Safe and Scalable Learning-based Control:
Theory and Application in Sustainable Energy Systems

Thesis by
Jing Yu

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2025
Defended July 19, 2024

© 2025

Jing Yu
ORCID: 0000-0003-1318-0189

To Andrea,

ACKNOWLEDGEMