the sequence r_i is decreasing.*

The above representation is useful as it captures the coupling between the several substates via the matrices $A_{i+1,i}$. If one of these matrices $A_{i+1,i}$ is close to zero then the system will be close to being uncontrollable. On the other hand, if a system is robustly controllable then these matrices are bounded away from being row-rank deficient. Since the similarity transformation is unitary it does not affect properties of the system like the minimum singular value of the controllability Gramian. The proof of Theorem 6.5 exploits the above ideas–see Section 6.10.5 for more details.

6.9 Distance from uncontrollability properties

In this section we review properties of the distance from uncontrollability. The main focus is to prove that standard systems, like the integrator, have distance to uncontrollability which degrades linearly with the dimension n.

Lemma 6.3. *Let $0 < \rho < 1$ and consider the perturbed $n-th$ order integrator:*

$$A = \rho \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}, \, H = \rho \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$$

The distance from uncontrollability is given by

$$d(A, H) = \rho \sin\left(\frac{\pi}{n+1}\right). \tag{6.14}$$

As a result the distance degrades linearly:

$$\rho \frac{2}{n+1} \leq d(A, H) \leq \rho \frac{\pi}{n+1}, \tag{6.15}$$

for $n \geq 1$.

Proof. The proof follows from the fact that the distance form uncontrollability is equivalently given by the formula Eising (1984):

$$d(A, H) = \inf_{s \in \mathbb{C}} \sigma_{\min}(\begin{bmatrix} A - sI & H \end{bmatrix}), \tag{6.16}$$

and results about the eigenvalues of Toeplitz matrices Kulkarni et al. (1999).

In more detail, let * denote the complex conjugate. We have:

$$\begin{bmatrix} A - sI & H \end{bmatrix} \begin{bmatrix} A - sI & H \end{bmatrix}^* = \mathcal{T}_s,$$

where

$$\mathcal{T}_s = \begin{bmatrix} |\rho - s|^2 + \rho^2 & \rho(\rho - s^*) & 0 & 0 \\ \rho(\rho + s^*) & |\rho - s|^2 + \rho^2 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & |\rho - s|^2 + \rho^2 & \rho(\rho - s^*) \\ 0 & 0 & \rho(\rho + s^*) & |\rho - s|^2 + \rho^2 \end{bmatrix} \tag{6.17}$$

is a tri-diagonal Toeplitz matrix, with all diagonal elements equal to $|\rho - s|^2 + \rho^2$, all superdiagonal elements equal to $\rho(\rho - s^*)$ and subdiagonal elements equal to $\rho(\rho + s^*)$. Based on (Kulkarni et al., 1999, Th 2.2), the smallest eigenvalue of $\mathcal{T}$ is equal to:

$$\sigma_{\min}(\mathcal{T}_s) = |\rho - s|^2 + \rho^2 - 2\,|\rho|\,|\rho - s| \cos(\pi/(n+1)).$$

The above quantity is minimized for $\hat{s} = \rho + |\rho| \cos(\pi/(n+1))$. Hence, we can compute the distance to uncontrollability:

$$d(A, H) = \sqrt{\sigma_{\min}(\mathcal{T}_{\hat{s}})} = |\rho| \sin(\pi/(n+1)).$$

Finally (6.15) follows from (6.14) using the elementary calculus inequality

$$\frac{2x}{\pi} \leq \sin x \leq x, \text{ for } 0 \leq x \leq \pi/2,$$

which completes the proof. □

Lemma 6.4. *System (6.9) is μ-bounded away from uncontrollability with $\mu^{-1} \leq \rho^{-1}(n+1)$.*

Proof. Let * denote the complex conjugate. Then we have:

$$\begin{bmatrix} A - sI & H \end{bmatrix} \begin{bmatrix} A - sI & H \end{bmatrix}^* = \mathcal{T}_s + e_1 e_1' \succeq \mathcal{T}_s$$

where $\mathcal{T}_s$ is a tridiagonal Toeplitz matrix defined above in (6.17). Now the proof is identical to the proof of Lemma 6.3 but we have inequality instead of equality:

$$d(A,H) \geq \sqrt{\sigma_{\min}(\mathcal{T}_{\hat{s}})} = |\rho| \sin(\pi/(n+1)) \geq 2\,|\rho|\,/(n+1) \geq |\rho|\,/(n+1).$$

□

Lemma 6.5 (Triangle inequality)**.** *Let $d(A,H)$ be the distance to uncontrollability for some matrices $A \in \mathbb{R}^{n\times n}, H \in \mathbb{R}^{r\times n}$ and let $\|\hat{A} - A\|_2 \leq \epsilon < d(A,H)$ for some matrix $\hat{A} \in \mathbb{R}^{n\times n}$. Then:*

$$d(\hat{A},H) \geq d(A,H) - \epsilon. \tag{6.18}$$

Proof. Assume that $d(\hat{A},H) < d(A,H) - \epsilon$ and let $\begin{bmatrix} \Delta\hat{A} & \Delta\hat{H} \end{bmatrix}$ be the perturbation such that $(\hat{A}+\Delta\hat{A}, H+\Delta\hat{H})$ is uncontrollable with $d(\hat{A},H) = \| \begin{bmatrix} \Delta\hat{A} & \Delta\hat{H} \end{bmatrix} \|_2$. Then, we can define a perturbation for the original pair (A,H) that contradicts the definition of $d(A,H)$:

$$\Delta A = A - \hat{A} + \Delta\hat{A},\ \Delta H = \Delta\hat{H}.$$

The perturbation makes (A,H) uncontrollable and by the triangle inequality, it has norm $\| \begin{bmatrix} \Delta A & \Delta H \end{bmatrix} \|_2 \leq d(\hat{A},H) + \epsilon < d(A,H)$. Since this is impossible (6.18) holds. □

6.10 Proofs

6.10.1 Proof of Lemma 6.1

In this section, we establish upper bounds on the gramian matrices Γ_k. Contrary to previous approaches we avoid using the Jordan form of matrix A. We do not want our bounds to depend on the condition number of the Jordan transformation which can be ill-posed and badly conditioned. Instead, we should use stable transformations like the Schur decomposition.

Proof. When $n = 1$ the proof is immediate. So let $n \geq 2$. Consider the Schur triangular form (Horn & Johnson, 2012, Chapter 2.3) of A:

$$A = UDU^*,$$

where D is upper triangular, U is unitary, and $*$ denotes complex conjugate. Let Λ be the diagonal part of D, which contains all eigenvalues of A as elements. Notice that $D - \Lambda$ is upper triangular with zero diagonal elements, while Λ is diagonal. Thus, any product of the form

$$\Lambda^{t_0}(D-\Lambda)^{s_1}\Lambda^{t_1}\cdots(D-\Lambda)^{s_k}\Lambda^{t_k} = 0, \text{ if } s_1 + \cdots + s_k \geq n.$$

where $s_1, \ldots, s_k$ and $t_0, t_1, \ldots, t_k$ are two collections of integers, for some $k \geq 1$. Now we can simplify the expression:

$$\begin{aligned} D^k = (\Lambda + D - \Lambda)^k &= \sum_{d_1,\ldots,d_k \in \{0,1\}^k} F_{d_1} \cdots F_{d_k} \\ &= \sum_{\substack{d_1,\ldots,d_k \in \{0,1\}^k \\ d_1+\cdots+d_k \leq n-1}} F_{d_1} \cdots F_{d_k}, \end{aligned}$$

where $F_1 = D - \Lambda$, $F_0 = \Lambda$. Notice that $\|D - \Lambda\|_2 \leq \|D\|_2 = \|A\|_2 \leq M$, where the first inequality follows from the fact that $D - \Lambda$ is a submatrix if D. Since the eigenvalues of A are inside or on the unit circle, we have $\left\|\Lambda^t\right\|_2 \leq 1$, for all $t \geq 0$. Hence, by a counting argument

$$\begin{aligned} \left\|A^k\right\|_2 = \left\|D^k\right\| &\leq \sum_{t=0}^{n-1} \binom{k}{t} \max\left\{M^t, 1\right\} \\ &\leq \sum_{t=0}^{n-1} \binom{k}{t} \max\left\{M^{n-1}, 1\right\}. \end{aligned}$$

To conclude, we use the known bound (Vershynin, 2018, Exercise 0.0.5):

$$\sum_{t=0}^{n-1} \binom{k}{t} \leq \left(\frac{ek}{n-1}\right)^{n-1}$$

□

Since we obtained a bound on the powers of matrix A, we can immediately obtain an upper bound on the Gramian as a corollary.

Corollary 6.1. *Let $A \in \mathbb{R}^{n\times n}$ have all eigenvalues inside or on the unit circle, with $\|A\|_2 \leq M$. Let $H \in \mathbb{R}^{n\times r}$, $r \leq n$ with $\|H\|_2 \leq M$. Then, the gramian $\Gamma_k(A,H)$ is upper bounded by:*

$$\|\Gamma_k(A,H)\|_2 \leq e^{2n-2}k^{2n-1}\max\left\{M^{2n},1\right\} \tag{6.19}$$

6.10.2 Proof of Theorem 6.2

Let β be any non-zero number. Fix an accuracy parameter $\epsilon > 0$ and a confidence $0 < \delta < 1$. Consider the systems:

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 2\epsilon & 0 \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{bmatrix},$$
$$H_1 = H_2 = \begin{bmatrix} e_1 & e_3 \end{bmatrix}.$$

Both systems are controllable and belong to the class $\mathcal{C}_n$ for any non-zero $\beta \neq 0$. However, they are arbitrarily close to uncontrollability for small β. Let $f_{S_i}(x_0,\dots,x_N)$ denote the probability density function of the distribution of the data under system S_i, $i=1,2$. Then the log-likelihood ratio under S_1, S_2 is:

$$L_N = \log\frac{f_{S_1}(x_0,\dots,x_N)}{f_{S_2}(x_0,\dots,x_N)}.$$

Due to the Markovian structure of the linear system, we can write $f_{S_i}(x_0,\dots,x_N) = \prod_{k=1}^{N} f_{S_i}(x_k|x_{k-1})$, for $i=1,2$. Moreover, due to the structure of the dynamical systems:

$$f_{S_i}(x_k|x_{k-1}) = f_{S_i}(x_{k,1}|x_{k-1,2})f_{S_i}(x_{k,2},x_{k,3}|x_{k-1,2}x_{k-1,3}).$$

However, systems A_1, A_2 have identical distributions for $x_{k,2}$ and $x_{k,3}$. As a result, the log-likelihood ratio becomes:

$$L_N = \sum_{k=1}^{N} \log \frac{f_{S_1}(x_{k,1}|x_{k-1,2})}{f_{S_2}(x_{k,1}|x_{k-1,2})}.$$

The KL divergence can now be computed:

$$\begin{aligned}
\mathrm{KL}(\mathbb{P}_{S_1}, \mathbb{P}_{S_2}) &= \mathbb{E}_{S_1} L_N \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} \mathbb{E}_{S_1} \left(\log \frac{f_{S_1}(x_{k,1}|x_{k-1,2})}{f_{S_2}(x_{k,1}|x_{k-1,2})} |\mathcal{F}_{k-1} \right) \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} \mathrm{KL}(\mathcal{N}(0,1), \mathcal{N}(2\epsilon x_{k-1,2}, 1)) \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} (2\epsilon x_{k-1,2})^2/2 \leq 2\epsilon^2 N \Gamma_{N,22}(A, H),
\end{aligned}$$

where we used $\mathbb{E}_{S_1} x_{k-1,2}^2 = \Gamma_{k-1,22} \leq \Gamma_{N,22}$ along with the fact that the KL-divergence between two scalar Gaussians is:

$$\mathrm{KL}(\mathcal{N}(\mu_1, 1), \mathcal{N}(\mu_2, 1)) = (\mu_1 - \mu_2)^2/2$$

A simple computation shows that $\Gamma_{k,22} = b^2$, for all $k \geq 1$. Then, it follows from Lemma 6.2 that (6.2) holds only if:

$$N \geq \frac{1}{\beta^2 2\epsilon^2} \log \frac{1}{3\delta}.$$

However β is arbitrary, which implies that (6.2) holds only if:

$$N \geq \sup_{\beta \neq 0} \frac{1}{\beta^2 4\epsilon^2} \log \frac{1}{3\delta} = \infty.$$

6.10.3 Proof of Theorem 6.3

Consider system (6.9) with $\rho = 1/4$ and the perturbed system $\tilde{A} = A + 2\epsilon e_1 e_2'$, $\tilde{H} = H$, where we modify A_{12} by 2ϵ. Both pairs (A, H), $(\hat{A}, \hat{H})$ are controllable. From Lemma 6.4, we obtain that $d(A, H) \geq (4(n+1))^{-1} \geq (8(n+1))^{-1}$. Fix an $\epsilon \leq (16(n+1))^{-1}$. Then, from Lemma 6.5, we also get that $d(\hat{A}, \hat{H}) \geq d(A, H) - 2\epsilon \geq (8(n+1))^{-1}$. Hence, both systems belong to the class $\mathcal{C}_n$.

Define $S_1 = (A, H)$, $S_2 = (\hat{A}, \hat{H})$. Following the same arguments as in the proof of Theorem 6.2, the KL divergence of the distribution of the data under A and $\hat{A}$ is equal to

$$\begin{aligned}
\mathrm{KL}(\mathbb{P}_{S_1}, \mathbb{P}_{S_2}) &= \mathbb{E}_{S_1} L_N \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} \mathbb{E}_{S_1} \left(\log \frac{f_{S_1}(x_{k,1} | x_{k-1,2})}{f_{S_2}(x_{k,1} | x_{k-1,2})} | \mathcal{F}_{k-1} \right) \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} \mathrm{KL}(\mathcal{N}(\rho x_{k-1,2}, 1), \mathcal{N}((\rho + 2\epsilon) x_{k-1,2}, 1)) \\
&= \mathbb{E}_{S_1} \sum_{k=1}^{N} (2\epsilon x_{k-1,2})^2 / 2 \leq 2\epsilon^2 N \Gamma_{N,22}(A, H).
\end{aligned}$$

From Lemma 6.6, we obtain the exponential decay bound:

$$\Gamma_{N,22}(A, H) \leq 4^{-n+2}/3.$$

Finally, from Lemma 6.2, equation (6.2) holds only if:

$$N \geq \frac{1}{2\epsilon^2 \Gamma_{N,22}(A, H)} \log \frac{1}{3\delta} \geq \frac{4^{n-2}}{6\epsilon^2} \log \frac{1}{3\delta}.$$

Lemma 6.6. *Consider system (6.9) with $\rho < 1/2$. Then*

$$\Gamma_{k,22}(A, H) \leq (2\rho)^{2n-2} / (1 - 4\rho^2).$$

Proof. Notice that $e_2' A^s H = 0$ for all $s \le n-2$ and $\|A\| \le 2\rho < 1$. Hence,

$$e_2' \Gamma_k(A,H) e_2 \le \sum_{s=n-1}^{k} e_2 A^s Q A'^s e_2'$$
$$\le \sum_{s=n-1}^{\infty} (2\rho)^{2s} = (2\rho)^{2n-2}/(1-4\rho^2).$$

□

6.10.4 Proof of Theorem 6.4

By $\Gamma_k = \Gamma_k(A,H) + \Gamma_k(A,B)$ we denote the Gramian under both H, B. Define also the sigma-algebra:

$$\bar{\mathcal{F}}_k = \sigma(w_0, u_0, \dots, w_k, u_k).$$

We will apply Theorem 2.4 in Simchowitz et al. (2018) to the combined state-input vectors with three modifications since the noise is not isotropic. First, we compute the sub-Gaussian parameter of the noise.

Definition 5. *A zero mean random vector $w \in \mathbb{R}^{r \times 1}$ is called σ^2-sub-Gaussian with respect to a sigma algebra $\mathcal{F}$ if for every unit vector $u \in \mathbb{R}^{r\times}$:*

$$\mathbb{E}\left(e^{su'w} | \mathcal{F}\right) \le e^{s^2\sigma^2/2}.$$

From the definition, it follows that the non-isotropic Gaussian vector Hw_k is sub-Gaussian with parameter $\|H\|_2^2$.

Lemma 6.7. *Let $w_k \in \mathbb{R}^{r\times 1}$ be 1-sub-Gaussian with respect to $\bar{\mathcal{F}}_{k-1}$. Then Hw_k is $\|H\|_2^2-$sub-Gaussian with respect to $\bar{\mathcal{F}}_{k-1}$.*

Proof. Let $u \in \mathbb{R}^{r\times 1}$ be a unit vector. Then:

$$\mathbb{E}\left(e^{su'Hw_k} | \bar{\mathcal{F}}_{k-1}\right) = \mathbb{E}\left(e^{s\|u'H\| \frac{u'H}{\|u'H\|} w_k} | \bar{\mathcal{F}}_{k-1}\right)$$
$$\le e^{s^2\|u'H\|_2^2/2} \le e^{s^2\|H\|_2^2/2}$$

□

Second, define $y_k = \begin{bmatrix} x_k' & u_k' \end{bmatrix}'$. It follows that for all $j \geq 0$ and all unit vectors $v \in \mathbb{R}^{(n+p)\times 1}$, the following small-ball condition is satisfied:

$$\frac{1}{2\kappa}\sum_{t=0}^{2\kappa}\mathbb{P}(\left|v'y_{t+j}\right| \geq \sqrt{v'\Gamma_{\text{sb}}v}|\bar{\mathcal{F}}_j) \geq \frac{3}{20}, \tag{6.20}$$

where

$$\Gamma_{\text{sb}} = \begin{bmatrix} \Gamma_\kappa & 0 \\ 0 & I_p \end{bmatrix}. \tag{6.21}$$

Equation (6.20) follows from the same steps as in Proposition 3.1 in Simchowitz et al. (2018) with the choice $k = 2\kappa$.

Finally, we determine an upper bound $\bar{\Gamma}$ for the gram matrix $\sum_{t=0}^{N-1} y_t y_t'$. Using a Markov inequality argument as in (Simchowitz et al., 2018, proof of Th 2.1), we obtain that

$$\mathbb{P}(\sum_{t=0}^{N-1} y_t y_t' \preceq \bar{\Gamma}) \geq 1-\delta,$$

where

$$\bar{\Gamma} = \frac{n+p}{\delta}N\begin{bmatrix} \Gamma_N & 0 \\ 0 & I_p \end{bmatrix}$$

Now we can apply Theorem 4.2 of Simchowitz et al. (2018). With probability at least $1-3\delta$ we have $\|A - \hat{A}_N\| \leq \epsilon$ if:

$$N \geq \frac{\text{poly}(n, \log 1/\delta, M)}{\epsilon^2\sigma_{\min}(\Gamma_\kappa)}\log\det(\bar{\Gamma}\Gamma_\kappa^{-1}),$$

where we have simplified the expression by including terms in the polynomial term. Based on Lemma 6.1 and Theorem 6.5, we can bound the right-hand side:

$$\begin{aligned}\frac{\mathrm{poly}(n,\log 1/\delta,M)}{\epsilon^2\sigma_{\min}(\Gamma_\kappa)}\log\det(\bar{\Gamma}\Gamma_\kappa^{-1}) &\le \mathrm{poly}(n,\epsilon^{-1},\log 1/\delta,M)\mathrm{poly}\left(\frac{M}{\mu}\right)^{\kappa}\log N\\ &\le \mathrm{poly}(n,\epsilon^{-1},\log 1/\delta,M)\mathrm{poly}\left(M,n\right)^{\kappa}\log N,\end{aligned}$$

where we used the fact that $\mu^{-1}\le \mathrm{poly}(n)$. Hence, it is sufficient to have:

$$N\ge \mathrm{poly}\left(n,\epsilon^{-1},\log 1/\delta,M\right)\mathrm{poly}\left(M,n\right)^{\kappa}\log N.$$

To obtain the final polynomial bound, we need to remove the logarithm of N. It is sufficient to apply the inequality:

$$N\ge c\log N \text{ if } N\ge 2c\log 2c,$$

for $c>0$ which follows from elementary calculus.

6.10.5 Proof of Theorem 6.5

Our goal is to upper bound the norm of the pseudo-inverse $\|\mathcal{C}_\kappa^\dagger\|=\sqrt{\sigma_{\min}(\Gamma_\kappa)}$, where the equality follows from the SVD decomposition and the definition of the gramian. Towards proving the result, we will work with the staircase form (6.13). First, we show that if the system is μ-away from uncontrollability, then the subdiagonal matrices in the staircase form are bounded away from zero.

Lemma 6.8 (Staircase form lower bound)**.** *Let $(A,H)\in\mathbb{R}^{n\times(n+r)}$ be controllable and let Assumption 6.3 hold. Consider the staircase form of (A,H), with $A_{i+1,i}$ the subdiagonal matrices, for $i=1,\dots,\kappa-1$, where κ is the controllability index. Then, we have $A_{i+1,i}A'_{i+1,i}\succeq\mu^2 I_{r_{i+1}}$ for all $i=1,\dots,\kappa-1$. Moreover, $H_1H_1'\succeq\mu^2 I_r$.*

Proof. Let $(\hat{A},\hat{H})$ be the staircase form of (A,H) under the unitary similarity transformation U. First, we show that the distance from losing index κ is invariant to unitary transformations. Denote $\Delta\hat{A} = U^*\Delta A U$, $\Delta\hat{H} = U^*\Delta H$. Then:

$$\begin{aligned}
&\min\left\{\| \begin{bmatrix} \Delta A & \Delta H \end{bmatrix} \|_2 :\ \kappa(A+\Delta A, H+\Delta H) > \kappa\right\} \\
&= \min\left\{\| \begin{bmatrix} \Delta\hat{A} & \Delta\hat{H} \end{bmatrix} \|_2 :\ \kappa(A+\Delta A, H+\Delta H) > \kappa\right\} \\
&= \min\left\{\| \begin{bmatrix} \Delta\hat{A} & \Delta\hat{H} \end{bmatrix} \|_2 :\ \kappa(\hat{A}+\Delta\hat{A}, \hat{H}+\Delta\hat{H}) > \kappa\right\}
\end{aligned}$$

where the first equality follows from $\| \begin{bmatrix} \Delta A & \Delta H \end{bmatrix} \|_2 = \| \begin{bmatrix} U^*\Delta A U & U^*\Delta H \end{bmatrix} \|_2$. The second equality follows from the fact that controllability index is preserved under similarity transformations As a result, $d_\kappa(\hat{A}, \hat{H}) = d_\kappa(A, H) \geq \mu$.

Note that $A_{i+1,i} \in \mathbb{R}^{r_{i+1}\times r_i}$. Hence, it is sufficient to show that $\sigma_{r_{i+1}}(A_{i+1,i}) \geq \mu$, where $\sigma_{r_{i+1}}$ denotes the r_{i+1} smallest singular value. Assume that the opposite is true $\sigma_{r_{i+1}}(A_{i+1,i}) < \mu$. We will show that this contradicts the fact that $(\hat{A}, \hat{H})$ is μ-away from losing controllability index. Let u and v be the singular vectors in the Singular Value Decomposition of $A_{i+1,i}$ corresponding to $\sigma_{r_{i+1}}$. Let $\Delta A_{i+1,i} \triangleq -\sigma_{r_{i+1}}(A_{i+1,i})uv'$. Then $A_{i+1,i} + \Delta A_{i+1,i}$ is rank deficient. Now let $\Delta\hat{A}$ be zero everywhere apart from the block $\Delta A_{i+1,i}$. Then, we have that $(\hat{A}+\Delta\hat{A}, \hat{H})$ at time κ has singular controllability matrix:

$$\operatorname{rank}(\mathcal{C}_\kappa(\hat{A}+\Delta\hat{A}, \hat{H})) < n.$$

In other words,

$$\kappa(\hat{A}+\Delta\hat{A}, \hat{H}) > \kappa$$

and we lose controllability index, with $\|\Delta\hat{A}\|_2 < \mu \leq d_\kappa(\hat{A}, \hat{H})$, which is impossible. The proof for H_1 is similar. □

The above result allows us to work with the staircase form (6.13), which has a nice triangular structure. In fact the controllability matrix is block-triangular and we can upper-bound its least singular value using a simple recursive bound. Since the least singular value of the Gramian is invariant to similarity transformations, we will now assume that the system (A,H) is now already in form (6.13) with $U = I$. Let us define some auxiliary matrices that will help us prove Theorem 6.5. With $\tilde{A}_k$, for $k \leq \kappa$ we denote the submatrix of A when we keep the k-upper left block matrices in (6.13) and we delete the remaining columns and rows, e.g.:

$$\tilde{A}_2 = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}, \tilde{A}_3 = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ 0 & A_{3,2} & A_{3,3} \end{bmatrix}, \dots$$

Similarly, we define the submatrices $\tilde{H}_k$ where we keep only the upper k blocks of the matrix H:

$$\tilde{H}_1 = H_1,\, \tilde{H}_2 = \begin{bmatrix} H_1' & 0 \end{bmatrix}', \dots.$$

Finally, define the upper-left controllability submatrices $\tilde{\mathcal{C}}_k$:

$$\tilde{\mathcal{C}}_k = \begin{bmatrix} \tilde{H}_k & \tilde{A}_k \tilde{H}_k & \dots & \tilde{A}_k^{k-1} \tilde{H}_k \end{bmatrix} \in \mathbb{R}^{\sum_{i=1}^{k} r_i \times (kr)}. \tag{6.22}$$

The benefit of working with the above matrices is that they are block upper-triangular. For example:

$$\tilde{\mathcal{C}}_1 = H_1,\, \tilde{\mathcal{C}}_2 = \begin{bmatrix} H_1 & A_{1,1} H_1 \\ 0 & A_{2,1} H_1 \end{bmatrix}, \dots$$

By definition $\tilde{A}_\kappa = A$, $\tilde{H}_\kappa = H$, and $\tilde{\mathcal{C}}_\kappa = \mathcal{C}_\kappa$.

Lemma 6.9 (Recursive definition of right-inverse)**.** *Assume the pair (A, K) is in the canonical representation (6.13) with $U = I$. Let $\tilde{\mathcal{C}}_k$ be the upper-left part of the controllability matrix as defined in (6.22), with $k \leq \kappa$, where κ is the controllability index. Let $\Pi_k = H_1^{-1} A_{2,1}^{\dagger} A_{3,2}^{\dagger} \cdots A_{k,k-1}^{\dagger}$, where $\dagger$ denotes the Moore-Penrose pseudo-inverse. Then, the following inequality holds recursively:*

$$\|\tilde{\mathcal{C}}_k^{\dagger}\|_2 \leq \|\tilde{\mathcal{C}}_{k-1}^{\dagger}\|_2 + \|\Pi_k\|_2 + \|\tilde{\mathcal{C}}_{k-1}^{\dagger} \tilde{A}_{k-1}^{k-1} \tilde{H}_{k-1} \Pi_k\|_2. \tag{6.23}$$

Proof. The upper-left controllability matrix $\tilde{\mathcal{C}}_k$, $k \leq \kappa$ has the following block triangular structure:

$$\begin{aligned}\tilde{\mathcal{C}}_k &= \left[\begin{array}{ccc|c} \tilde{H}_k & \dots & \tilde{A}_k^{k-1}\tilde{H}_k & \tilde{A}_k^{k-1}\tilde{H}_k \end{array}\right] \\ &= \left[\begin{array}{c|c} \tilde{\mathcal{C}}_{k-1} & \tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1} \\ 0 & A_{k,k-1}A_{k-1,k-2}\dots H_1 \end{array}\right].\end{aligned} \tag{6.24}$$

Based on the above form, we can construct a right-inverse of matrix $\tilde{\mathcal{C}}_k$:

$$\tilde{\mathcal{C}}_k^\sharp \triangleq \begin{bmatrix} \tilde{\mathcal{C}}_{k-1}^\dagger & -\tilde{\mathcal{C}}_{k-1}^\dagger \tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k \\ 0 & \Pi_k \end{bmatrix},$$

which satisfies $\tilde{\mathcal{C}}_k\tilde{\mathcal{C}}_k^\sharp = I$. By the definition of $\tilde{\mathcal{C}}_k^\sharp$:

$$\|\tilde{\mathcal{C}}_k^\sharp\|_2 \leq \|\tilde{\mathcal{C}}_{k-1}^\dagger\|_2 + \|\Pi_k\|_2 + \|\tilde{\mathcal{C}}_{k-1}^\dagger\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k\|_2.$$

To conclude the proof, we invoke Lemma 6.10. □

Lemma 6.10. *Let $M \in \mathbb{R}^{s\times t}$ be any matrix with full column rank $s \leq t$. Let $M^\sharp$ be any right inverse of M, i.e. $MM^\sharp = I_s$. Then the following inequality is true:*

$$\|M^\dagger\|_2 \leq \|M^\sharp\|_2,$$

where $M^\dagger$ is the Moore Penrose pseudo-inverse.

Proof. Notice that $M(M^\dagger - M^\sharp) = 0$. As a result, we can write $M^\sharp = M^\dagger + M_{\text{null}}$, where M_{null} is any matrix in the null space $MM_{\text{null}} = 0$. However, the Moorse-Penrose pseudoinverse and M_{null} are orthogonal

$$(M^\dagger)'M_{\text{null}} = 0.$$

By orthogonality, for every $x \in R^{t\times 1}$ we have $\|M^\sharp x\|_2 = \sqrt{\|M^\dagger x\|^2 + \|M_{\text{null}}x\|^2} \geq \|M^\dagger x\|^2$.

□

Since all coupling matrices $A_{k,k-1}, \dots, A_{2,1}, H_1$ have least singular value lower bounded by μ, the product of their pseudo-inverses is upper bounded by:

$$\|\Pi_k\| \leq \mu^{-k}.$$

So, we should expect (6.23) to grow no faster than exponentially with κ. However, the main challenge is to control the last term in (6.23). Unless we follow a careful analysis, if we just apply the submultiplicative property of the norm we will get bounds which are exponential with κ^2 instead of κ. The idea is the following. Since by definition $\tilde{C}_{k-1}$ has full rank, then there exists an appropriate matrix $\Lambda_{k-1} \in \mathbb{R}^{(k-1)r \times r_k}$ such that

$$\tilde{A}_{k-1}^{k-1} \tilde{H}_{k-1} \Pi_k = \tilde{C}_{k-1} \Lambda_{k-1}.$$

Then the above bound becomes:

$$\|\tilde{\mathcal{C}}_k^\dagger\| \leq \|\tilde{\mathcal{C}}_{k-1}^\dagger\| + \mu^{-k} + \|\Lambda_{k-1}\|, \tag{6.25}$$

where we used the fact that $\|\tilde{\mathcal{C}}_{k-1}^\dagger \tilde{\mathcal{C}}_{k-1}\| \leq 1$. For the remaining proof, we need to construct such a matrix Λ_{k-1} and upper bound it.

Lemma 6.11. *Let $\Lambda_{k-2} \in \mathbb{R}^{(k-2)r \times r_{k-1}}$ be any matrix such that:*

$$\tilde{A}_{k-2}^{k-2} \tilde{H}_{k-2} \Pi_{k-1} = \tilde{\mathcal{C}}_{k-2} \Lambda_{k-2}$$

There exists a matrix $\Lambda_{k-1} \in \mathbb{R}^{(k-1)r \times r_k}$ such that:

$$\tilde{A}_{k-1}^{k-1} \tilde{H}_{k-1} \Pi_k = \tilde{\mathcal{C}}_{k-1} \Lambda_{k-1}$$

with

$$\|\Lambda_{k-1}\|_2 \leq \frac{2+M}{\mu}\,\|\Lambda_{k-2}\|_2 + \frac{M}{\mu}\|\tilde{\mathcal{C}}_{k-2}^{\dagger}\|_2 + \mu^{-k}M. \tag{6.26}$$

Proof. **Part A: algebraic expression for $\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k$.** Observe that every matrix $\tilde{A}_{k-1}$ includes the previous as an upper-left submatrix:

$$\tilde{A}_{k-1} = \begin{bmatrix} \tilde{A}_{k-2} & A_{1:k-2,k-1} \\ A_{k-1,1:k-2} & A_{k-1,k-1} \end{bmatrix},$$

with

$$A_{1:k-1,k-1} = \begin{bmatrix} A_{1,k-1} \\ \vdots \\ A_{k-2,k-1} \end{bmatrix},\ A_{k-1,1:k-2} = \begin{bmatrix} 0 & \cdots & 0 & A_{k-1,k-2} \end{bmatrix}$$

Let also:

$$Q_k = A_{k,k-1}A_{k-1,k-2}\cdots H_1.$$

A direct computation gives:

$$\tilde{\mathcal{C}}_k = \begin{bmatrix} \tilde{\mathcal{C}}_{k-2} & \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} & \tilde{A}_{k-2}^{k-1}\tilde{H}_{k-2} + A_{1:k-2,k-1}Q_{k-1} \\ 0 & Q_{k-1} & A_{k-1,1:k-2}\tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} + A_{k-1,k-1}Q_{k-1} \\ 0 & 0 & Q_k \end{bmatrix}. \tag{6.27}$$

As a result of (6.24) and (6.27),

$$\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k = \begin{bmatrix} \tilde{A}_{k-2}^{k-1}\tilde{H}_{k-2} + A_{1:k-2,k-1}Q_{k-1} \\ A_{k-1,1:k-2}\tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} + A_{k-1,k-1}Q_{k-1} \end{bmatrix}\Pi_k.$$

We can simplify the above expression using $Q_{k-1}\Pi_{k-1} = I$ and $\tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Pi_{k-1} = \tilde{\mathcal{C}}_{k-2}\Lambda_{k-2}$:

$$\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k = \begin{bmatrix} \tilde{A}_{k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{1:k-2,k-1} \\ A_{k-1,1:k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{k-1,k-1} \end{bmatrix}\tilde{A}_{k,k-1}^{\dagger}. \tag{6.28}$$

Part B: last rows as linear combination.

Our goal is to express (6.28) as a linear combination of the columns of:

$$\tilde{\mathcal{C}}_{k-1} = \begin{bmatrix} \tilde{\mathcal{C}}_{k-2} & \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} \\ 0 & Q_{k-1} \end{bmatrix}.$$

Since $\tilde{\mathcal{C}}_{k-1}$ has a triangular structure, we start from the last r_{k-1} rows of $\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k$ Exploiting the structure of $A_{k-1,1:k-2}$, which includes many zeros we can write:

$$\begin{aligned}
&A_{k-1,1:k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{k-1,k-1} \\
&= \begin{bmatrix} 0 & \cdots & 0 & A_{k-1,k-2} \end{bmatrix} \left[\begin{array}{c|c} \tilde{\mathcal{C}}_{k-3} & \tilde{A}_{k-3}^{k-3}\tilde{H}_{k-1} \\ 0 & A_{k-2,k-3}A_{k-3,k-4}\dots H_1 \end{array} \right] \Lambda_{k-2} + A_{k-1,k-1} \\
&= A_{k-1,k-2}Q_{k-2}\Lambda_{k-2,k-2} + A_{k-1,k-1} \\
&= Q_{k-1}\Lambda_{k-2,k-2} + A_{k-1,k-1},
\end{aligned}$$

where $\Lambda_{k-2,k-2} \in \mathbb{R}^{r \times r_{k-1}}$ are the last r rows of matrix Λ_{k-2}:

$$\Lambda_{k-2} = \begin{bmatrix} \Lambda_{k-2,1} \\ \vdots \\ \Lambda_{k-2,k-2} \end{bmatrix}.$$

Finally, we car rewrite the last r_{k-1} rows of $\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k$ as:

$$(A_{k-1,1:k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{k-1,k-1})\tilde{A}_{k,k-1}^{\dagger} = Q_{k-1}(\Lambda_{k-2,k-2} + \Pi_{k-1}A_{k-1,k-1})\tilde{A}_{k,k-1}^{\dagger} \quad (6.29)$$

Part c: remaining rows.

From (6.29), we can eliminate the last rows:

$$\tilde{A}_{k-1}^{k-1}\tilde{H}_{k-1}\Pi_k - \begin{bmatrix} \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} \\ Q_{k-1} \end{bmatrix} (\Lambda_{k-2,k-2} + \Pi_{k-1}A_{k-1,k-1})\tilde{A}_{k,k-1}^{\dagger} =$$

$$\begin{bmatrix} \tilde{A}_{k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{1:k-2,k-1} - \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Lambda_{k-2,k-2} - \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Pi_{k-1}A_{k-1,k-1} \\ 0 \end{bmatrix} \tilde{A}_{k,k-1}^{\dagger}$$

$$= \begin{bmatrix} \tilde{A}_{k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} + A_{1:k-2,k-1} - \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Lambda_{k-2,k-2} - \tilde{\mathcal{C}}_{k-2}\Lambda_{k-2}A_{k-1,k-1} \\ 0 \end{bmatrix} \tilde{A}_{k,k-1}^{\dagger}$$

Notice that by the shift structure of the controllability matrix:

$$\begin{aligned} &\tilde{A}_{k-2}\tilde{\mathcal{C}}_{k-2}\Lambda_{k-2} - \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Lambda_{k-2,k-2} \\ &= \begin{bmatrix} \tilde{A}_{k-2}\tilde{H}_{k-2} & \ldots & \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2} \end{bmatrix} \Lambda_{k-2} - \tilde{A}_{k-2}^{k-2}\tilde{H}_{k-2}\Lambda_{k-2,k-2} \\ &= \begin{bmatrix} \tilde{A}_{k-2}\tilde{H}_{k-2} & \ldots & \tilde{A}_{k-2}^{k-3}\tilde{H}_{k-2} & 0 \end{bmatrix} \Lambda_{k-2} \\ &= \begin{bmatrix} \tilde{H}_{k-2} & \tilde{A}_{k-2}\tilde{H}_{k-2} & \ldots & \tilde{A}_{k-2}^{k-3}\tilde{H}_{k-2} \end{bmatrix} \Lambda_{k-2}^{\text{shift}} \\ &= \tilde{\mathcal{C}}_{k-2}\Lambda_{k-2}^{\text{shift}}. \end{aligned}$$

where

$$\Lambda_{k-2}^{\text{shift}} = \begin{bmatrix} 0 \\ \Lambda_{k-2,1} \\ \vdots \\ \Lambda_{k-2,k-3} \end{bmatrix}.$$

Moreover, we can write $A_{1:k-2,k-1} = \tilde{\mathcal{C}}_{k-2}\tilde{\mathcal{C}}_{k-2}^{\dagger}A_{1:k-2,k-1}$

Part d: construction of Λ_{k-1}.

Combining the above equalities:

$$\tilde{A}^{k-1}_{k-1}\tilde{H}_{k-1}\Pi_k = \begin{bmatrix} \tilde{A}^{k-2}_{k-2}\tilde{H}_{k-2} \\ Q_{k-1} \end{bmatrix} (\Lambda_{k-2,k-2} + \Pi_{k-1}A_{k-1,k-1})\tilde{A}^{\dagger}_{k,k-1}$$

$$+ \begin{bmatrix} \tilde{\mathcal{C}}_{k-2} \\ 0 \end{bmatrix} (\Lambda^{\text{shift}}_{k-2} + \tilde{\mathcal{C}}^{\dagger}_{k-2}A_{1:k-2,k-1} - \Lambda_{k-2}A_{k-1,k-1})\tilde{A}^{\dagger}_{k,k-1}.$$

Hence we can select:

$$\Lambda_{k-1} = \begin{bmatrix} \left(\Lambda^{\text{shift}}_{k-2} + \tilde{\mathcal{C}}^{\dagger}_{k-2}A_{1:k-2,k-1} - \Lambda_{k-2}A_{k-1,k-1}\right)\tilde{A}^{\dagger}_{k,k-1} \\ \left(\Lambda_{k-2,k-2} + \Pi_{k-1}A_{k-1,k-1}\right)\tilde{A}^{\dagger}_{k,k-1} \end{bmatrix},$$

with

$$\|\Lambda_{k-1}\| \leq (2+M)\mu^{-1}\|\Lambda_{k-2}\| + M\mu^{-1}\|\tilde{\mathcal{C}}^{\dagger}_{k-2}\| + \mu^{-k}M$$

□

Now we can complete the proof of Theorem 6.5. It is sufficient to select Λ_1:

$$A_{1,1}H_1\Pi_2 = H_1H_1^{-1}A_{1,1}A^{\dagger}_{2,1} = \tilde{\mathcal{C}}_1\Lambda_1,$$

with $\|\Lambda_1\|_2 \leq M\mu^{-2}$. Let $\alpha_k = \begin{bmatrix} \|\tilde{\mathcal{C}}^{\dagger}_k\| & \|\Lambda_k\| & \mu^{-k} \end{bmatrix}'$. From (6.25), (6.26) we obtain the following recursion:

$$\alpha_k \leq \begin{bmatrix} 1 & 1 & \mu^{-1} \\ \frac{M}{\mu} & \frac{2+M}{\mu} & \frac{M}{\mu} \\ 0 & 0 & \mu^{-1} \end{bmatrix} \alpha_{k-1},$$

where the inequality is interpreted coordinate-wise. Let Ξ be the matrix of the above recursion. We have the crude bound:

$$\|\mathcal{C}^{\dagger}_{\kappa}\|_2 = \|\tilde{\mathcal{C}}^{\dagger}_{\kappa}\|_2 \leq \|\Xi^{\kappa-1}\|_2\|\alpha_1\|_2,$$

where $\|\Xi^\kappa\|_2\|\alpha_1\|_2 \leq \mathrm{poly}^\kappa(M/\mu)$. This completes the proof.

Chapter 7

Difficulty of Learning to Control Linear Systems

7.1 Introduction

Let us consider again linear systems of the form

$$S: \qquad x_{k+1} = Ax_k + Bu_k + Hw_k, \tag{7.1}$$

where $x_k \in \mathbb{R}^n$ is the system state, $u_k \in \mathbb{R}^p$ is some exogenous input, and $w_k \in \mathbb{R}^r$ is some random disturbance sequence. Control theory has a long history of studying how to design controllers for system (7.1) when its model is *known* (Bertsekas, 2017). However, in reality system (7.1) might be *unknown* and we might not have access to its model. In this case, we have to learn how to control (7.1) based on data. In the previous chapter we only focused on the problem of recovering the model of (7.1). In this chapter, we will focus on the problem of learning to control (7.1).

Controlling unknown dynamical systems has also been studied from the perspective of Reinforcement Learning (RL). Although the setting of tabular RL is relatively well-understood (Jaksch et al., 2010), it has been challenging to analyze the continuous setting, where the state and/or action spaces are infinite (Ortner & Ryabko, 2012; Kakade et al., 2020). Recently, there has been renewed interest in learning to control linear systems. Indeed, linear systems are simple enough to allow for an in-depth theoretical analysis, yet exhibit sufficiently rich behavior so that we can draw conclusions about continuous control of more general system classes (Recht, 2019). Here, we focus on the following two problems.

Regret of online LQR. A fundamental benchmark for continuous control is the Linear Quadratic Regulator (LQR) problem, where the goal is to compute a policy π that minimizes

$$J^*(S) \triangleq \min_{\pi} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}_{S,\pi}\left[\sum_{t=0}^{T-1}(x_t'Qx_t + u_t'Ru_t) + x_T'Q_Tx_T\right]. \tag{7.2}$$

When model (7.1) is known, LQR enjoys a closed-form solution; the optimal policy is a linear feedback law $\pi_{\star,t}(x_t) = K_\star x_t$, where the control gain $K_\star$ is given by solving the celebrated Algebraic Riccati Equation (7.7). If model (7.1) is unknown, we have to learn the optimal policy from data. In the online learning setting, the goal of the learner is to find a policy that adapts online and competes with the optimal LQR policy that has access to the true model. The suboptimality of the online learning policy at time T is captured by the *regret*

$$R_T(S) \triangleq \sum_{t=0}^{T-1}(x_t'Qx_t + u_t'Ru_t) + x_T'Q_Tx_T - TJ^*(S). \tag{7.3}$$

The learning task is to find a policy with as small regret as possible.

Sample Complexity of Stabilization Another important benchmark is the problem of stabilization from data. The goal is to learn a linear gain $K \in \mathbb{R}^{m\times n}$ such that the closed-loop system $A + BK$ is stable, i.e., such that its spectral radius $\rho(A + BK)$ is less than one. Many algorithms for online LQR require the existence of such a stabilizing gain to initialize the online learning policy (Simchowitz & Foster, 2020; Jedra & Proutiere, 2021). Furthermore, stabilization is a problem of independent interest (Faradonbeh et al., 2018b). In this setting, the learner designs an exploration policy π and an algorithm that uses batch state-input data $x_0, \dots, x_N, u_0, \dots, u_{N-1}$ to output a control gain $\hat{K}_N$, at the end of the exploration phase. Here we focus on *sample complexity*, i.e., the minimum number of samples N required to find a stabilizing gain.

Since the seminal papers by Abbasi-Yadkori & Szepesvári (2011) and Dean et al. (2017) both LQR and stabilization have been studied extensively in the literature – see Section 7.1.1. Current state-of-the-art results state that the regret of online LQR and the sample complexity of stabilization scale at most polynomially with system dimension n

$$R_T(S) \lesssim C_1^{\text{sys}} \text{poly}(n)\sqrt{T}, \quad N \lesssim C_2^{\text{sys}} \text{poly}(n), \tag{7.4}$$

where C_1^{sys}, C_2^{sys} are system specific constants that depend on several control theoretic quantities of system (7.1). However, the above statements might not reveal the whole picture.

In fact, system theoretic parameters C_1^{sys}, C_2^{sys} can actually hide dimensional dependence on n. This dependence has been overlooked in prior work. As we show in this chapter, there exist non-trivial classes of linear systems for which system theoretic parameters scale dramatically, i.e. exponentially, with the dimension n. As a result, the system theoretic quantities C_1^{sys}, C_2^{sys} might be very large and in fact *dominate* the poly(n) term in the upper bounds (7.4). This phenomenon especially arises in systems which are structurally difficult to control, such as for example underactuated systems. Then, the upper bounds (7.4) suggest that learning might be difficult for such instances. This brings up the following questions. *Can learning LQR or stabilizing controllers indeed be hard for such systems? How does system structure affect difficulty of learning?*

To answer the first question, we need to establish lower bounds. As we discuss in Section 7.1.1, existing lower bounds for online LQR (Simchowitz & Foster, 2020) might not always reveal the dependence on control theoretic parameters. Chen & Hazan (2021) provided exponential lower bounds for the start-up regret of stabilization. Still, to the best of our knowledge, there are no existing lower bounds for the *sample complexity* of stabilization. Recently, it was shown that the sample complexity of system identification can grow exponentially with the dimension n (Tsiamis & Pappas, 2021). However, it is not clear if difficulty of identification translates into difficulty of control. Besides, we do not always need to identify the whole system in order to control it (Gevers, 2005). To answer the second question, we need to provide upper bounds for several control theoretic parameters. Our contributions are the following:

Exp(n) Stabilization Lower Bounds. We prove an information-theoretic lower bound for the problem of learning stabilizing controllers, showing that it can indeed be statistically hard for underactuated systems. In particular, we show that the sample complexity of stabilizing an unknown underactuated linear system can scale exponentially with the state dimension n. To the best of our knowledge this is the first work to address this issue and consider lower bounds in this setting.

Exp(n) LQR Regret Lower Bounds. We show that the regret of online LQR can scale exponentially with the dimension as $\exp(n)\sqrt{T}$. In fact, even common integrator-like systems can exhibit this behavior. To prove our result, we leverage recent regret lower bounds (Ziemann & Sandberg, 2022), which provide a more refined analysis linking regret to system theoretic parameters. Chen & Hazan (2021) first showed that the start-up cost of the regret (terms of low order) can scale exponentially with n. Here, we show that this exponential dependence can also affect multiplicatively the dominant $\sqrt{T}$ term.

Exponential Upper Bounds. Under some additional structural assumptions (bounding systems away from uncontrollability), we provide matching global upper bounds. We show that the sample complexity of stabilization and the regret of online LQR can be at most exponential with the dimension n. In fact, we prove a stronger result, that they can be at most exponential with the *controllability index* of the system, which captures the structural difficulty of control – see Section 7.3. This implies that if the controllability index is small with respect to the dimension n, then learning is guaranteed to be easy.

7.1.1 Related Work

System Identification. A related problem is that of system identification, where the learning objective is to recover the model parameters A, B, H from data (Matni & Tu, 2019). The sample complexity of system identification was studied extensively in the setting of fully observed linear systems (Dean et al., 2017; Simchowitz et al., 2018; Faradonbeh et al., 2018a; Sarkar & Rakhlin, 2018; Fattahi et al., 2019; Jedra & Proutiere, 2019; Wagenmaker & Jamieson, 2020; Efroni et al., 2021) as well as partially-observed systems (Oymak & Ozay, 2018; Sarkar et al., 2019; Simchowitz et al., 2019; Tsiamis & Pappas, 2019; Lee & Lamperski, 2020; Zheng & Li, 2020; Lee, 2020; Lale et al., 2020b). Recently, it was shown that the sample complexity of system identification can grow exponentially with the dimension n (Tsiamis & Pappas, 2021).

Learning Feedback Laws. The problem of learning stabilizing feedback laws from data was studied before in the case of stochastic (Dean et al., 2017; Tu et al., 2017; Faradonbeh et al., 2018b; Mania et al., 2019) as well as adversarial (Chen & Hazan, 2021) disturbances. The standard paradigm has been to perform system identification, followed by a robust control or certainty equivalent gain design. Prior work is limited to sample complexity upper bounds. To the best of our knowledge, there have been no sample complexity lower bounds.

Online LQR. While adaptive control in the LQR framework has a rich history (Matni et al., 2019), the recent line of work on regret minimization in online LQR begins with Abbasi-Yadkori & Szepesvári (2011). They provide a computationally intractable algorithm based on optimism attaining $O(\sqrt{T})$ regret. Algorithms based on optimism have since been improved and made more tractable (Ouyang et al., 2017a; Abeille & Lazaric, 2018; Abbasi-Yadkori et al., 2019; Cohen et al., 2019; Abeille & Lazaric, 2020). In a closely related line of work, Dean et al. (2018) provide an $O(T^{2/3})$ regret bound for robust adaptive LQR control, drawing inspiration from classical methods in system identification and robust adaptive control. It has since been shown that certainty equivalent control, without robustness, can attain the (locally) minimax optimal $O(\sqrt{T})$ regret (Mania et al., 2019; Faradonbeh et al., 2020a; Lale et al., 2020a; Jedra & Proutiere, 2021). In particular, by providing nearly matching upper and lower bounds, Simchowitz & Foster (2020) refine this analysis and establish that the optimal rate, without taking system theoretic quantities into account, is $R_T = \Theta(\sqrt{p^2 nT})$. In this work, we rely on the lower bounds by Ziemann & Sandberg (2022), which provide a refined instance specific analysis and also lower bounds for the partially observed setting. Since their bounds exhibit more direct dependencies on control-theoretic parameters, we may use them here to show that certain non-local minimax complexities can be far worse than $R_T = \Omega(\sqrt{p^2 nT})$ and scale exponentially in the problem dimension. Indeed, an exponential start-up cost has already been observed by Chen & Hazan (2021). Here we show that this exponential dependency can persist multiplicatively even for large T.

7.1.2 Notation

The transpose of X is denoted by X'. For vectors $v \in \mathbb{R}^d$, $\|v\|_2$ denotes the ℓ_2-norm. For matrices $X \in \mathbb{R}^{d_1 \times d_2}$, the spectral norm is denoted by $\|X\|_2$. For comparison with respect to the positive semi-definite cone we will use $\succeq$ or $\succ$ for strict inequality. By $\mathbb{P}$ we will denote probability measures and by $\mathbb{E}$ expectation. By $\mathrm{poly}(\cdot)$ we denote a polynomial function of its arguments. By $\exp(\cdot)$ we denote a exponential function of its arguments.

7.2 Problem Statement

System (7.1) is characterized by the matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $H \in \mathbb{R}^{n \times r}$. We assume that $w_k \sim \mathcal{N}(0, I_r)$ is i.i.d. Gaussian with unit covariance. Without loss of generality the initial state is assumed to be zero $x_0 = 0$. In a departure from prior work, we do not necessarily assume that the noise is isotropic. Instead, we consider a more general model, where the noise Hw_k is allowed to be anisotropic and potentially degenerate.

Assumption 7.1. *Matrices A, B, H and the noise dimension $r \leq n$ are all unknown. The unknown matrices are bounded, i.e. $\|A\|_2, \|B\|_2, \|H\|_2 \leq M$, for some positive constant $M \geq 1$. Matrices B, H have full column rank $\mathrm{rank}(B) = p \leq n$, $\mathrm{rank}(H) = r \leq n$. We also assume that the system is non-explosive $\rho(A) \leq 1$.*

The boundedness assumption on the state parameters allows us to argue about global sample complexity upper bounds. To simplify the presentation, we make the assumption that the system is non-explosive $\rho(A) \leq 1$. This setting includes marginally stable systems and is rich enough to provide insights about the difficulty of learning more general systems.

A policy is a sequence of functions $\pi = \{\pi_t\}_{t=0}^{N-1}$. Every function π_t maps previous state-input values $x_0, \dots, x_t, u_0, \dots, u_{t-1}$ and potentially an auxiliary randomization signal AUX to the new input u_t. Hence all inputs u_t are $\mathcal{F}_t$-measurable, where $\mathcal{F}_t \triangleq \sigma(x_0, \dots, x_t, u_0, \dots, u_{t-1}, \mathrm{AUX})$. For brevity we will use the symbol S to denote a system $S = (A, B, H)$. Let $\mathbb{P}_{S,\pi}$ $(\mathbb{E}_{S,\pi}(\cdot))$ denote the probability distribution (expectation) of the input-state data when the true system is equal to S and we apply a policy π.

7.2.1 Difficulty of Stabilization

In the stabilization problem, the goal is to find a state-feedback control law $u = Kx$, where K renders the closed-loop system $A + BK$ stable with spectral radius less than one, i.e., $\rho(A+BK) < 1$. We assume that we collect data $x_0, \dots, x_N, u_0, \dots, u_N$, which are generated by system (7.1) using any exploration policy π, e.g. white-noise excitation, active learning etc. Since we care only about sample complexity, the policy is allowed to be maximally exploratory. To make the problem meaningful, we restrict the average control energy.

Assumption 7.2. *The control energy is bounded $\mathbb{E}_{S,\pi}\|u_t\|_2^2 \le \sigma_u^2$, for some $\sigma_u > 0$.*

Next, we define a notion of learning difficulty for classes of linear systems. By $\mathcal{C}_n$ we will denote a class of systems with dimension n. We will define as easy, classes of linear system that exhibit $\mathrm{poly}(n)$ sample complexity.

Definition 6 (Poly(n)-stabilizable classes)**.** *Let $\mathcal{C}_n$ be a class of systems. Let $\hat{K}_N$ be a function that maps input-state data $(u_0, x_1), \dots, (u_{N-1}, x_N)$ to a control gain. We call the class $\mathcal{C}_n$ $\mathrm{poly}(n)-$stabilizable if there exists an algorithm $\hat{K}_N$ and an exploration policy π satisfying Assumption 7.2, such that for any confidence $0 \le \delta < 1$:*

$$\sup_{S \in \mathcal{C}_n} \mathbb{P}_{S,\pi}\left(\rho(A + B\hat{K}_N) \ge 1\right) \le \delta, \quad \text{if} \quad N \ge \mathrm{poly}(n, \log 1/\delta, M). \tag{7.5}$$

The above class-specific definition can be turned into a local, instance-specific, definition of sample complexity by considering a neighborhood around an unknown system. The question then arises whether linear systems are generally poly(n)-stabilizable.

Problem 7.1. *Are there linear system classes which are not* poly(n)*-stabilizable? When can we guarantee* poly(n)*-stabilizability?*

7.2.2 Difficulty of Online LQR

Consider the LQR objective (7.2). Let the state penalty matrix $Q \in \mathbb{R}^{n\times n} \succ 0$ be positive definite, with the input penalty matrix $R \in \mathbb{R}^{p\times p}$ also positive definite. When the model is known, the optimal policy is a linear feedback law $\pi_\star = \{K_\star x_k\}_{k=0}^{T-1}$, where $K_\star$ is given by

$$K_\star = -(B'PB+R)^{-1}B'PA, \tag{7.6}$$

and P is the unique positive definite solution to the Discrete Algebraic Riccati Equation

$$P = A'PA + Q - A'PB(B'PB+R)^{-1}B'PA. \tag{7.7}$$

Throughout the chapter, we will assume that $Q_T = P$. If the model of (7.1) is unknown, the goal of the learner is to find an online learning policy π that leads to minimum regret $R_T(S)$. In the setting of online LQR, the data are revealed sequentially, i.e. x_{t+1} is revealed after we select u_t. Contrary to the stabilization problem, here we study regret, i.e. there is a tradoff between exploration and exploitation. We will define a class-specific notion of learning difficulty based on the ratio between the regret and $\sqrt{T}$.

Definition 7 (Poly(n)-Regret)**.** *Let $\mathcal{C}_n$ be a class of systems of dimension n. We say that the class $\mathcal{C}_n$ exhibits poly(n) minimax expected regret if*

$$\min_{\pi} \sup_{S\in\mathcal{C}_n} \mathbb{E}_{S,\pi} R_T(S) \le \text{poly}(n, M, \log T)\sqrt{T} + \tilde{O}(1), \tag{7.8}$$

where $\tilde{O}(1)$ hides $\operatorname{poly}\log T$ *terms.*

Our definition here is based on expected regret, but we could have a similar definition based on high probability regret guarantees – see Dann et al. (2017) for distinctions between the two definitions. Similar to the stabilization problem, we pose the following questions.

Problem 7.2. *Are there classes of systems for which poly(n)-regret is impossible? When is poly(n)-regret guaranteed?*

7.3 Classes with Rich Controllability Structure

Before we present our learning guarantees, we need to find classes of systems, where learning is meaningful. To make sure that the stabilization and the LQR problems are well-defined, we assume that system (7.1) is controllable[4].

Assumption 7.3. *System (7.1) is (A, B)* controllable, *i.e. matrix*

$$\mathcal{C}_k(A,B) \triangleq \begin{bmatrix} B & AB & \cdots & A^{k-1}B \end{bmatrix} \tag{7.9}$$

has full column rank $\mathrm{rank}(\mathcal{C}_k(A,B)) = n$*, for some* $k \leq n$*.*

Unsurprisingly, the class of all controllable systems does not exhibit finite sample complexity/regret, let alone polynomial sample complexity/regret. The main issue is that there exist systems which satisfy the rank condition but are arbitrarily close to uncontrollability. For example, consider the following controllable system, which we want to stabilize

$$x_{k+1} = \begin{bmatrix} 1 & \alpha \\ 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k + w_k.$$

The only way to stabilize the system is indirectly by using the second state $x_{k,2}$, via the coupling coefficient α. However, we need to know the sign of α. If α is allowed to be arbitrarily small, i.e. the system is arbitrarily close to uncontrollability, then an arbitrarily large number of samples is required to learn the sign of α, leading to infinite complexity. To obtain classes with finite sample complexity/regret we need to bound the system instances away from uncontrollability. One way is to consider the least singular value of the

[4] We can slightly relax the condition to (A, B) stabilizable (Lale et al., 2020a; Simchowitz & Foster, 2020; Efroni et al., 2021). To avoid technicalities we leave that for future work.

controllability Gramian $\Gamma_k(A, B)$ at time k:

$$\Gamma_k(A, B) \triangleq \sum_{t=0}^{k-1} A^t B B'(A')^t. \tag{7.10}$$

An implicit assumption in prior literature is that $\sigma_{\min}^{-1}(\Gamma_k(A, B)) \leq \text{poly}(n)$. We will not assume this here, since it might exclude many systems of interest, such as integrator-like systems, also known as underactuated systems. Instead, we will relax this requirement to allow richer system structures.

To avoid pathologies, we will lower bound the coupling between states in the case of indirectly controlled systems. To formalize this idea, let us review some notions from system theory. The *controllability index* is defined as follows

$$\kappa(A, B) \triangleq \min \{k \geq 1 : \text{rank}(\mathcal{C}_k(A, B)) = n\}, \tag{7.11}$$

i.e., it is the minimum time such that the controllability rank condition is satisfied. It captures the degree of underactuation and reflects the structural difficulty of control.

Based on the fact that the rank of the controllability matrix at time κ is n, we can show that the pair (A, B) admits the following canonical representation, under a unitary similarity transformation (Dooren, 2003). It is called the Staircase or Hessenberg form of system (7.1).

Proposition 7.1 (Staircase form). *Consider a controllable pair (A, B) with controllability index κ and controllability matrix $\mathcal{C}_k$, $k \geq 0$. There exists a unitary similarity transforma-*

tion $U \in \mathbb{R}^{n\times n}$ such that $U'U = UU' = I$ and:

$$U'B = \begin{bmatrix} B_1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad U'AU = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,\kappa-1} & A_{1,\kappa} \\ A_{2,1} & A_{2,2} & \cdots & A_{3,\kappa-1} & A_{2,\kappa} \\ 0 & A_{3,2} & \cdots & A_{3,\kappa-1} & A_{3,\kappa} \\ 0 & 0 & \cdots & A_{4,\kappa-1} & A_{4,\kappa} \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & A_{\kappa,\kappa-1} & A_{\kappa,\kappa} \end{bmatrix}, \tag{7.12}$$

where $A_{i,j} \in \mathbb{R}^{p_i \times p_j}$ are block matrices, with $p_i = \operatorname{rank}(\mathcal{C}_i) - \operatorname{rank}(\mathcal{C}_{i-1})$, $p_1 = p$, $B_1 \in \mathbb{R}^{p\times p}$. Matrices $A_{i+1,i}$ have full row rank $\operatorname{rank}(A_{i+1,i}) = p_{i+1}$ and the sequence p_i is decreasing.

Matrix U is the orthonormal matrix of the QR decomposition of the first n independent columns of $\mathcal{C}_\kappa(A, B)$. It is unique up to sign flips of its columns. The above representation captures the coupling between the several sub-states via the matrices $A_{i+1,i}$. It has been used before as a test of controllability Dooren (2003). This motivates the following definition, wherein we bound the coupling matrices $A_{i+1,i}$ away from zero.

Definition 8 (Robustly coupled systems)**.** *Consider a controllable system (A, B) with controllability index κ. It is called $\mu-$robustly coupled if and only if for some positive $\mu > 0$:*

$$\sigma_p(B_1) \geq \mu, \quad \sigma_{p_{i+1}}(A_{i+1,i}) \geq \mu, \;\; \textit{for all } 2 \leq i \leq \kappa - 1, \tag{7.13}$$

where B_1, $A_{i+1,i}$ are defined as in the Staircase form (7.12).

In the previous example, by introducing the $\mu-$robust coupling requirement, we enforce a lower bound on the coupling coefficient $\alpha \geq \mu$, thus, avoiding pathological systems.

7.4 Difficulty of Stabilization

In this section, we show that there exist non-trivial classes of linear systems for which the problem of stabilization from data is hard. In fact, the class of robustly coupled systems requires at least an exponential, in the state dimension n, number of samples.

Theorem 7.1 (Stabilization can be Hard). *Consider the class $\mathcal{C}^{\mu}_{n,\kappa}$ of all μ-robustly coupled systems $S = (A, B, H)$ of dimension n and controllability index κ. Let Assumption 7.2 hold and let $\mu < 1$. Then, for any stabilization algorithm, the sample complexity is exponential in the index κ. For any confidence $0 \leq \delta < 1/2$ the requirement*

$$\sup_{S \in \mathcal{C}^{\mu}_{n,\kappa}} \mathbb{P}_{S,\pi}\left(\rho(A + B\hat{K}_N) \geq 1\right) \leq \delta$$

is satisfied only if

$$N\sigma_u^2 \geq \frac{1}{2}\left(\frac{1}{\mu}\right)^{2\kappa-2}\left(\frac{1-\mu}{\mu}\right)^2 \log\frac{1}{3\delta}.$$

Theorem 7.1 implies that system classes with large controllability index, e.g. $\kappa = n$, suffer in general from sample complexity which is exponential with the dimension n. In other words, learning difficulty arises in the case of under-actuated systems. Only a limited number of system states are directly driven by inputs and the remaining states are only indirectly excited, leading to a hard learning and stabilization problem. Consider now systems

$$S_i: \qquad x_{k+1} = \begin{bmatrix} 1 & \alpha_i\mu & 0 & \cdots & 0 \\ 0 & 0 & \mu & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & \mu \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \mu \end{bmatrix} u_k + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} w_k,\ i \in \{1, 2\}, \qquad (7.14)$$

where $0 < \mu < 1$, $\alpha_1 = 1$, $\alpha_2 = -1$. Systems S_1, S_2 are almost identical with the exception of element A_{12} where they have different signs. Both systems have one marginally stable mode corresponding to state $x_{k,1}$. The only way to stabilize $x_{k,1}$ with state feedback is indirectly, via $x_{k,2}$. Given system S_1, since $\alpha_1\mu > 0$, it is necessary that the first component of the gain is negative $\hat{K}_{N,1} < 0$. This follows from the Jury stability criterion, a standard stability test in control theory (Fadali & Visioli, 2013, Ch. 4.5)–see Section 7.9. On the other hand, we can only stabilize S_2 if $\hat{K}_{N,1} > 0$. Hence, the only way to stabilize the system is to identify the sign of α_i. In other words, we transform the stabilization problem into a system identification problem. However, identification of the correct sign is very hard since the excitation of $x_{k,2} = \mu^{n-1} u_{k-n+1}$ scales with μ^{n-1}. The proof relies on Birgé's inequality (Boucheron et al., 2013). In Section 7.9 we construct a slightly more general example with non-zero diagonal elements. Our construction relies on the fact that $\mu < 1$. It is an open question whether we can construct hard learning instances for $\mu \geq 1$.

One insight that we obtain from the above example is that lack of excitation might lead to large sample complexity of stabilization. In particular, this can happen when we have an unstable/marginally stable mode, which can only be controlled via the system identification bottleneck, like $A_{1,2}$ in the above example.

7.4.1 Sample complexity upper bounds

As we show below, sample complexity cannot be worse than exponential under the assumption of robust coupling. If the exploration policy is a white noise input sequence, then using a least squares identification algorithm (Simchowitz et al., 2018), and a robust control design scheme (Dean et al., 2017), the sample complexity can be upper bounded by a function which is at most exponential with the dimension n. In fact, we provide a more refined result, directly linking sample complexity to the controllability index κ. Our proof relies on bounding control theoretic quantities like the least singular value of the controllablility Gramian. The details of the proof and the algorithm can be found in Section 7.10.

Theorem 7.2 (Exponential Upper Bounds)**.** *Consider the class $\mathcal{C}^{\mu}_{n,\kappa}$ of all μ-robustly coupled systems $S = (A, B, H)$ of dimension n and controllability index κ. Let Assumption 7.2 hold. Then, the sample complexity is at most exponential with κ. There exists an exploration policy π and algorithm $\hat{K}_N$ such that for any $\delta < 1$:*

$$\sup_{S \in \mathcal{C}^{\mu}_{n,\kappa}} \mathbb{P}_{S,\pi}\left(\rho(A + B\hat{K}_N) \geq 1\right) \leq \delta, \quad \textit{if} \quad N\sigma_u^2 \geq \operatorname{poly}\left(\left(\frac{M}{\mu}\right)^{\kappa}, M^{\kappa}, n, \log 1/\delta\right).$$

Assume that the constants μ and M are dimensionless. Then, our upper and lower bounds match qualitatively with respect to the dependence on κ. Theorem 7.2 implies that if the degree of underactuation is mild, i.e. $\kappa = O(\log n)$, then robustly coupled systems are guaranteed to be poly(n)-stabilizable. Our upper bound picks up a dependence on the quantity M/μ. Recall that M upper-bounds the norm of A. Hence, it captures a notion of sensitivity of the dynamics A to inputs/noise. In the lower bounds only the coupling term μ appears. It is an open question to prove or disprove whether the sensitivity of A affects stabilization or it is an artifact of our analysis. Another important open problem is to determine the optimal constant that multiplies κ in the exponent. Our lower bound suggests that the exponent can be at least of the order of 2 times κ. In our upper bounds, by following the proof, we get an exponent which is larger than 2.

7.5 Difficulty of online LQR

In the following theorem, we prove that classes of robustly coupled systems can exhibit minimax expected regret which grows at least exponentially with the dimension n. Let $\mathcal{C}^{\mu}_{n,\kappa}$ denote the class of μ-robustly coupled systems $S = (A, B, H)$ of state dimension n and controllability index κ. Define the ϵ-dilation $\mathcal{C}^{\mu}_{n,\kappa}(\epsilon)$ of $\mathcal{C}^{\mu}_{n,\kappa}$ as

$$\mathcal{C}^{\mu}_{n,\kappa}(\epsilon) \triangleq \left\{(A, B, H) : \| \begin{bmatrix} A - \tilde{A} & B - \tilde{B} \end{bmatrix} \|_2 \leq \epsilon, \text{ for some } (\tilde{A}, \tilde{B}, H) \in \mathcal{C}^{\mu}_{n,\kappa}\right\},$$

which consists of every system in $\mathcal{C}^{\mu}_{n,\kappa}$ along with its $\epsilon-$ball around it.

Theorem 7.3 (Exponential Regret Lower Bounds)**.** *Consider the class $\mathcal{C}^{\mu}_{n,\kappa}$ of all μ-robustly coupled systems $S = (A, B, H)$ of state dimension n and controllability index κ, with $\kappa \leq n - 1$. For every $\epsilon > 0$ define the ϵ-dilation $\mathcal{C}^{\mu}_{n,\kappa}(\epsilon)$. Let $Q_T = P$, the solution to the ARE* (7.7)*, and assume $\mu < 1$. Let $0 < \alpha < 1/4$. For any policy π*

$$\liminf_{T\to\infty} \sup_{S\in\mathcal{C}^{\mu}_{n,\kappa}(T^{-\alpha})} \mathbb{E}_{S,\pi} \frac{R_T(S)}{\sqrt{T}} \geq \frac{1}{4\sqrt{n}} 2^{\frac{\kappa-1}{2}}.$$

When the controllability index is large, e.g. $\kappa = n$, then the lower bounds become exponential with n. Hence, achieving poly(n)-regret is impossible in the case of general linear systems. Let us now explain when learning can be difficult. Consider the following $1-$strongly coupled system, which consists of two independent subsystems

$$A = \left[\begin{array}{c|cccc} 0 & 0 & 0 & & 0 \quad 0 \\ \hline 0 & 1 & 1 & & 0 \quad 0 \\ & & & \ddots & \\ 0 & 0 & 0 & & 1 \quad 1 \\ 0 & 0 & 0 & & 0 \quad 1 \end{array}\right], \; B = \left[\begin{array}{c|c} 1 & 0 \\ 0 & 0 \\ \vdots & \\ 0 & 1 \end{array}\right] u_k, \; H = I_n, \; Q = I_n, \; R = I_2, \tag{7.15}$$

where the first subsystem is a memoryless system, while the second one is the discrete integrator of order $n-1$. Since the sub-systems are decoupled, the optimal LQR controller will also be decoupled and structured

$$K_\star = \begin{bmatrix} 0 & 0 \\ 0 & K_{\star,0} \end{bmatrix},$$

where $K_{\star,0}$ is the optimal gain of the second subsystem. The first subsystem (upper-left) is memoryless and does not require any regulation, that is, $[K_\star]_{11} = 0$.

Consider now a perturbed system $\tilde{A} = A - \Delta K_\star$, $\tilde{B} = B + \Delta$, for some $\Delta \in \mathbb{R}^{p\times n}$. Such perturbations are responsible for the $\sqrt{T}$ term in the regret of LQR (Simchowitz & Foster, 2020; Ziemann & Sandberg, 2022); systems (A, B) and $(\tilde{A}, \tilde{B})$ are indistinguishable under the control law $u_t = K_\star x_t$ since $A + BK_\star = \tilde{A} + \tilde{B}K_\star$. Now, informally, to get an $\exp(n)\sqrt{T}$ regret bound it is sufficient to satisfy two conditions: i) the system is sensitive to inputs or noise, in the sense that any exploratory signal can incur extra cost, which grows exponentially with n. ii) the difference $\tilde{A} - A$, $\tilde{B} - B$ is small enough, i.e. polynomial in n, so that identification of Δ requires significant deviation from the optimal policy.

The $n-1$-th integrator is very sensitive to inputs or noises. As inputs $u_{k,2}$ and noises w_k get integrated $(n-1)$-times, this will result in accumulated values that grow exponentially as we move up the integrator chain. Hence, the first informal condition is satisfied. To satisfy the second condition we let the perturbation Δ have the following structure

$$\Delta = \begin{bmatrix} 0 & 0 \\ \Delta_1 & 0 \end{bmatrix}, \tag{7.16}$$

where we only perturb the matrix of the first input $u_{k,1}$. By using two subsystems and the above construction, we make it harder to detect Δ. In particular, because of the structure of the system ($[K_\star]_{11} = 0$) and the perturbation Δ, we have $\tilde{A} = A - \Delta K_\star = A$. Hence $\| \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} \tilde{A} & \tilde{B} \end{bmatrix} \|_2 = \|\Delta\|_2 \leq \text{poly}(n)\|\Delta\|_2$, i.e., the perturbed system does not lie too far away from the nominal one. This last condition might be crucial. If $\|\Delta K_\star\| \geq \exp(n)\|\Delta\|_2$, then it might be possible to distinguish between (A, B) and $(\tilde{A}, \tilde{B})$ without deviating too much from the optimal policy. This may happen if we use only one subsystem, since $\|K_{\star,0}\|_2$ might be large. By using two subsystems, we cancel the effect of $K_{\star,0}$ in $\Delta K_\star$.

In the stabilization problem, we show that the lack of excitation during the system identification stage might hurt sample complexity. Here, we show that if a system is too sensitive to inputs and noises, i.e. some state subspaces are too easy to excite, this can lead to large regret. Both lack of excitation and too much excitation of certain subspaces can hurt learning performance. This was observed before in control (Skogestad et al., 1988).

7.5.1 Sketch of Lower Bound Proof

Let $S_0 = (A_0, B_0, I_{n-1}) \in \mathcal{C}^{\mu}_{n-1,\kappa}$ be a $\mu-$robustly coupled system of state dimension $n-1$, input dimension $p-1$ and controllability index $\kappa \leq n-1$. Let P_0 be the solution of the Riccati equation for $Q_0 = I_{n-1}$, $R_0 = I_{p-1}$, with $K_{\star,0}$ the corresponding optimal gain. Define the steady-state covariance of the closed-loop system

$$\Sigma_{0,x} = (A_0 + B_0 K_{\star,0})\Sigma_{0,x}(A_0 + B_0 K_{\star,0})' + I_{n-1}. \tag{7.17}$$

Now, consider the composite system:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & A_0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & B_0 \end{bmatrix}, H = I_n, \tag{7.18}$$

with $Q = I_n$, $R = I_p$. Let Δ be structured as in (7.16), for some arbitrary Δ_1 of unit norm $\|\Delta_1\|_2 = 1$. The Riccati matrix of the composite system is denoted by P and the corresponding gain by $K_\star$. Consider the parameterization:

$$A(\theta) = A - \theta\Delta K_\star, \qquad B(\theta) = B + \theta\Delta, \tag{7.19}$$

for any $\theta \in \mathbb{R}$. Let $\mathcal{B}(\theta, \epsilon)$ denote the open Euclidean ball of radius ϵ around θ. For every $\epsilon > 0$, define the local class of systems around S as $\mathcal{C}_S(\epsilon) \triangleq \{(A(\theta), B(\theta), I_n),\, \theta \in \mathcal{B}(0, \epsilon)\}$. Based on the above construction and Theorem 1 of Ziemann & Sandberg (2022), a general information-theoretic regret lower bound, we prove the following lemma.

Lemma 7.1 (Two-Subsystems Lower Bound)**.** *Consider the parameterized family of linear systems defined in (7.19), for $n, p \geq 2$ where Δ is structured as in (7.16). Let $Q = I_n$, $R = I_p$. Let $Q_T = P(\theta)$, where $P(\theta)$ is the solution to the Riccati equation for $(A(\theta), B(\theta))$. Then, for any policy π and any $0 < a < 1/4$ the expected regret is lower bounded by*

$$\liminf_{T\to\infty} \sup_{\hat{S}\in\mathcal{C}_S(T^{-a})} \mathbb{E}_{\hat{S},\pi} \frac{R_T(\hat{S})}{\sqrt{T}} \geq \frac{1}{4\sqrt{n}} \sqrt{\Delta_1' P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0 \Delta_1}.$$

Optimizing over Δ_1, we obtain a lower bound on the order of $\|P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0\|_2$. What remains to show is that for the $(n-1)$-th order integrator (second subsystem in (7.15)) the product $\|P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0\|_2$ is exponentially large with n.

Lemma 7.2 (System Theoretic Parameters can be Large)**.** *Consider the $(n-1)-th$ order integrator (second subsystem in (7.15)). Let P_0 be the Riccati matrix for $Q_0 = I_{n-1}, R_0 = 1$,* *with $K_{\star,0}$, $\Sigma_{0,x}$ the corresponding LQR control gain and steady-state covariance. Then*

$$\|P_0\left[\Sigma_{0,x} - I_{n-1}\right]P_0\|_2 \geq \sum_{j=1}^{n-1}\sum_{i=0}^{j}\binom{j}{i}^2 \geq 2^{n-1}$$

Our lemma shows that control theoretic parameters can scale exponentially with the dimension n. The $(n-1)-$th order integrator is a system which is mildly unstable. In Section 7.11.4, we show that **stable** systems can also suffer from the same issue.

7.5.2 Regret Upper Bounds

Similar to the stabilization problem, we show that under the assumption of robust coupling, the regret cannot be worse than $\exp(\kappa)\sqrt{T}$ with high probability. As we prove in Lemma 7.3, the solution P to the Riccati equation has norm $\|P\|_2$ that scales at most exponentially with the index κ in the case of robustly-coupled systems. This result combined with the regret upper bounds of Simchowitz & Foster (2020), give us the following result.

Theorem 7.4 (Exponential Upper Bounds)**.** *Consider a μ-robustly coupled system $S = (A, B, H)$ of dimension n, controllability index κ. Assume that we are given an initial stabilizing gain K_0. Let $Q = I_n$, $R \succeq I_p$, and $Q_T = 0$. Assume that the noise is non-singular $HH' = I_n$. Let $\delta \in (0, 1/T)$. Using the Algorithm 1 of Simchowitz & Foster (2020) with probability at least $1-\delta$:*

$$R_T(A,B) \leq \mathrm{poly}(n, (\frac{M}{\mu})^{\kappa}, M^{\kappa}, \log 1/\delta)\sqrt{T} + \mathrm{poly}(n, (\frac{M}{\mu})^{\kappa}, M^{\kappa}, \log 1/\delta, P(K_0)),$$

where $P(K_0) = (A + BK_0)'P(K_0)(A + BK) + Q + K_0'RK_0$.

The result follows immediately by our Lemma 7.3 and the upper bounds of Theorem 2 in Simchowitz & Foster (2020). Assuming that the plant sensitivity M and the coupling coefficient μ are dimensionless, then if we have a mild degree of underactuation, i.e. $\kappa = O(\log n)$, we get poly(n)-regret with high probability. Note that the above guarantees are for high probability regret which is not always equivalent to expected regret (Dann et al., 2017). Our upper-bounds are almost global for all robustly coupled systems, in the sense that the dominant $\sqrt{T}$-term is globally bounded. To provide truly global regret guarantees it is sufficient to add an initial exploration phase to Algorithm 1 of Simchowitz & Foster (2020), which first learns a stabilizing gain K_0. For this stage we could use the results of Section 7.4.1, and Section 7.10. We leave this for future work.

7.6 Conclusion

We prove that learning to control linear systems can be hard for non-trivial system classes. The problem of stabilization might require sample complexity which scales exponentially with the system dimension n. Similarly, online LQR might exhibit regret which scales exponentially with n. This difficulty arises in the case of underactuated systems. Such systems are structurally difficult to control; they can be very sensitive to inputs/noise or very hard to excite. If the system is robustly coupled and has a mild degree of underactuation (small controllability index), then we can guarantee that learning will be easy.

We stress that system theoretic quantities might not be dimensionless. On the contrary, they might grow very large with the dimension and dominate any poly(n) terms. Hence, going forward, an important direction of future work is to find policies with optimal dependence on such system theoretic quantities. Although the optimal dependence is known for the problem of system identification (Simchowitz et al., 2018; Jedra & Proutiere, 2019), it is still not clear what is the optimal dependence in the case of control. For example, an interesting open problem is to find the optimal dependence of the regret R_T on the Riccati equation solution P. For the problem of stabilization, it is open to find how sample complexity optimally scales with the least singular value of the controllability Gramian.

7.7 Preliminaries: System Theoretic Concepts

In this section, we review briefly some system theoretic concepts. A system $(A, B) \in \mathbb{R}^{n\times(n+p)}$ is **controllable** if and only if the controllability matrix

$$\mathcal{C}_k(A,B) = \begin{bmatrix} B & AB & \cdots & A^{k-1}B \end{bmatrix}$$

has full column rank for some $k \leq n$. The minimum such index κ that the rank condition is satisfied is called the controllability index, and it is always less or equal than the state dimension n. A system (A, B) is called **stabilizable** if and only if there exists a matrix $K \in \mathbb{R}^{p\times n}$ such that $A+BK$ is stable, i.e. has spectral radius $\rho(A+BK)$. Any controllable system is also stabilizable. A system (A', B') is called **observable** if and only if (A, B) is controllable. Similarly (A', B') is **detectable** if and only if (A, B) is stabilizable.

Let A be stable ($\rho(A) < 1$) and consider the transfer matrix $(zI - A)^{-1}, z \in \mathcal{C}$ in the frequency domain. The $\mathcal{H}_\infty$-norm is given by

$$\|(zI - A)^{-1}\|_{\mathcal{H}_\infty} = \sup_{|z|=1} \|(zI - A)^{-1}\|_2.$$

Using the identity $(I - D)^{-1} = I + D + D^2 \ldots$ for $\rho(D) < 1$, we can upper bound the $\mathcal{H}_\infty$-norm by

$$\|(zI - A)^{-1}\|_{\mathcal{H}_\infty} \leq \sum_{t=0}^{\infty} \|A^t\|_2.$$

7.7.1 Properties of the Riccati Equation

Consider the infinite horizon LQR problem defined in (7.2). Let (A, B) be controllable and assume that $Q \succ 0$ is positive semi-definite and $R \succ 0$ is positive definite. As we stated in Section 7.2, the optimal policy $K_\star x_k$ has the following closed-form solution

$$K_\star = -(B'PB + R)^{-1}B'PA,$$

where P is the unique positive definite solution to the **Discrete Algebraic Riccati Equation**

$$P = A'PA + Q - A'PB(B'PB + R)^{-1}B'PA.$$

Moreover, $A + BK_\star$ is stable, i.e. $\rho(A + BK_\star) < 1$. The above solution is well-defined under the conditions of (A, B) controllable, $Q \succ 0$, $R \succ 0$. Note that we can relax the conditions to $Q \succeq 0$ being positive semi-definite, $(A, Q^{1/2})$ detectable, and (A, B) stabilizable, which is a well-known result in control theory (Chan et al., 1984, Th. 3.1).

Consider now the **finite-horizon** LQR problem, under the same assumptions of (A, B) controllable, $Q \succ 0$, and $R \succ 0$

$$J_T^*(S) \triangleq \min_\pi \mathbb{E}_{S,\pi} \left[\sum_{t=0}^{T-1} (x_t' Q x_t + u_t' R u_t) + x_T' Q_T x_T \right]. \tag{7.20}$$

The optimal policy is a feedback law $K_t x_t$, $t \leq T - 1$, with time varying gains. The gains satisfy the following closed-form expression

$$K_t = -(B'P_{t+1}B + R)^{-1}B'P_{t+1}A,$$

where P_t satisfies the **Riccati Difference Equation**

$$P_t = A'P_{t+1}A + Q - A'P_{t+1}B(B'P_{t+1}B + R)^{-1}B'P_{t+1}A, \; P_T = Q_T.$$

It turns out that as we take the horizon to infinity $T \to \infty$, then we get $\lim_{T\to\infty} P_k = P$ exponentially fast, for any fixed k, where P is the positive definite solution to the Algebraic Riccati Equation. The convergence is true under the conditions of (A, B) controllable, $Q \succ 0$, $R \succ 0$. Again we could relax the conditions to $Q \succeq 0$ being positive semi-definite, $(A, Q^{1/2})$ detectable, and (A, B) stabilizable (Chan et al., 1984, Th. 4.1). Note that if we select the terminal cost $Q_T = P$, then trivially $P_t = P$ for all $t \leq T$, and we recover the same controller as in the infinite horizon case.

Finally, a nice property of the Riccati recursion is that the right-hand side is orderpreserving with respect to the matrices P, Q. In particular, define the operator:

$$g(X,Y) = A'XA + Y - A'YB(B'XB+R)^{-1}B'YA.$$

Then, if $X_1 \succeq X_2$, we have that $g(X_1,Y) \succeq g(X_2,Y)$ (Anderson & Moore, 2005, Ch. 4.4). Similarly, if $Y_1 \succeq Y_2$ then $g(X,Y_1) \succeq g(X,Y_2)$.

7.8 Proofs: System Theoretic Bounds for Robustly Coupled Systems

The first result lower bounds the least singular value of the controllability Gramian in terms of the sensitivity M, the coupling coefficient μ, and the controllability index κ of the system.

Theorem 7.5 (Gramian lower bound (Tsiamis & Pappas, 2021))**.** *Consider a system (A,B,H) that satisfies Assumption 7.1, with κ its controllability index. Assume that (A,B) is μ-robustly coupled. Then, the least singular value of the Gramian $\Gamma_\kappa = \Gamma_\kappa(A,B)$ is lower bounded by:*

$$\sigma_{\min}^{-1}(\Gamma_\kappa) \le \mu^{-2}\left(\frac{3M}{\mu}\right)^{2\kappa}.$$

Proof. The result follows from Theorem 5 in Tsiamis & Pappas (2021). The theorem statement requires a different condition, called robust controllability. However, the proof still goes through if we have $\mu-$robust coupling instead. Recall that $\mathcal{C}_\kappa = \mathcal{C}_\kappa(A,B)$ is the controllability matrix (7.9) of (A,B) at κ. Following the proof in (Tsiamis & Pappas, 2021), we arrive at

$$\sqrt{\sigma_{\min}(\Gamma_\kappa)} \le \|\mathcal{C}_\kappa^\dagger\|_2 \le \|\Xi^{\kappa-1}\|_2\|\alpha\|_2,$$

where

$$\Xi = \begin{bmatrix} 1 & 1 & \mu^{-1} \\ \frac{M}{\mu} & \frac{2+M}{\mu} & \frac{M}{\mu} \\ 0 & 0 & \mu^{-1} \end{bmatrix}, \alpha = \begin{bmatrix} \frac{1}{\mu} \\ \frac{M}{\mu^2} \\ \frac{1}{\mu} \end{bmatrix}.$$

The result follows from the crude bounds $\|\Xi\|_2 \le 3M/\mu$, $\|\alpha\|_2 \le \sqrt{3}M/\mu^{-2}$ where we assumed that $M > 1$. □

The following result, upper bounds the solution P to the LQR Riccati equation in terms of the sensitivity M, the coupling coefficient μ, and the controllability index κ of the system.

Lemma 7.3 (Riccati Upper Bounds)**.** *Let the system $(A, B) \in \mathbb{R}^{n\times(n+p)}$ be controllable and $\mu-$robustly coupled with controllability index κ. Let $R \in \mathbb{R}^{p\times p}$ be positive definite and $Q \in \mathbb{R}^{n\times n}$ be positive semi-definite. Assume $T > \kappa$ and consider the Riccati difference equation:*

$$P_{k-1} = A'P_kA + Q - A'P_kB(B'P_kB + R)^{-1}B'P_kA,\ P_T = Q.$$

Then, the Riccati matrix evaluated at time 0 *is upper-bounded by*

$$\|P_0\|_2 \le \text{poly}\Big((\frac{M}{\mu})^\kappa, M^\kappa, \kappa, \|Q\|_2, \|R\|_2\Big).$$

As a result, if $Q \succ 0$, then the unique positive definite solution P of the algebraic Riccati equation:

$$P = A'PA + Q - A'PB(B'PB + R)^{-1}B'PA$$

satisfies the same bound

$$\|P\|_2 \le \text{poly}\Big((\frac{M}{\mu})^\kappa, M^\kappa, \kappa, \|Q\|_2, \|R\|_2\Big).$$

Proof. The optimal policy of the LQR problem does not depend on the noise. Even for deterministic systems, the optimal policy still have the same form $u_t = K_\star x_t$. This property is known as certainty equivalence (Bertsekas, 2017, Ch. 4). In fact, for deterministic systems, the cost of regulation is given explicitly by $x_0'Px_0$. We leverage this idea to upper bound the stabilizing solution of the Riccati equation P.

Step a) Noiseless system upper bound. Consider the noiseless version of system (7.1)

$$x_{k+1} = Ax_k + Bu_k, \quad \|x_0\|_2 = 1. \tag{7.21}$$

Let $u_{0:t}$ be the shorthand notation for

$$u_{0:t} = \begin{bmatrix} u_t \\ \vdots \\ u_0 \end{bmatrix}.$$

Consider the deterministic LQR objective

$$\begin{aligned} \min_{u_{0:T-1}} \quad & J(u_{0:T-1}) \triangleq x_T' Q x_T + \sum_{k=0}^{N-1} x_k' Q x_k + u_k' R u_k \\ \text{s.t.} \quad & \text{dynamics (7.21)}. \end{aligned}$$

The optimal cost of the problem is given by (Bertsekas, 2017, Ch. 4)

$$\min_{u_{0:T-1}} J(u_{0:T-1}) = x_0' P_0 x_0,$$

where P_0 is the value of P_t at time $t = 0$. Let $u_{0:T-1}$ be any input sequence. Immediately, by optimality, we obtain an upper bound for the Riccati matrix P_0:

$$x_0' P_0 x_0 \leq J(u_{0:T-1}). \tag{7.22}$$

Hence, it is sufficient to find a suboptimal policy that incurs a cost which is at most exponential with the controllability index κ.

Step b) Suboptimal Policy. It is sufficient to drive the state x_κ to zero at time κ with minimum energy $u_{0:\kappa-1}$ and then keep $x_{t+1} = 0$, $u_t = 0$, for $t \geq \kappa$. Recall that $\mathcal{C}_k$ is the controllability matrix at time k. By unrolling the state x_κ:

$$x_\kappa = A^\kappa x_0 + \mathcal{C}_\kappa u_{0:\kappa-1}.$$

To achieve $x_\kappa = 0$, it is sufficient to apply the minimum norm control

$$u_{0:\kappa-1} = -\mathcal{C}_\kappa^\dagger A^\kappa x_0,$$

which leads to input penalties

$$\sum_{k=0}^{T-1} u_k' R u_k \leq \|R\|_2 \sigma_{\min}^{-1}(\Gamma_\kappa) M^{2\kappa},$$

where we used the fact that $\|x_0\|_2 = 1$. For the state penalties, we can write in batch form

$$x_{1:\kappa} \triangleq \begin{bmatrix} x_\kappa \\ \vdots \\ x_1 \end{bmatrix} = \begin{bmatrix} B & AB & \cdots & A^{\kappa-1}B \\ 0 & B & \cdots & A^{\kappa-2}B \\ \vdots & & & \\ 0 & 0 & \cdots & B \end{bmatrix} u_{0:\kappa-1} + \begin{bmatrix} A^\kappa \\ A^{\kappa-1} \\ \vdots \\ A \end{bmatrix} x_0.$$

Exploiting the Toeplitz structure of the first matrix above and by Cauchy-Schwartz

$$\begin{aligned} \sum_{t=0}^{T} x_t' Q x_t &\leq \|Q\|_2 (\|x_{1:\kappa}\|_2^2 + 1) \\ &\leq 2\|Q\|_2 \big((\sum_{t=0}^{\kappa-1} \|A^t B\|_2)^2 \|u_{0:\kappa-1}\|_2^2 + \sum_{t=0}^{\kappa} \|A^t\|_2 \big) \\ &\leq 2\kappa^2 \|Q\|_2 (M^{4\kappa} \|R\|_2 \sigma_{\min}^{-1}(\Gamma_\kappa) + M^{2\kappa}). \end{aligned}$$

Putting everything together and since x_0 is arbitrary, we finally obtain

$$\|P_0\|_2 \leq \frac{\|R\|_2}{\sigma_{\min}(\Gamma_\kappa)} (M^{2\kappa} + 2\kappa^2 \|Q\|_2 M^{4\kappa}) + 2\kappa^2 \|Q\|_2 M^{2\kappa}. \tag{7.23}$$

The result for P_0 now follows from Theorem 7.5.

Step c) Steady State Riccati. If the pair $(A, Q^{1/2})$ is observable, then from standard LQR theory-see Section 7.7.1, $\lim_{T\to\infty} P_0 = P$ and the bound for P follows directly. □

Similar results have been reported before (Cohen et al., 2018; Chen & Hazan, 2021).

However, instead of κ and $(M/\mu)^\kappa$, the least singular value $\sigma_{\min}^{-1}(\Gamma_k)$ shows up in the bounds, for some $k \geq \kappa$.

Finally, based on Lemmas B.10, B.11 of Simchowitz & Foster (2020), we provide some upper bounds on the $\mathcal{H}_\infty-$norm of the closed loop response $(zI - A + BK)^{-1}$, where K is the control gain of the optimal LQR controller for some Q and R.

Lemma 7.4 (LQR Robustness Margins)**.** *Let the system $(A, B) \in \mathbb{R}^{n\times(n+p)}$ be controllable and $\mu-$robustly coupled. Let $R = I_p$, $Q = I_n$. Let P be the stabilizing solution of the algebraic Riccati equation:*

$$P = A'PA + Q - A'PB(B'PB + R)^{-1}B'PA$$

with $K_\star$ the respective control gain $K_\star = -(B'PB + R)^{-1}B'PA$. The spectral radius and the $\mathcal{H}_\infty$-norm of the closed loop response are upper bounded by

$$(1 - \rho(A + BK_\star))^{-1} \leq \text{poly}\Big(\big(\frac{M}{\mu}\big)^\kappa, M^\kappa, \kappa\Big) \tag{7.24}$$

$$\|(zI - A - BK_\star)^{-1}\|_{\mathcal{H}_\infty} \leq \text{poly}\Big(\big(\frac{M}{\mu}\big)^\kappa, M^\kappa, \kappa\Big) \tag{7.25}$$

Proof. First, note that since $Q = I$, immediately $(A, Q^{1/2})$ is observable and the stabilizing solution P is well-defined. Note that the Riccati solution P also satisfies the Lyapunov equation

$$P = (A + BK_\star)'P(A + BK_\star) + I + K_\star'K_\star \succeq (A + BK_\star)'P(A + BK_\star) + I \succeq I.$$

As a result,

$$(A + BK_\star)'(A + BK_\star) \overset{i)}{\preceq} (A + BK_\star)'P(A + BK_\star) = P - I \overset{ii)}{\preceq} (1 - \|P\|_2^{-1})P, \tag{7.26}$$

where i) follows from $P \succeq I$. To prove ii) observe that $P - I = P^{1/2}(I - P^{-1})P^{1/2}$ and

$P^{-1} \succeq \|P\|_2^{-1} I$. Hence

$$P - I \succeq P^{1/2}(I - \|P\|_2^{-1} I)P^{1/2} = (1 - \|P\|_2^{-1})P.$$

Applying inequality (7.26) recursively

$$(A + BK_\star)^{t'}(A + BK_\star)^t = \|(A + BK_\star)^t\|_2^2 \leq \left(1 - \|P\|_2^{-1}\right)^t P.$$

From here, we immediately deduce that

$$\rho(A + BK_\star) \leq \sqrt{1 - \|P\|_2^{-1}},$$

which by Lemma 7.3 proves (7.24). For the $\mathcal{H}_\infty$ norm bound

$$\begin{aligned} \|(zI - A - BK_\star)^{-1}\|_{\mathcal{H}_\infty} &\leq \sum_{t \geq 0} \|(A + BK_\star)^t\|_2 \leq \|P\|_2^{1/2} \frac{1}{1 - \sqrt{1 - \|P\|_2^{-1}}} \\ &\leq \|P\|_2^{1/2} \frac{1 + \sqrt{1 - \|P\|_2^{-1}}}{\|P\|_2^{-1}} \leq 2\|P\|_2^{3/2}. \end{aligned}$$

The proof of (7.25) now follows from Lemma 7.3. □

7.9 Proofs: Lower Bounds for the problem of Stabilization

In this section, we prove Theorem 7.1 by using information theoretic methods. The main idea is to find systems that are nearly indistinguishable from data but require completely different stabilization schemes. We rely on Birgé's inequality (Boucheron et al., 2013), which we review below for convenience.

Definition 9 (KL divergence)**.** *Let $\mathbb{P}$, $\mathbb{Q}$ be two probability measures on some space $(\Omega, \mathcal{A})$. Let $\mathbb{Q}$ be absolutely continuous with respect to $\mathbb{P}$, that is $\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(Y 1_A)$ for some integrable* *non-negative random variable with $\mathbb{E}_{\mathbb{P}}(Y) = 1$. The KL divergence $D(\mathbb{Q}||\mathbb{P})$ is given by*

$$D(\mathbb{Q}||\mathbb{P}) \triangleq \mathbb{E}_{\mathbb{Q}}(\log Y).$$

Theorem 7.6 (Birgé's Inequality (Boucheron et al., 2013))**.** *Let $\mathbb{P}_0$, $\mathbb{P}_1$ be probability measures on $(\Omega, \mathcal{E})$ and let E_0, $E_1 \in \mathcal{E}$ be disjoint events. If $1 - \delta \triangleq \min_{i=0,1} \mathbb{P}_i(E_i) \geq 1/2$ then*

$$(1-\delta)\log\frac{1-\delta}{\delta} + \delta\log\frac{\delta}{1-\delta} \leq D(\mathbb{P}_1||\mathbb{P}_0).$$

The KL divergence between two Gaussian distributions with same variance is given below.

Lemma 7.5 (Gaussian KL divergence)**.** *Let $\mathbb{P} = \mathcal{N}(\mu_1, \sigma^2)$ and $\mathbb{Q} = \mathcal{N}(\mu_2, \sigma^2)$ then*

$$D(\mathbb{Q}||\mathbb{P}) = \frac{1}{2\sigma^2}(\mu_1 - \mu_2)^2.$$

7.9.1 Proof of Theorem 7.1

It is sufficient to prove it for $\kappa = n$. The proof for $\kappa < n$ is similar. Let $\alpha > 0$ be such that $\alpha + \mu < 1$. Consider the systems:

$$S_0: \quad x_{k+1} = \begin{bmatrix} 1 & \mu & 0 & \cdots & 0 \\ 0 & \alpha & \mu & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & \mu \\ 0 & 0 & 0 & \cdots & \alpha \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \mu \end{bmatrix} u_k + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} w_k,$$

$$S_1: \quad x_{k+1} = \begin{bmatrix} 1 & -\mu & 0 & \cdots & 0 \\ 0 & \alpha & \mu & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & \mu \\ 0 & 0 & 0 & \cdots & \alpha \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \mu \end{bmatrix} u_k + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} w_k.$$

By construction, the systems are $\mu-$robustly coupled. Denote the state matrices by A_0, A_1 for S_0, S_1 respectively. Let $\phi_1(z) = \det(zI - A_0 - B\hat{K}_N)$, $\phi_2(z) = \det(zI - A_1 - B\hat{K}_N)$ be the respective characteristic polynomials. By Jury's criterion (Fadali & Visioli, 2013, Ch. 4.5), a necessary (but not sufficient) condition for stability is:

$$\phi_0(1) > 0,\ \phi_1(1) > 0.$$

An direct computation gives:

$$\phi_0(1) = \begin{vmatrix} 0 & -\mu & 0 & \cdots & 0 \\ 0 & 1-\alpha & -\mu & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & -\mu \\ -\hat{K}_{N,1} & -\hat{K}_{N,2} & -\hat{K}_{N,3} & \cdots & 1-\alpha-\hat{K}_{N,n} \end{vmatrix} = -\hat{K}_{N,1}\mu^{n-1},$$

$$\phi_1(1) = \hat{K}_{N,1}\mu^{n-1}.$$

As a result, the events:

$$E_0 = \left\{\rho(A_0 + B\hat{K}_N) < 1\right\} \subseteq \left\{\hat{K}_{N,1} < 0\right\}, \quad E_1 = \left\{\rho(A_1 + B\hat{K}_N) < 1\right\} \subseteq \left\{\hat{K}_{N,1} > 0\right\}$$

are disjoint. By Theorem 7.6, a necessary condition for stabilizing both systems with probability larger than $1-\delta$ is:

$$D(\mathbb{P}_0||\mathbb{P}_1) \geq (1-2\delta)\log\frac{1-\delta}{\delta} \geq \log\frac{1}{2.4\delta} \geq \log\frac{1}{3\delta}. \tag{7.27}$$

Here $\mathbb{P}_i$ is a shorthand notation for $\mathbb{P}_{S_i,\pi}$, for $i = 1, 2$.

Meanwhile, by the chain rule of KL divergence (see Exercise 4.4 in Boucheron et al. (2013)):

$$\begin{aligned}
D(\mathbb{P}_0||\mathbb{P}_1) = \mathbb{E}_{\mathbb{P}_0}\Big(& D(\mathbb{P}_0(\mathrm{AUX})||\mathbb{P}_1(\mathrm{AUX})) \\
& + \sum_{k=0}^{N} D(\mathbb{P}_0(x_k|x_{0:k-1}, u_{0:k-1}, \mathrm{AUX})||\mathbb{P}_1(x_k|x_{0:k-1}, u_{0:k-1}, \mathrm{AUX})) \\
& + \sum_{k=0}^{N-1} D(\mathbb{P}_0(u_k|x_{0:k}, u_{0:k-1}, \mathrm{AUX})||\mathbb{P}_1(u_k|x_{0:k}, u_{0:k-1}, \mathrm{AUX})\Big),
\end{aligned}$$

where $x_{0:k}$ is a shorthand notation for $x_0, \dots, x_k$ (same for $u_{0:k}$). By $\mathbb{P}(X|Y)$ we denote the conditional distribution of X given Y. Note that the inputs have the same conditional distributions under both measures hence their KL divergence is zero. As a result

$$\begin{aligned}
D(\mathbb{P}_0||\mathbb{P}_1) &= \mathbb{E}_{\mathbb{P}_0} \sum_{k=0}^{N} D(\mathbb{P}_0(x_k|x_{0:k-1}, u_{0:k-1}, \mathrm{AUX})||\mathbb{P}_1(x_k|x_{0:k-1}, u_{0:k-1}, \mathrm{AUX})) \\
&\overset{1)}{=} \mathbb{E}_{\mathbb{P}_0} \sum_{k=0}^{N} D(\mathbb{P}_0(x_k|x_{k-1}, u_{k-1})||\mathbb{P}_1(x_k|x_{k-1}, u_{k-1}) \\
&\overset{2)}{=} \mathbb{E}_{\mathbb{P}_0} \sum_{k=0}^{N} D(\mathbb{P}_0(x_{k,1}|x_{k-1,1}, x_{k-1,2})||\mathbb{P}_1(x_{k,1}|x_{k-1,1}, x_{k-1,2})\Big),
\end{aligned}$$

where 1) follows from the Markov property of the linear system and 2) follows from an application of the chain rule, the structure of the dynamics, and the fact that all $x_{k,j}$ have the same distribution for $j \geq 2$. Recall that the normal distribution is denoted by $\mathcal{N}(\mu, \Sigma)$. Now we can explicitly compute the KL divergence:

$$\begin{aligned}
D(\mathbb{P}_0||\mathbb{P}_1) &= \mathbb{E}_{\mathbb{P}_0} \sum_{k=1}^{N} D(\mathcal{N}(\alpha x_{k-1,1} + \mu x_{k-1,2}, 1)||\mathcal{N}(\alpha x_{k-1,1} - \mu x_{k-1,2}, 1)) \\
&\overset{i)}{=} \mathbb{E}_{\mathbb{P}_0} \sum_{k=1}^{N} 2\mu^2 x_{k-1,2}^2 = 2\mu^2 \sum_{k=1}^{N} \mathbb{E}_{\mathbb{P}_0} x_{k-1,2}^2, \qquad (7.28)
\end{aligned}$$

where $i)$ follows by Lemma 7.5. By (7.27), (7.28), and Lemma 7.6, it is necessary to have

$$N\sigma_u^2 \geq \frac{1}{2}\left(\frac{1}{\alpha+\mu}\right)^{2n-2}\left(\frac{1-a-\mu}{\mu}\right)^2 \log\frac{1}{3\delta}$$

Since we are free to choose α, it is sufficient to choose $\alpha = 0$. ■

Lemma 7.6. *Consider system S_0 as defined above. Recall that $\mathbb{P}_0$ is a shorthand notation for $\mathbb{P}_{S_0,\pi}$. Then, under Assumption 7.2, we have*

$$\mathbb{E}_{\mathbb{P}_0} x_{k,2}^2 \leq \sigma_u^2(\alpha+\mu)^{2n-2}\left(\frac{1}{1-(a+\mu)}\right)^2$$

Proof. Let e_2 denote the canonical vector $e_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$. Then

$$x_{k,2} = \sum_{t=1}^{k} e_2' A^{t-1} B u_{k-t} = \sum_{t=n-1}^{k} e_2' A^{t-1} B u_{k-t},$$

where the second equality follows from the fact that $e_2' A^{t-1} B$, for $t \leq n-1$. Moreover, we can upper bound:

$$\left|e_2' A^{t-1} B\right| \leq (\alpha+\mu)^{t-1},$$

which follows from the fact that the sub-matrix $[A_0]_{2:n,2:n}$ of A_0 if we delete the first row and column is bi-diagonal and Toeplitz hence $\|[A_0]_{2:n,2:n}\|_2 \leq \alpha+\mu$. Define $c_t \triangleq (\alpha+\mu)^{t-1}$. Then, we can upper bound $|x_{k,2}|$ by

$$|x_{k,2}| \leq \sum_{t=n-1}^{k} c_t \left|u_{k-t}\right|.$$

By Cauchy-Schwartz and Assumption 7.2

$$\mathbb{E}_{S_0,\pi} u_k^2 \leq \sigma_u^2, \quad \mathbb{E}_{S_0,\pi}\left|u_k u_t\right| \leq \sigma_u^2.$$

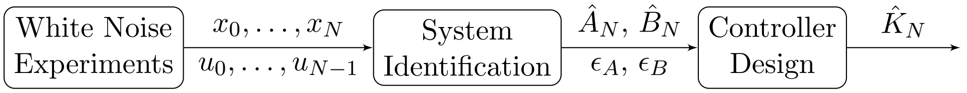


Figure 7.1: The block diagram of the stabilization scheme. First, we generate white noise inputs $u_t \sim \mathcal{N}(0, \bar{\sigma}_u^2 I)$ to excite the system. Then we perform system identification based on least squares to obtain estimates $\hat{A}_N, \hat{B}_N$ of the true system matrices. Finally, we design a controller gain $\hat{K}_N$, based on the system estimates and upper bounds ϵ_A, ϵ_B on the estimation error.

Finally, combining the above results

$$\mathbb{E}_{S_0,\pi} x_{k,2}^2 \le \sigma_u^2 (\sum_{t=n-1}^{k} c_t)^2 \le \sigma_u^2 (\alpha + \mu)^{2n-2} \left(\frac{1}{1 - (a + \mu)} \right)^2,$$

which completes the proof. □

7.10 Proofs: Upper Bounds for the problem of Stabilization

We employ a naive passive learning algorithm, where we employ a white-noise exploration policy to excite the state. Our gain design proceeds in two parts. First, we perform system identification based on least squares (Simchowitz et al., 2018). Second, we use robust control to design the gain based on the identified model and bounds on the identification error of A and B, similar to Dean et al. (2017).

7.10.1 Algorithm

The block diagram for the algorithm is shown in Fig. 7.1. To generate the input data $u_0, \dots, u_{N-1}$, we employ white noise inputs $u_k \sim \mathcal{N}(0, \bar{\sigma}_u^2 I)$, $\bar{\sigma}_u^2 = \sigma_u^2 / p$, where we normalize with p in order to satisfy Assumption 7.2. For the system identification part, we use a least squares algorithm

$$\left[\begin{array}{cc} \hat{A}_N & \hat{B}_N \end{array} \right] = \arg \min_{\{F \in \mathbb{R}^{n \times n}, G \in \mathbb{R}^{n \times p}\}} \sum_{t=0}^{N-1} \| x_{t+1} - F x_t - G u_t \|_2^2, \tag{7.29}$$

to obtain estimates of the matrices A, B. Now, let ϵ_A, ϵ_B be large enough constants such that $\|A - \hat{A}_N\|_2 \le \epsilon_A$, $\|B - \hat{B}_N\|_2 \le \epsilon_B$. To design the controller gain $\hat{K}_N$, it is sufficient to solve the following problem

$$\begin{aligned} \min_{K\in\mathbb{R}^{p\times n}} \quad & 0 \\ \text{s.t.} \quad & \left\| \begin{bmatrix} \sqrt{2}\epsilon_A(zI-\hat{A}_N-\hat{B}_NK)^{-1} \\ \sqrt{2}\epsilon_B K(zI-\hat{A}_N-\hat{B}_NK)^{-1} \end{bmatrix} \right\|_{\mathcal{H}_\infty} < 1. \end{aligned} \tag{7.30}$$

The idea behind the scheme is the following. Let $\hat{K}_N$ be a gain that stabilizes the estimated plant $(\hat{A}_N, \hat{B}_N)$. To make sure that it also stabilizes the nominal plant (A, B) we impose some additional robustness conditions. In fact, as we show in Theorem 7.8, any feasible gain of problem (7.30) will stabilize any plant $(\hat{A}, \hat{B})$ that satisfies $\|\hat{A}-\hat{A}_N\|_2 \le \epsilon_A$, $\|\hat{B}-\hat{B}_N\|_2 \le \epsilon_B$, including the nominal one. In this work, we do not study how to efficiently solve (7.30). For efficient implementations one can refer to Dean et al. (2017). Note that the certainty equivalent LQR design (Mania et al., 2019) or the SDP relaxation method (Cohen et al., 2018; Chen & Hazan, 2021) could also work as stabilization schemes.

7.10.2 System Identification Analysis

Here we review a fundamental system identification result from Simchowitz et al. (2018). The original proof can be easily adapted to the case of singular noise matrices H (Tsiamis & Pappas, 2021).

Theorem 7.7 (Identification Sample Complexity)**.** *Consider a system $S = (A, B, H)$ such that Assumption 7.1 is satisfied. Let (A, B) be controllable with $\Gamma_k = \Gamma_k(A, B)$ the respective controllability Gramian and $\kappa = \kappa(A, B)$ the respective controllability index. Then, under the least squares system identification algorithm (7.29) and white noise inputs $u_k \sim \mathcal{N}(0, \bar{\sigma}_u^2 I_p)$, we obtain*

$$\mathbb{P}_{S,\pi}(\| \begin{bmatrix} A-\hat{A}_N & B-\hat{B}_N \end{bmatrix} \|_2 \ge \epsilon) \le \delta$$

if we have a large enough sample size

$$N\bar{\sigma}_u^2 \geq \frac{\mathrm{poly}(n, \log 1/\delta, M)}{\epsilon^2 \sigma_{\min}(\Gamma_\kappa)} \log N.$$

Proof. The proof is almost identical to the one of Theorem 4 in Tsiamis & Pappas (2021). The difference is that here we consider only the Gramian and index of (A, B) in the final bound, while in Tsiamis & Pappas (2021) the Gramian and index of $(A \begin{bmatrix} H & B \end{bmatrix})$ appears. We repeat the proof here to avoid notation ambiguity. Our goal is to apply Theorem 2.4 in (Simchowitz et al., 2018). Define the noise-controllability Gramian $\Gamma_t^h = \Gamma_t(A, H)$ as well as the combined controllability Gramian

$$\Gamma_t^c = \Gamma_t(A, \begin{bmatrix} \bar{\sigma}_u B & H \end{bmatrix}) = \bar{\sigma}_u^2 \Gamma_t + \Gamma_t^h.$$

Define $y_k = \begin{bmatrix} x_k' & u_k' \end{bmatrix}'$. It follows that for all $j \geq 0$ and all unit vectors $v \in \mathbb{R}^{(n+p)\times 1}$, the following small-ball condition is satisfied:

$$\frac{1}{2\kappa} \sum_{t=0}^{2\kappa} \mathbb{P}(\left|v' y_{t+j}\right| \geq \sqrt{v' \Gamma_{\mathrm{sb}} v} | \bar{\mathcal{F}}_j) \geq \frac{3}{20}, \tag{7.31}$$

where

$$\Gamma_{\mathrm{sb}} = \begin{bmatrix} \Gamma_\kappa^c & 0 \\ 0 & \bar{\sigma}_u^2 I_p \end{bmatrix}. \tag{7.32}$$

Equation (7.31) follows from the same steps as in Proposition 3.1 in Simchowitz et al. (2018) with the choice $k = 2\kappa$.

Next, we determine an upper bound $\bar{\Gamma}$ for the gram matrix $\sum_{t=0}^{N-1} y_t y_t'$. Using a Markov inequality argument as in (Simchowitz et al., 2018, proof of Th 2.1), we obtain that

$$\mathbb{P}(\sum_{t=0}^{N-1} y_t y_t' \preceq \bar{\Gamma}) \geq 1 - \delta,$$

where

$$\bar{\Gamma} = \frac{n+p}{\delta} N \begin{bmatrix} \Gamma_N^c & 0 \\ 0 & \bar{\sigma}_u^2 I_p \end{bmatrix}.$$

Now, we can apply Theorem 2.4 of Simchowitz et al. (2018). With probability at least $1 - 3\delta$ we have $\| \begin{bmatrix} A - \hat{A}_N & B - \hat{B}_N \end{bmatrix} \|_2 \leq \epsilon$ if:

$$N \geq \frac{\text{poly}(n, \log 1/\delta, M)}{\epsilon^2 \sigma_{\min}(\Gamma_\kappa^c)} \log\det(\bar{\Gamma}\Gamma_{\text{sb}}^{-1}),$$

where we have simplified the expression by including terms in the polynomial term. Using Lemma 1 in Tsiamis & Pappas (2021), we obtain

$$\log\det(\bar{\Gamma}\Gamma_{\text{sb}}^{-1}) = \text{poly}(n, M, \log 1/\delta) \log N.$$

Moreover, we use the lower bound $\Gamma_k^c \succeq \bar{\sigma}_u^2 \Gamma_k$, which holds for every $k \geq 0$. □

We note that we can easily obtain sharper bounds by considering the combined controllability Gramian $\Gamma_k(A, \begin{bmatrix} \bar{\sigma}_u B & H \end{bmatrix})$ for the identification stage. For the economy of the presentation, we omit such an analysis here.

7.10.3 Sensitivity of Stabilization

Here we prove that when (7.30) is feasible, then $\hat{K}_N$ stabilizes all plants (A, B) such that $\|A - \hat{A}_N\|_2 \leq \epsilon_A$, $\|B - \hat{B}_N\|_2 \leq \epsilon_B$. We also show that feasibility is guaranteed as long as we can achieve small enough error bounds ϵ_A, ϵ_B.

Theorem 7.8. *Let $\hat{K}_N$ be a feasible solution to problem (7.30) for some $\epsilon_A, \epsilon_B > 0$. Then for any system (A, B) such that $\|A - \hat{A}_N\|_2 \leq \epsilon_A$, $\|B - \hat{B}_N\|_2 \leq \epsilon_B$ we have that*

$$\rho(A + B\hat{K}_N) < 1.$$

Moreover, there exists an $\epsilon_0 > 0$ *such that*

$$\epsilon_0 = \text{poly}\Big((\frac{M}{\mu})^{\kappa}, M^{\kappa}, \kappa \Big)$$

and Problem (7.30) *is feasible if* $\epsilon_A, \epsilon_B \leq \epsilon_0$.

Proof. Let $\hat{K}_N$ be a feasible solution to problem (7.30). Define $\mathbf{\Phi}_x = (zI - \hat{A}_N - \hat{B}_N\hat{K}_N)^{-1}$, which is well-defined and stable since $\epsilon_A > 0$ and $\|\mathbf{\Phi}_x\|_{\mathcal{H}_\infty} < 1/(\sqrt{2}\epsilon_A)$. Define the system difference

$$\mathbf{\Delta} \triangleq (\hat{A}_N - A)\mathbf{\Phi}_x + (\hat{B}_N - B)\hat{K}_N\mathbf{\Phi}_x$$

It follows from simple algebra that:

$$\begin{aligned} zI - A - B\hat{K}_N &= zI - \hat{A}_N - \hat{B}_N\hat{K}_N + (\hat{A}_N - A) + (\hat{B}_N - B)\hat{K}_N \\ &= (I + \mathbf{\Delta})(zI - \hat{A}_N - \hat{B}_N\hat{K}_N). \end{aligned}$$

If $(I + \mathbf{\Delta})^{-1}$ is stable then the closed loop response is stable and well-defined

$$(zI - A - B\hat{K}_N)^{-1} = (zI - \hat{A}_N - \hat{B}_N\hat{K}_N)^{-1}(I + \mathbf{\Delta})^{-1}.$$

But $(I + \mathbf{\Delta})^{-1}$ being stable is equivalent to

$$\|(I + \mathbf{\Delta})^{-1}\|_{\mathcal{H}_\infty} < \infty.$$

A sufficient condition for this to occur is to require (Dean et al., 2017)

$$\|\mathbf{\Delta}\|_{\mathcal{H}_\infty} < 1.$$

By Proposition 3.5 (select $\alpha = 1/2$) of (Dean et al., 2017)

$$\|\boldsymbol{\Delta}\|_{\mathcal{H}_\infty} < \left\| \begin{bmatrix} \sqrt{2}\epsilon_A (zI - \hat{A}_N - \hat{B}_N K)^{-1} \\ \sqrt{2}\epsilon_B K (zI - \hat{A}_N - \hat{B}_N K)^{-1} \end{bmatrix} \right\|_{\mathcal{H}_\infty} < 1.$$

This completes the proof of $\rho(A + B\hat{K}_N) < 1$.

To prove feasibility consider the optimal LQR gain $K_\star$, for $Q = I_n$, $R = I_p$. Following Lemma 4.2 in Dean et al. (2017), if the following sufficient condition holds

$$(\epsilon_A + \epsilon_B \|K_\star\|_2)\|(zI - A - BK_\star)^{-1}\|_{\mathcal{H}_\infty} \le 1/5,$$

then $K_\star$ is a feasible solution

$$\left\| \begin{bmatrix} \sqrt{2}\epsilon_A (zI - \hat{A}_N - \hat{B}_N K_\star)^{-1} \\ \sqrt{2}\epsilon_B K_\star (zI - \hat{A}_N - \hat{B}_N K_\star)^{-1} \end{bmatrix} \right\|_{\mathcal{H}_\infty} < 1.$$

Hence, we can choose

$$\epsilon_0 = \left(5(1 + \|K_\star\|_2)\|(zI - A - BK_\star)^{-1}\|_{\mathcal{H}_\infty}\right)^{-1}. \tag{7.33}$$

The fact that $\epsilon_0 = \mathrm{poly}\left((\frac{M}{\mu})^\kappa, M^\kappa, \kappa\right)$ follows from Lemmas 7.3, 7.4. □

7.10.4 Proof of Theorem 7.2

Let $u_t \sim \mathcal{N}(0, \bar{\sigma}_u^2 I)$, with $\bar{\sigma}_u^2 = \sigma_u^2/p$. Consider the stabilization algorithm as described in (7.29), (7.30). Consider the ϵ_0 defined in (7.33). By Theorems 7.7, 7.8, if

$$N\sigma_u^2 \ge \underbrace{\triangleq \frac{\mathrm{poly}(n, \log 1/\delta, M)}{\epsilon_0^2 \sigma_{\min}(\Gamma_\kappa)}}_{\mathcal{N}} \log N$$

we have with probability at least $1-\delta$ that $\|A-\hat{A}_N\|_2, \|B-\hat{B}_N\|_2 \le \epsilon_0$ and problem (7.30) is feasible with $\epsilon_B = \epsilon_A = \epsilon_0$. By Theorems 7.5 7.8,

$$\mathcal{N} = \text{poly}\left(\left(\frac{M}{\mu}\right)^{\kappa}, M^{\kappa}, n, \log 1/\delta\right).$$

To complete the proof we use the fact that

$$N \ge c \log N \text{ if } N \ge 2c \log 2c.$$

7.11 Proofs: Regret Lower Bounds

First let us state an application of the main result of Ziemann & Sandberg (2022). Consider a system $(A, B, H) \in \mathbb{R}^{n\times(n+p+n)}$, where (A, B) is controllable and $H = I_n$. Let P be the respective Riccati matrix for $Q = I_n$, $R = I_p$, with $K_\star$ the respective optimal LQR gain. Fix a matrix $\Delta \in \mathbb{R}^{p\times n}$ and define the family of systems:

$$A(\theta) = A - \theta B\Delta,\ B(\theta) = B + \theta\Delta,\ H(\theta) = I_n, \tag{7.34}$$

where $\theta \in \mathcal{B}(0, \epsilon)$, for some small ϵ. Assume that ϵ is small enough, such that the Riccati equation has a stabilizing solution for every system in the above family. The respective Riccati matrix is denoted by $P(\theta)$ and the LQR gain by $K(\theta)$. The derivative of $K_\star(\theta)$ with respect to θ at point $\theta = 0$ is given by the following formula.

Lemma 7.7 (Lemma 2.1 (Simchowitz & Foster, 2020)). *If the system (A, B) is stabilizable, then*

$$\frac{d}{d\theta} K_\star(\theta)|_{\theta=0} = -(B'PB + R)^{-1}\Delta' P(A + BK_*).$$

Finally, let Σ_x be the solution to the Lyapunov equation:

$$\Sigma_x = (A + BK_\star)\Sigma_x(A + BK_\star)' + I_n. \tag{7.35}$$

Theorem 7.9 (Application of Theorem 1 in Ziemann & Sandberg (2022))**.** *Consider a system $S = (A, B, H) \in \mathbb{R}^{n\times(n+p+n)}$, where (A, B) is controllable and $H = I_n$. Let P be the respective solution of the algebraic Riccati equation for $Q = I_n$, $R = I_p$, with $K_\star$ the respective optimal LQR gain. Recall the definition of Σ_x in (7.35). Define the family of systems $\mathcal{C}_S(\epsilon) \triangleq \{(A(\theta), B(\theta), I_n),\ \theta \in \mathcal{B}(0, \epsilon)\}$ as defined in (7.34), for any $\epsilon > 0$ sufficiently small such that $P(\theta)$ and $K_\star(\theta)$ are well-defined. Let $Q_T = P(\theta)$. Then for any $\alpha \in (0, 1/4)$:*

$$\liminf_{T\to\infty} \sup_{\hat{S}\in\mathcal{C}_S(T^{-a})} \mathbb{E}_{\hat{S},\pi} \frac{R_T(\hat{S})}{\sqrt{T}} \geq \frac{1}{2\sqrt{2}}\sqrt{\frac{F}{L}}, \tag{7.36}$$

where

$$F = \operatorname{tr}\left((B'PB+R)^{-1}\Delta' P\left[\Sigma_x - I_n\right]P\Delta\right)$$
$$L = n(\|\Delta K_\star\|_2^2 + \|\Delta\|_2^2)\|(B'PB+R)^{-1}\|_2$$

Proof. Note that if $\Delta' P(A + BK_\star) = 0$, then since $\Sigma_x \succeq I_n$ is invertible

$$\begin{aligned}\Delta' P(A+BK_\star) = 0 &\Leftrightarrow \Delta' P(A+BK_\star)\Sigma_x(A+BK_\star)' P\Delta = 0\\ &\Leftrightarrow \Delta' P(\Sigma_x - I_n)P\Delta = 0.\end{aligned}$$

This implies that $F = 0$ and the regret lower bound becomes 0, in which case the claim of the theorem is trivially true. Hence, we will assume that $\Delta' P(A + BK_\star) \neq 0$.

All systems in the family have the same closed-loop response under the control policy $u = K_\star x$. In particular, for all $\theta \in \mathcal{B}(0, \epsilon)$:

$$\frac{d}{d\theta}\begin{bmatrix} A(\theta) & B(\theta)\end{bmatrix}\begin{bmatrix} I_n \\ K_\star\end{bmatrix} = \begin{bmatrix} -\Delta K_\star & \Delta\end{bmatrix}\begin{bmatrix} I_n \\ K_\star\end{bmatrix} = 0.$$

Moreover, by Lemma 7.7

$$\frac{d}{d\theta}K_\star(\theta)|_{\theta=0} = (B'PB+R)^{-1}\Delta' P(A+BK_\star) \neq 0.$$

By Proposition 3.4 in Ziemann & Sandberg (2022), the above two conditions imply that the family $\mathcal{C}_S(\epsilon)$ is $\epsilon-$uninformative (see Section 3 in Ziemann & Sandberg (2022) for definition).

Next, by Lemma 3.6 in Ziemann & Sandberg (2022), the family is also $L-$information regret bounded (see Section 3 in Ziemann & Sandberg (2022) for the definition), where

$$L = \mathrm{tr}(I_n)\| \begin{bmatrix} -\Delta K_\star & \Delta \end{bmatrix} \|_2^2 \|(B'PB+R)^{-1}\|_2 \overset{i)}{\leq} n(\|\Delta K_\star\|_2^2 + \|\Delta\|_2^2)\|(B'PB+R)^{-1}\|_2.$$

Inequality $i)$ follows from $\mathrm{tr}(I_n) = n$ and the norm property

$$\| \begin{bmatrix} M_1 & M_2 \end{bmatrix} \|_2^2 = \| \begin{bmatrix} M_1 & M_2 \end{bmatrix} \begin{bmatrix} M_1 & M_2 \end{bmatrix}' \|_2 = \|M_1M_1' + M_2M_2'\|_2 \leq \|M_1\|_2^2 + \|M_2\|_2^2.$$

Applying Theorem 1 in Ziemann & Sandberg (2022), we get (7.36), for L defined as above and

$$F = \mathrm{tr}\left(\left[\Sigma_x \otimes (B'P(\theta)B + R)\right] (\frac{d}{d\theta}\mathrm{vec}K_\star(\theta)|_{\theta=0})(\frac{d}{d\theta}\mathrm{vec}K_\star(\theta)|_{\theta=0})' \right),$$

where $\otimes$ is the Kronecker product and vec is the vectorization operator (mapping a matrix into a column vector by stacking its columns). Using the identities:

$$\mathrm{vec}(XYZ) = (Z' \otimes X)\mathrm{vec}(Y), \qquad \mathrm{tr}(\mathrm{vec}(X)\mathrm{vec}(Y)') = \mathrm{tr}(XY'),$$

we can rewrite F as

$$F = \mathrm{tr}\left((B'P(\theta)B + R)\frac{d}{d\theta}K(\theta)|_{\theta=0}\Sigma_x \frac{d}{d\theta}K'(\theta)|_{\theta=0} \right).$$

By Lemma 7.7 and the property $\mathrm{tr}(XY) = \mathrm{tr}(YX)$, we finally get

$$F = \mathrm{tr}\left((B'PB + R)^{-1}\Delta' P(A + BK_*)\Sigma_x(A + BK_*)P\Delta \right).$$

The result follows from $(A + BK_*)\Sigma_x(A + BK_*)' = \Sigma_x - I_n$. □

7.11.1 Proof of Lemma 7.1

The result follows by Theorem 7.9. We only need to compute and simplify F, L. Due to the structure of system (7.18), we have

$$P = \begin{bmatrix} 1 & 0 \\ 0 & P_0 \end{bmatrix}, K_\star = \begin{bmatrix} 0 & 0 \\ 0 & K_{0,\star} \end{bmatrix}.$$

Moreover, due to the structure of the perturbation Δ in (7.16)

$$B'PB + R = \begin{bmatrix} 2 & 0 \\ 0 & B_0'P_0B_0 + R_0 \end{bmatrix}, P\Delta(B'PB+R)^{-1}\Delta' P = \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & P_0\Delta_1\Delta_1' P_0 \end{bmatrix}.$$

Hence

$$F = \frac{1}{2}\operatorname{tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & P_0\Delta_1\Delta_1' P_0 \end{bmatrix}(\Sigma_x - I_n)\right) = \frac{1}{2}\Delta_1' P_0(\Sigma_{0,x} - I_{n-1})P_0\Delta_1$$

Finally we have $L \leq n$, since $\Delta K_\star = 0$, Δ_1 has unit norm, and $R = I_p$. ■

7.11.2 Proof of Lemma 7.2

First note that $P_0 \succeq Q_0 = I_{n-1}$. As a result, we have

$$\|P_0(\Sigma_{0,x} - I_{n-1})P_0\|_2 \geq \|\Sigma_{0,x} - I_{n-1}\|_2.$$

It is sufficient to lower bound $\|\Sigma_{0,x} - I_{n-1}\|_2$. Consider the recursion:

$$\Sigma_k = (A_0 + B_0K_{0,\star})\Sigma_{k-1}(A_0 + B_0K_{0,\star})' + I_{n-1},\ \Sigma_0 = 0.$$

Then $\Sigma_{0,x} = \lim_{k\to\infty} \Sigma_k \succeq \Sigma_{n-1} \succeq I_{n-1}$. The second inequality follows from monotonicity of the Lyapunov operator:

$$g(X) = (A_0 + B_0 K_{0,\star}) X (A_0 + B_0 K_{0,\star})' + I_{n-1},$$

i.e. $g(X) \succeq g(Y)$ if $X \succeq Y$. What remains is to lower bound $\|\Sigma_{n-1} - I_{n-1}\|_2$. Let $e_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}'$ be the first canonical vector. Due to the structure of A_0, B_0

$$e_1'(A_0 + B_0 K_{0,\star})^i = e_1'(A_0)^i, \text{ for } i \le n-1.$$

Hence

$$\begin{aligned} \|\Sigma_{n-1} - I_{n-1}\|_2 &\ge e_1'(\Sigma_{n-1} - I_{n-1}) e_1 \\ &= \sum_{k=1}^{n-1} e_1' A_0^k (A_0')^k e_1. \end{aligned}$$

After some algebra we can compute analytically

$$\|\Sigma_{n-1} - I_{n-1}\|_2 \ge \sum_{k=1}^{n-1} \sum_{t=0}^{k} \binom{k}{t}^2 = \sum_{k=1}^{n-1} \binom{2k}{k} \ge \binom{2(n-1)}{n-1} \ge \left(2\frac{n-1}{n-1}\right)^{n-1} = 2^{n-1},$$

which completes the proof. ■

7.11.3 Proof of Theorem 7.3

It is sufficient to prove the result for the class $\mathcal{C}^{\mu}_{n,n-1}$. If $n > \kappa + 1$, then we can consider the system:

$$\tilde{A} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A \end{array}\right], \tilde{B} = \left[\begin{array}{c|c} I_{n-\kappa-1} & 0 \\ \hline 0 & B \end{array}\right], \tilde{H} = \left[\begin{array}{c|c} I_{n-\kappa-1} & 0 \\ \hline 0 & H \end{array}\right]$$

where $(A, B, H) \in \mathcal{C}^{\mu}_{\kappa,\kappa-1}$ and repeat the same arguments.

The proof follows from Lemma 7.1 and Lemma 7.2. What remains to show that for every ϵ

$$\mathcal{C}_S(\epsilon) \subseteq \mathcal{C}^{\mu}_{n,n-1}(\epsilon).$$

This follows from the fact that $\Delta K_\star = 0$, hence $A = A(\theta)$ and $\|B - B(\theta)\| = \theta\|\Delta\|_2 = \theta \leq \epsilon$. Thus,

$$\| \begin{bmatrix} A - A(\theta) & B - B(\theta) \end{bmatrix} \|_2 \leq \epsilon.$$

Since $\mathcal{C}_S(\epsilon) \subseteq \mathcal{C}^{\mu}_{n,n-1}(\epsilon)$, we get

$$\liminf_{T\to\infty} \sup_{S \in \mathcal{C}^{\mu}_{n,n-1}(T^{-a})} \mathbb{E}_{\hat{S},\pi} \frac{R_T(\hat{S})}{\sqrt{T}} \geq \liminf_{T\to\infty} \sup_{\hat{S} \in \mathcal{C}_S(T^{-a})} \mathbb{E}_{\hat{S},\pi} \frac{R_T(\hat{S})}{\sqrt{T}} \qquad \blacksquare$$

7.11.4 Stable System Example

Here we show that the local minimax expected regret can be exponential in the dimension even for stable systems. Using again the two subsystems trick, consider the following stable system

$$S: \qquad x_{k+1} = \left[\begin{array}{c|ccccc} 0 & 0 & 0 & & 0 & 0 \\ \hline 0 & \rho & 2 & & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & & \rho & 2 \\ 0 & 0 & 0 & & 0 & \rho \end{array}\right] x_k + \left[\begin{array}{c|c} 1 & 0 \\ 0 & 0 \\ \vdots & \\ 0 & 1 \end{array}\right] u_k + w_k,\ 0 < \rho < 1, \tag{7.37}$$

with $Q = I_n$, $R = I_2$. Following the notation of (7.18) let:

$$A_0 = \begin{bmatrix} \rho & 2 & 0 & & 0 & 0 \\ 0 & \rho & 2 & & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & & \rho & 2 \\ 0 & 0 & 0 & & 0 & \rho \end{bmatrix},\ B_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},\ Q_0 = I_{n-1},\ R_0 = 1, \tag{7.38}$$

where $A_0 \in \mathbb{R}^{(n-1)\times(n-1)}$ and $B_0 \in \mathbb{R}^{n-1}$. Note that A_0 has spectral radius $\rho < 1$. Let $\Delta = \begin{bmatrix} 0 & 0 \\ \Delta_1 & 0 \end{bmatrix}$. Then, by Lemma 7.1, the local minimax expected regret for system S, given the perturbation Δ_1 is lower bounded by

$$\liminf_{T\to\infty} \sup_{\hat{S}\in\mathcal{C}_S(T^{-a})} \mathbb{E}_{\hat{S},\pi} \frac{R_T(\hat{S})}{\sqrt{T}} \geq \frac{1}{4\sqrt{n}} \sqrt{\Delta_1' P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0 \Delta_1}.$$

As we show in the following lemma, the quantity $\sqrt{\Delta_1' P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0 \Delta_1}$ is exponential with n if we choose Δ_1 appropriately. Although the system is stable, it is very sensitive to inputs and noises. Any signal $u_{k,2}$ that we apply gets amplified by 2 as we move up the chain from state $x_{k,n}$ to state $x_{k,2}$. As a result, any suboptimal policy will result in excessive excitation of the state.

Lemma 7.8 (Stable systems can be hard to learn)**.** *Consider system* (7.38) *Let P_0 be the Riccati matrix for $Q_0 = I_{n-1}, R_0 = 1$, with $K_{\star,0}$, $\Sigma_{0,x}$ the corresponding LQR control gain and steady-state covariance, respectively. Then*

$$\|P_0 \left[\Sigma_{0,x} - I_{n-1}\right] P_0\|_2 \geq 2^{4n-8} + o(1),$$

where $o(1)$ goes to zero as $n \to \infty$.

Proof. Let $\Delta_1 = \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 \end{bmatrix}'$. It is sufficient to prove that

$$\Delta_1' P_0 (\Sigma_{0,x} - I_{n-1}) P_0 \Delta_1$$

is exponential. Using the identity $\Sigma_{0,x} - I_{n-1} = (A_0 + B_0 K_{\star,0})\Sigma_{0,x}(A_0 + B_0 K_{\star,0})'$, $\Sigma_{0,x} \succeq I$, we have:

$$\Delta_1' P_0 (\Sigma_{0,x} - I_{n-1}) P_0 \Delta_1 \geq \|\Delta_1' P_0 (A_0 + B_0 K_{\star,0})\|_2^2.$$

By Lemma 7.10 and Lemma 7.9 it follows that

$$\|\Delta_1' P_0 (A_0 + B_0 K_{\star,0})\|_2^2 \geq 2^{4n-8} + o(1).$$

□

Lemma 7.9 (Riccati matrix can grow exponentially)**.** *For system* (7.38) *we have:*

$$B_0'P_0B_0 + R_0 \geq 2^{2n-4} + 1.$$

Proof. Consider the Riccati operator:

$$g(X,Y) = A_0'XA_0 + Y - A_0'XB_0(B_0'XB_0 + R_0)^{-1}B_0'XA_0.$$

Based on the above notation, we have $P_0 = g(P_0, Q_0)$. The Riccati operator is monotone (Anderson & Moore, 2005), i.e

$$X_1 \succeq X_2 \Rightarrow g(X_1, Y) \succeq g(X_1, Y).$$

It is also trivially monotone with respect to Y. Let $X_0 = 0$, then the recursion $X_{t+1} = g(X_t, Q_0)$ converges to P_0. By monotonicity

$$P_0 \succeq X_t, \text{ for all } t \geq 0$$

Let e_i denote the i-th canonical vector in $\mathbb{R}^{n-1}$. By monotonicity, we also have:

$$X_1 = g(X_0, Q_0) \succeq g(X_0, e_1e_1') = \underbrace{e_1e_1'}_{\tilde{X}_1}$$

Repeating the argument:

$$X_2 = g(X_1, Q_0) \succeq g(\tilde{X}_1, Q_0) \succeq g(\tilde{X}_1, e_1e_1') = \underbrace{A_0'\tilde{X}_1A_0 + e_1e_1'}_{\tilde{X}_2} = A_0'e_1e_1'A_0 + e_1e_1'$$
$$= 2^2e_2e_2' + \rho^2e_1e_1' + 2\rho e_1e_2' + 2\rho e_2e_1'$$

Similarly,

$$X_{n-1} = g(X_{n-2}, Q_0) \succeq g(\tilde{X}_{n-2}, e_1 e_1') = (A_0')^{n-2} e_1 e_1' A_0^{n-2} + (A_0')^{n-1} e_1 e_1' A_0^{n-1} + \cdots + e_1 e_1',$$

where we use the fact that every $\tilde{X}_k$ is orthogonal to B_0 for $k \leq n-2$. As a result:

$$\begin{aligned} [P_0]_{n-1,n-1} \geq [X_n]_{n-1,n-1} &\geq e_{n-1}'(A_0')^{n-2} e_1 e_1' A_0^{n-2} e_{n-1} \\ &= (e_1' A_0^{n-2} e_{n-1})^2 = ([A_0^{n-2}]_{1,n-1})^2 \end{aligned} \tag{7.39}$$

What remains is to compute $[A_0^{n-2}]_{1,n-1}$. Define by $J \in \mathbb{R}^{(n-1)\times(n-1)}$ the companion matrix:

$$J = \begin{bmatrix} 0 & 1 & 0 & & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & & 0 & 1 \\ 0 & 0 & 0 & & 0 & 0 \end{bmatrix}.$$

Since $A_0 = \rho I + 2J$ and I commutes with J by the binomial expansion formula:

$$A_0^{n-2} = 2^{n-2} J^{n-2} + \sum_{t=0}^{n-3} 2^t \binom{n-2}{t} J^t.$$

Since $e_1' J^{n-1} e_{n-1} = 1$, $e_1' J^t e_{n-1} = 0$, for $t \leq n-2$, we obtain:

$$([A_0^{n-2}]_{1,n-1})^2 = 2^{2n-4}. \tag{7.40}$$

By (7.39) and (7.40) we finally get

$$B_0' P_0 B_0 + R_0 = [P_0]_{n-1,n-1} + 1 \geq 2^{2n-4} + 1$$

□

Lemma 7.10. *We have:*

$$\|\Delta_1' P_0(A_0 + B_0 K_{\star,0})\|_2 \geq (0.5 + o(1))(B_0' P_0 B_0 + R_0),$$

where the $o(1)$ is in the large n regime.

Proof. Let e_i denote the i-th canonical vector in $\mathbb{R}^{n-1}$. It is sufficient to show that

$$\left|(B_0' P_0 B_0 + R_0)^{-1} \Delta_1' P_0(A_0 + B_0 K_{\star,0}) e_{n-1}\right| \geq 0.5 + o(1).$$

For simplicity we will denote:

$$\alpha \triangleq [P_0]_{n-1,n-1}, \quad \beta \triangleq [P_0]_{n-2,n-2}, \quad \gamma \triangleq [P_0]_{n-1,n-2}.$$

Due to the structure of A_0, we have

$$A_0 e_{n-1} = \rho e_{n-1} + 2e_{n-2}.$$

Using this, we obtain

$$\begin{aligned} K_{\star,0} e_{n-1} &= -(B_0' P_0 B_0 + 1)^{-1} B_0' P_0 A_0 e_{n-1} = -(\alpha + 1)^{-1} e_{n-1}' P_0(\rho e_{n-1} + 2e_{n-2}) \\ &= -(\alpha + 1)^{-1}(\rho\alpha + 2\gamma). \end{aligned} \tag{7.41}$$

Combining the above results

$$\begin{aligned} &(B_0' P_0 B_0 + R)^{-1} \Delta_1' P_0(A_0 + B_0 K_{\star,0}) e_{n-1} = (B_0' P_0 B_0 + 1)^{-1} e_{n-2}' P_0(A_0 + B_0 K_{\star,0}) e_{n-1} \\ &= (\alpha + 1)^{-1} \Big\{ e_{n-2}' P_0(\rho e_{n-1} + 2e_{n-2}) - e_{n-2}' P_0 e_{n-1}(\alpha + 1)^{-1}(\rho\alpha + 2\gamma) \Big\} \\ &= (\alpha + 1)^{-1} \left\{ \rho\gamma + 2\beta - \gamma(\alpha + 1)^{-1}(\rho\alpha + 2\gamma) \right\} \\ &= 2(\alpha + 1)^{-1} \left\{ \beta - (\alpha + 1)^{-1}\gamma^2 \right\} - (\alpha + 1)^{-2}\rho\gamma \\ &\overset{i)}{=} \frac{2}{\alpha + 1} \left\{ \beta - \frac{\gamma^2}{\alpha + 1} \right\} + o(1), \end{aligned}$$

where i) follows from Lemma 7.11. What remains to show is that

$$\frac{2}{\alpha+1}\left\{\beta-\frac{\gamma^2}{\alpha+1}\right\}=0.5+o(1). \tag{7.42}$$

Using the algebraic Riccati equation:

$$\begin{aligned}\alpha &= e_{n-1}'A_0'P_0A_0e_{n-1}+1-e_{n-1}'A_0'P_0B_0(\alpha+1)^{-1}B_0'P_0A_0e_{n-1}\\ &=(\rho e_{n-1}+2e_{n-2})'P_0(\rho e_{n-1}+2e_{n-2})+1\\ &-(\rho e_{n-1}+2e_{n-2})'P_0e_{n-1}(\alpha+1)^{-1}e_{n-1}'P_0(\rho e_{n-1}+2e_{n-2})\\ &=\rho^2\alpha+4\beta+4\rho\gamma+1-\frac{(\rho\alpha+2\gamma)^2}{\alpha+1}\\ &=4\beta+\frac{\rho^2\alpha+4\rho\gamma+\alpha+1-4\gamma^2}{\alpha+1}.\end{aligned}$$

Dividing both sides with $\alpha+1$:

$$\frac{\alpha}{1+\alpha}=\frac{4}{\alpha+1}\left\{\beta-\frac{\gamma^2}{\alpha+1}\right\}+\frac{4\rho\gamma}{(\alpha+1)^2}+\frac{1+\rho^2\alpha}{(1+\alpha)^2}$$

Rearranging the terms gives:

$$\frac{2}{\alpha+1}\left\{\beta-\frac{\gamma^2}{\alpha+1}\right\}-0.5=-\frac{0.5}{1+\alpha}-\frac{2\rho\gamma}{(\alpha+1)^2}-\frac{1+\rho^2\alpha}{2(1+\alpha)^2}$$

By Lemma 7.11 the second term in the right-hand side is $o(1)$. By Lemma 7.9, $\alpha=\Omega(2^{2n})$, hence all remaining terms also go to zero, which completes the proof of (7.42).

□

Lemma 7.11. *Recall the notation in the proof of Lemma 7.10*

$$\alpha\triangleq[P_0]_{n-1,n-1},\quad \gamma\triangleq[P_0]_{n-1,n-2}.$$

Then, we have:

$$\left|\frac{\gamma}{(\alpha+1)^2}\right|=o(1)$$

Proof. We use the relation:

$$P_0 = (A_0 + B_0 K_{\star,0})' P_0 (A_0 + B_0 K_{\star,0}) + Q_0 + K_{\star,0}' R_0 K_{\star,0} \succeq K_{\star,0}' R_0 K_{\star,0}.$$

Multiplying from the left and right by e_{n-1} and by invoking (7.41) we obtain:

$$\alpha \geq \left(\frac{\rho\alpha + 2\gamma}{\alpha + 1}\right)^2 = (\xi + \lambda)^2,$$

where for simplicity we define $\xi = \frac{\rho\alpha}{\alpha+1}$, $\lambda = \frac{2\gamma}{\alpha+1}$. We can further lower bound the above expression by:

$$\alpha \geq (\xi + \lambda)^2 \geq \xi^2 + \lambda^2 - 2\xi\left|\lambda\right|.$$

This is a quadratic inequality and holds if and only if:

$$\xi - \sqrt{\alpha} \leq |\lambda| \leq \xi + \sqrt{\alpha}.$$

As a result:

$$2\frac{|\gamma|}{\alpha + 1} \leq \rho + \sqrt{\alpha + 1}$$

which leads to

$$\frac{|\gamma|}{\alpha + 1} \leq 0.5\frac{\rho + \sqrt{\alpha + 1}}{\alpha + 1} = O(1/\sqrt{\alpha}) = o(1)$$

since $\alpha = \Omega(2^{2n})$. □

Chapter 8

Conclusion and Open Problems

In this thesis, we studied the statistical complexity of learning linear systems for the tasks of system identification, learning to predict, and learning to control. In the first part of this thesis, we provided the first finite-sample analysis of stochastic system identification and the first end-to-end guarantees for the offline learning of the Kalman filter. We also provided the first logarithmic regret upper bounds for the problem of online Kalman filtering. In the second part, we studied when systems are statistically easy or hard to learn. We proved that control theoretic parameters, especially the controllability structure of the system, can affect the statistical difficulty of learning dramatically. Our approach was based on control theoretic tools as well as modern tools from statistical learning and high-dimensional statistics. Going forward, there are multiple directions for future work.

Time-varying systems/Continuous adaptation In the online control literature, most existing algorithms rely on the assumption that the unknown system is time-invariant. As a result, adaptation essentially only happens initially. If a change occurs in the system, then, the online control algorithms might not adapt fast enough. It is an interesting topic to study the regret of online control in the case of time-varying systems. There has already been some work on this topic Jadbabaie et al. (2021); Luo et al. (2022); Gradu et al. (2020).

Nonlinear systems In this thesis, we focused on linear systems. In reality, most systems are nonlinear. However, the statistical tools used in this thesis, might not be directly applicable to nonlinear systems. Recently, there has been work on developing new tools for nonlinear systems (Foster et al., 2020; Boffi et al., 2021; Mania et al., 2022; Ziemann et al., 2022). It seems that the stability properties of the nonlinear system might be crucial to provide generalization guarantees.

System Identification of Unstable Partially-Observed Systems It has been known that identification of unstable systems from single trajectory data is possible in the case of fully-observed systems under certain conditions (Faradonbeh et al., 2018a; Sarkar & Rakhlin, 2018). It is an open problem if identification of unstable systems is possible in the case of partially-observed systems. The main limitation is that the technique of unrolling and truncating the Kalman filter might no longer work in the case of single trajectory data; the initial state and the noise might have an exponentially increasing influence on the future measurements.

Learning with Structure Most of the work so far has considered unstructured problems, in the sense that A, B, C etc were assumed to be entirely unknown. In many real world applications we might have partial knowledge of the model, e.g. we might know that matrix A is sparse (Fattahi et al., 2019) or that A has a graph structure etc. It is an open problem to study how structural knowledge would affect the statistical difficulty of learning. For example, if we know the structure of matrices A, B, it might be possible to avoid exponential sample complexity, in the case of the difficult learning instances studied in Chapters 6, 7.

Learning under Constraints In this thesis, we considered unconstrained systems. In reality, systems have physical limits and should satisfy safety specifications. Optimizing performance, i.e. minimizing the LQR cost, might not be enough. It might be even more important to satisfy constraints, e.g. in a Model Predictive Control (MPC) framework.

The problem of adaptive MPC has been studied before (Bujarbaruah et al., 2019), however, understanding its statistical complexity is still an open problem. The related problem of constrained LQR control was studied before in both the offline (Dean et al., 2019) and the adaptive setting (Li et al., 2021).

Application Oriented Bounds. The sample complexity/regret guarantees in this thesis are mainly theoretical; they reveal how various system theoretic properties qualitatively affect the difficulty of learning. Our guarantees are data-independent and they might be loose due to overestimated universal constants. As a result, they might not be sharp enough to be used in applications. A different line of work focuses on application oriented bounds, e.g. data-dependent bounds, see for example Carè et al. (2018) or bootstrapping (Dean et al., 2017). Such bounds might be more suitable for applications, but they might not necessarily reveal how system properties affect learnability. It is still an open problem how to successfully utilize such bounds in real-world applications.

Bibliography

Abbasi-Yadkori, Y. and Szepesvári, C. Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th Annual Conference on Learning Theory*, pp. 1–26, 2011.

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, pp. 2312–2320, 2011.

Abbasi-Yadkori, Y., Lazic, N., and Szepesvári, C. Model-Free Linear Quadratic Control via Reduction to Expert Prediction. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 3108–3117, 2019.

Abeille, M. and Lazaric, A. Improved regret bounds for thompson sampling in linear quadratic control problems. In *International Conference on Machine Learning*, pp. 1–9, 2018.

Abeille, M. and Lazaric, A. Efficient Optimistic Exploration in Linear-Quadratic Regulators via Lagrangian Relaxation. *arXiv preprint arXiv:2007.06482*, 2020.

Anava, O., Hazan, E., Mannor, S., and Shamir, O. Online learning for time series prediction. In *Conference on learning theory*, pp. 172–184, 2013.

Anderson, B. and Moore, J. *Optimal Filtering*. Dover Publications, 2005.

Anderson, B. D. and Dehghani, A. Historical, generic and current challenges of adaptive control. *IFAC Proceedings Volumes*, 40(14):1–12, 2007.

Anderson, J., Doyle, J. C., Low, S. H., and Matni, N. System level synthesis. *Annual Reviews in Control*, 2019.

Åström, K. J. and Wittenmark, B. On self tuning regulators. *Automatica*, 9(2):185–199, 1973.

Baggio, G., Katewa, V., and Pasqualetti, F. Data-driven minimum-energy controls for linear systems. *IEEE Control Systems Letters*, 3(3):589–594, 2019.

Bai, E.-W. and Sastry, S. S. Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control. *Systems & control letters*, 6(3):153–163, 1985.

Bauer, D. and Wagner, M. Estimating cointegrated systems using subspace algorithms. *Journal of Econometrics*, 111(1):47–84, 2002.

Bauer, D., Deistler, M., and Scherrer, W. Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs. *Automatica*, 35(7):1243–1254, 1999.

Bertsekas, D. P. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 4th edition, 2017.

Boffi, N. M., Tu, S., and Slotine, J.-J. E. Regret Bounds for Adaptive Nonlinear Control. In *Learning for Dynamics and Control*, pp. 471–483. PMLR, 2021.

Boucheron, S., Lugosi, G., and Massart, P. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

Bujarbaruah, M., Zhang, X., Tanaskovic, M., and Borrelli, F. Adaptive MPC under time varying uncertainty: Robust and Stochastic. *arXiv preprint arXiv:1909.13473*, 2019.

Campi, M. C. and Weyer, E. Finite sample properties of system identification methods. *IEEE Transactions on Automatic Control*, 47(8):1329–1334, 2002.

Carè, A., Csáji, B. C., Campi, M. C., and Weyer, E. Finite-sample system identification: An overview and a new correlation method. *IEEE Control Systems Letters*, 2(1):61–66, 2018.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, learning, and games*. Cambridge university press, 2006.

Chan, S., Goodwin, G., and Sin, K. Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems. *IEEE Transactions on Automatic Control*, 29(2):110–118, 1984.

Chen, X. and Hazan, E. Black-Box Control for Linear Dynamical Systems. In *Conference on Learning Theory*, pp. 1114–1143. PMLR, 2021.

Chiuso, A. and Picci, G. The asymptotic variance of subspace estimates. *Journal of Econometrics*, 118(1-2):257–291, 2004.

Chiuso, A. and Pillonetto, G. System identification: A machine learning perspective. *Annual Review of Control, Robotics, and Autonomous Systems*, 2(1), 2019.

Cohen, A., Hasidim, A., Koren, T., Lazic, N., Mansour, Y., and Talwar, K. Online Linear Quadratic Control. In *International Conference on Machine Learning*, pp. 1029–1038. PMLR, 2018.

Cohen, A., Koren, T., and Mansour, Y. Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. *arXiv preprint arXiv:1902.06223*, 2019.

Coskun, H., Achilles, F., DiPietro, R., Navab, N., and Tombari, F. Long short-term memory Kalman filters: Recurrent neural estimators for pose regularization. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5524–5532, 2017.

Dahleh, M. A., Theodosopoulos, T. V., and Tsitsiklis, J. N. The sample complexity of worst-case identification of FIR linear systems. In *Proceedings of the 32nd IEEE Conference on Decision and Control, 1993.*, pp. 2082–2086, 1993.

Dann, C., Lattimore, T., and Brunskill, E. Unifying PAC and regret: Uniform PAC bounds for episodic reinforcement learning. *arXiv preprint arXiv:1703.07710*, 2017.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. On the sample complexity of the linear quadratic regulator. *arXiv preprint arXiv:1710.01688*, 2017.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. Regret bounds for robust adaptive control of the linear quadratic regulator. In *Advances in Neural Information Processing Systems*, pp. 4188–4197, 2018.

Dean, S., Tu, S., Matni, N., and Recht, B. Safely Learning to Control the Constrained Linear Quadratic Regulator. In *2019 American Control Conference (ACC)*, pp. 5582–5588. IEEE, 2019.

Deistler, M., Peternell, K., and Scherrer, W. Consistency and relative efficiency of subspace methods. *Automatica*, 31(12):1865–1875, 1995.

Ding, F., Shi, Y., and Chen, T. Performance analysis of estimation algorithms of non-stationary ARMA processes. *IEEE Transactions on Signal Processing*, 54(3):1041–1053, 2006.

Dooren, P. M. V. Numerical linear algebra for signals systems and control. *Draft notes prepared for the Graduate School in Systems and Control*, 2003.

Du, S. S., Kakade, S. M., Wang, R., and Yang, L. F. Is a good representation sufficient for sample efficient reinforcement learning? *arXiv preprint arXiv:1910.03016*, 2019.

Durrant-Whyte, H. and Bailey, T. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110, 2006.

Efroni, Y., Kakade, S., Krishnamurthy, A., and Zhang, C. Sparsity in Partially Controllable Linear Systems. *arXiv preprint arXiv:2110.06150*, 2021.

Eising, R. Between controllable and uncontrollable. *Systems & Control Letters*, 4(5):263–264, 1984.

El Ghaoui, L. and Calafiore, G. Robust filtering for discrete-time systems with bounded noise and parametric uncertainty. *IEEE Transactions on Automatic Control*, 46(7):1084–1089, 2001.

Fadali, M. S. and Visioli, A. *Digital Control Engineering: Analysis and Design.* Academic Press, 2013.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. Finite time identification in unstable linear systems. *Automatica*, 96:342–353, 2018a.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. Finite-time Adaptive Stabilization of Linear Systems. *IEEE Transactions on Automatic Control*, 64(8):3498–3505, 2018b.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. On Adaptive Linear–Quadratic Regulators. *Automatica*, 117:108982, 2020a.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. Optimism-based adaptive regulation of linear-quadratic systems. *IEEE Transactions on Automatic Control*, 66(4):1802–1808, 2020b.

Fattahi, S., Matni, N., and Sojoudi, S. Learning sparse dynamical systems from a single sample trajectory. *arXiv preprint arXiv:1904.09396*, 2019.

Foster, D., Sarkar, T., and Rakhlin, A. Learning nonlinear dynamical systems from a single trajectory. In *Learning for Dynamics and Control*, pp. 851–861. PMLR, 2020.

Gevers, M. Identification for Control: From the Early Achievements to the Revival of Experiment Design. *European journal of control*, 11(4-5):335–352, 2005.

Ghai, U., Lee, H., Singh, K., Zhang, C., and Zhang, Y. No-regret prediction in marginally stable systems. In *Conference on Learning Theory*, pp. 1714–1757. PMLR, 2020.

Goodwin, G. C., Ramadge, P. J., and Caines, P. E. Discrete time stochastic adaptive control. *SIAM Journal on Control and Optimization*, 19(6):829–853, 1981.

Gradu, P., Hazan, E., and Minasyan, E. Adaptive Regret for Control of Time-Varying Dynamics. *arXiv preprint arXiv:2007.04393*, 2020.

Hardt, M., Ma, T., and Recht, B. Gradient descent learns linear dynamical systems. *Journal of Machine Learning Research*, 19(29):1–44, 2018.

Harvey, A. C. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

Hazan, E., Lee, H., Singh, K., Zhang, C., and Zhang, Y. Spectral filtering for general linear dynamical systems. In *Advances in Neural Information Processing Systems*, pp. 4634–4643, 2018.

Heaton, J. B., Polson, N. G., and Witte, J. H. Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1):3–12, 2017.

Horn, R. A. and Johnson, C. R. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, 2 edition, 2012.

Jadbabaie, A., Mania, H., Shah, D., and Sra, S. Time varying regression with hidden linear dynamics. *arXiv preprint arXiv:2112.14862*, 2021.

Jaksch, T., Ortner, R., and Auer, P. Near-optimal Regret Bounds for Reinforcement Learning. *Journal of Machine Learning Research*, 11:1563–1600, 2010.

Jansson, M. and Wahlberg, B. On consistency of subspace methods for system identification. *Automatica*, 34(12):1507–1519, 1998.

Jedra, Y. and Proutiere, A. Sample complexity lower bounds for linear system identification. In *IEEE 58th Conference on Decision and Control (CDC)*, pp. 2676–2681. IEEE, 2019.

Jedra, Y. and Proutiere, A. Minimal Expected Regret in Linear Quadratic Control. *arXiv preprint arXiv:2109.14429*, 2021.

Jiang, N., Krishnamurthy, A., Agarwal, A., Langford, J., and Schapire, R. E. Contextual decision processes with low Bellman rank are PAC-learnable. In *International Conference on Machine Learning*, pp. 1704–1713. PMLR, 2017.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.

Kailath, T., Sayed, A. H., and Hassibi, B. *Linear estimation*. Prentice Hall, 2000.

Kakade, S., Krishnamurthy, A., Lowrey, K., Ohnishi, M., and Sun, W. Information Theoretic Regret Bounds for Online Nonlinear Control. *Advances in Neural Information Processing Systems*, 33:15312–15325, 2020.

Kalman, R. E. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(1):35–45, 1960.

Knudsen, T. Consistency analysis of subspace identification methods based on a linear regression approach. *Automatica*, 37(1):81–89, 2001.

Kozdoba, M., Marecek, J., Tchrakian, T., and Mannor, S. On-line learning of linear dynamical systems: Exponential forgetting in kalman filters. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 4098–4105, 2019.

Krahmer, F., Mendelson, S., and Rauhut, H. Suprema of chaos processes and the restricted isometry property. *Communications on Pure and Applied Mathematics*, 67(11):1877–1904, 2014.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

Kulkarni, D., Schmidt, D., and Tsui, S.-K. Eigenvalues of tridiagonal pseudo-Toeplitz matrices. *Linear Algebra and its Applications*, 297(1):63–80, 1999.

Lai, T. L. and Wei, C. Z. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics*, 10(1):154–166, 1982.

Lai, T. L. and Ying, Z. Recursive identification and adaptive prediction in linear stochastic systems. *SIAM Journal on Control and Optimization*, 29(5):1061–1090, 1991.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Explore more and improve regret in Linear Quadratic Regulators. *arXiv preprint arXiv:2007.12291*, 2020a.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Logarithmic regret bound in partially observable linear dynamical systems. *arXiv preprint arXiv:2003.11227*, 2020b.

Lee, B. and Lamperski, A. Non-asymptotic Closed-Loop System Identification using Autoregressive Processes and Hankel Model Reduction. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 3419–3424. IEEE, 2020.

Lee, H. Improved rates for identification of partially observed linear dynamical systems. *arXiv preprint arXiv:2011.10006*, 2020.

Lee, H. and Zhang, C. Robust Guarantees for Learning an Autoregressive Filter. In *Algorithmic Learning Theory*, pp. 490–517. PMLR, 2020.

Levine, S., Finn, C., Darrell, T., and Abbeel, P. End-to-End Training of Deep Visuomotor Policies. *The Journal of Machine Learning Research*, 17(1):1334–1373, 2016.

Levy, B. C. and Nikoukhah, R. Robust state space filtering under incremental model perturbations subject to a relative entropy tolerance. *IEEE Transactions on Automatic Control*, 58(3):682–695, 2012.

Li, Y., Das, S., Shamma, J., and Li, N. Safe Adaptive Learning-based Control for Constrained Linear Quadratic Regulators with Regret Guarantees. *arXiv preprint arXiv:2111.00411*, 2021.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Ljung, L. Convergence of an adaptive filter algorithm. *International Journal of Control*, 27 (5):673–693, 1978.

Ljung, L. *System Identification: Theory for the User*. Prentice Hall, 1999.

Ljung, L. Perspectives on System identification. *Annual Reviews in Control*, 34(1):1–12, 2010.

Lu, Y. and Mo, Y. Safe Linear-Quadratic Dual Control with Almost Sure Performance Guarantee. *arXiv preprint arXiv:2103.13278*, 2021.

Luo, Y., Gupta, V., and Kolar, M. Dynamic Regret Minimization for Control of Non-stationary Linear Dynamical Systems. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 6(1):1–72, 2022.

Mania, H., Tu, S., and Recht, B. Certainty Equivalence is Efficient for Linear Quadratic Control. *arXiv preprint arXiv:1902.07826*, 2019.

Mania, H., Jordan, M. I., and Recht, B. Active Learning for Nonlinear System Identification with Guarantees. *Journal of Machine Learning Research*, 23(32):1–30, 2022.

Matni, N. and Tu, S. A tutorial on concentration bounds for system identification. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 3741–3749. IEEE, 2019.

Matni, N., Proutiere, A., Rantzer, A., and Tu, S. From self-tuning regulators to reinforcement learning and back again. *arXiv preprint arXiv:1906.11392*, 2019.

Mehra, R. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on automatic control*, 15(2):175–184, 1970.

Moden, E. Experiences with adaptive control since 1982. In *Proceedings of 1995 34th IEEE Conference on Decision and Control*, volume 1, pp. 667–672. IEEE, 1995.

Moore, J. and Ledwich, G. Multivariable adaptive parameter and state estimators with convergence analysis. *The ANZIAM Journal*, 21(2):176–197, 1979.

Ortner, R. and Ryabko, D. Online Regret Bounds for Undiscounted Continuous Reinforcement Learning. *Advances in Neural Information Processing Systems*, 25, 2012.

Ouyang, Y., Gagrani, M., and Jain, R. Control of Unknown Linear Systems with Thompson Sampling. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1198–1205. IEEE, 2017a.

Ouyang, Y., Gagrani, M., and Jain, R. Learning-based control of unknown linear systems with Thompson sampling. *arXiv preprint arXiv:1709.04047*, 2017b.

Oymak, S. and Ozay, N. Non-asymptotic Identification of LTI Systems from a Single Trajectory. *arXiv preprint arXiv:1806.05722*, 2018.

Pernebo, L. and Silverman, L. Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382–387, 1982.

Peternell, K., Scherrer, W., and Deistler, M. Statistical analysis of novel subspace identification methods. *Signal Processing*, 52(2):161–177, 1996.

Plevrakis, O. and Hazan, E. Geometric Exploration for Online Control. *Advances in Neural Information Processing Systems*, 33:7637–7647, 2020.

Poolla, K. and Tikku, A. On the Time Complexity of Worst-Case System Identification. *IEEE Transactions on Automatic Control*, 39(5):944–950, 1994.

Qin, S. J. An overview of subspace identification. *Computers & chemical engineering*, 30 (10-12):1502–1513, 2006.

Rantzer, A. Concentration bounds for single parameter adaptive control. In *2018 Annual American Control Conference (ACC)*, pp. 1862–1866. IEEE, 2018.

Rashidinejad, P., Jiao, J., and Russell, S. SLIP: Learning to Predict in Unknown Dynamical Systems with Long-Term Memory. In *34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada*, 2020.

Recht, B. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2(1):253–279, 2019.

Sarkar, T. and Rakhlin, A. Near optimal finite time identification of arbitrary linear dynamical systems. *arXiv preprint arXiv:1812.01251*, 2018.

Sarkar, T., Rakhlin, A., and Dahleh, M. A. Finite-Time System Identification for Partially Observed LTI Systems of Unknown Order. *arXiv preprint arXiv:1902.01848*, 2019.

Sayed, A. H. et al. A framework for state-space estimation with uncertain models. *IEEE Transactions on Automatic Control*, 46(7):998–1013, 2001.

Shalev-Shwartz, S. and Ben-David, S. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. Mastering the game of Go without human knowledge. *nature*, 550(7676):354–359, 2017.

Simchowitz, M. and Foster, D. J. Naive Exploration is Optimal for Online LQR. *arXiv preprint arXiv:2001.09576*, 2020.

Simchowitz, M., Mania, H., Tu, S., Jordan, M. I., and Recht, B. Learning Without Mixing: Towards A Sharp Analysis of Linear System Identification. *arXiv preprint arXiv:1802.08334*, 2018.

Simchowitz, M., Boczar, R., and Recht, B. Learning Linear Dynamical Systems with Semi-Parametric Least Squares. *arXiv preprint arXiv:1902.00768*, 2019.

Skogestad, S., Morari, M., and Doyle, J. C. Robust Control of Ill-Conditioned Plants: High-Purity Distillation. *IEEE transactions on automatic control*, 33(12):1092–1105, 1988.

Talebi, S., Alemzadeh, S., Rahimi, N., and Mesbahi, M. On Regularizability and its Application to Online Control of Unstable LTI Systems. *IEEE Transactions on Automatic Control*, 2021.

Tsiamis, A. and Pappas, G. Online learning of the Kalman filter with logarithmic regret. *arXiv preprint arXiv:2002.05141*, 2020.

Tsiamis, A. and Pappas, G. J. Finite Sample Analysis of Stochastic System Identification. In *IEEE 58th Conference on Decision and Control (CDC)*, 2019.

Tsiamis, A. and Pappas, G. J. Linear Systems can be Hard to Learn. *arXiv preprint arXiv:2104.01120*, 2021.

Tsiamis, A., Matni, N., and Pappas, G. Sample Complexity of Kalman Filtering for Unknown Systems. In *Learning for Dynamics and Control*, pp. 435–444. PMLR, 2020.

Tsiamis, A., Ziemann, I., Morari, M., Matni, N., and Pappas, G. J. Learning to Control Linear Systems Can Be Hard, 2022. Submitted.

Tu, S., Boczar, R., Simchowitz, M., Soltanolkotabi, M., and Recht, B. Low-rank solutions of linear matrix equations via procrustes flow. In *International Conference on Machine Learning*, pp. 964–973, 2016.

Tu, S., Boczar, R., Packard, A., and Recht, B. Non-Asymptotic Analysis of Robust Control from Coarse-Grained Identification. *arXiv preprint arXiv:1707.04791*, 2017.

Van Overschee, P. and De Moor, B. A Unifying Theorem for Three Subspace System Identification Algorithms. *Automatica*, 31(12):1853–1864, 1995.

Van Overschee, P. and De Moor, B. *Subspace identification for linear systems: Theory–Implementation–Applications*. Springer Science & Business Media, 2012.

Verhaegen, M. and Verdult, V. *Filtering and system identification: a least squares approach*. Cambridge university press, 2007.

Vershynin, R. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.

Viberg, M., Wahlberg, B., and Ottersten, B. Analysis of state space system identification methods based on instrumental variables and subspace fitting. *Automatica*, 33(9):1603–1616, 1997.

Vidyasagar, M. and Karandikar, R. L. A learning theory approach to system identification and stochastic adaptive control. *Journal of Process Control*, 18(3-4):421–430, 2008.

Wagenmaker, A. and Jamieson, K. Active learning for identification of linear dynamical systems. In *Conference on Learning Theory*, pp. 3487–3582. PMLR, 2020.

Wang, F. and Janson, L. Exact Asymptotics for Linear Quadratic Adaptive Control. *Journal of Machine Learning Research*, 22(265):1–112, 2021.

Wang, Y.-S., You, S., and Matni, N. Localized distributed Kalman filters for large-scale systems. *IFAC-PapersOnLine*, 48(22):52–57, 2015.

Wang, Y.-S., Matni, N., and Doyle, J. C. A system level approach to controller synthesis. *IEEE Transactions on Automatic Control*, 2019.

Wedin, P.-Å. Perturbation bounds in connection with singular value decomposition. *BIT Numerical Mathematics*, 12(1):99–111, 1972.

Wedin, P.-Å. Perturbation theory for pseudo-inverses. *BIT Numerical Mathematics*, 13(2): 217–232, 1973.

Weyer, E., Williamson, R. C., and Mareels, I. M. Finite sample properties of linear model identification. *IEEE Transactions on Automatic Control*, 44(7):1370–1383, 1999.

Young, T., Hazarika, D., Poria, S., and Cambria, E. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13(3):55–75, 2018.

Yu, C., Ljung, L., and Verhaegen, M. Identification of structured state-space models. *Automatica*, 90:54–61, 2018.

Zhang, S., Yao, L., Sun, A., and Tay, Y. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, 52(1):1–38, 2019.

Zheng, Y. and Li, N. Non-Asymptotic Identification of Linear Dynamical Systems Using Multiple Trajectories. *IEEE Control Systems Letters*, 5(5):1693–1698, 2020.

Zheng, Y., Furieri, L., Kamgarpour, M., and Li, N. Sample Complexity of Linear Quadratic Gaussian (LQG) Control for Output Feedback Systems. In *Learning for Dynamics and Control*, pp. 559–570. PMLR, 2021.

Ziemann, I. and Sandberg, H. On a phase transition of regret in linear quadratic control: The memoryless case. *IEEE Control Systems Letters*, 5(2):695–700, 2020.

Ziemann, I. and Sandberg, H. Regret Lower Bounds for Learning Linear Quadratic Gaussian Systems. *arXiv preprint arXiv:2201.01680*, 2022.

Ziemann, I., Sandberg, H., and Matni, N. Single Trajectory Nonparametric Learning of Nonlinear Dynamics. *arXiv preprint arXiv:2202.08311*, 2022.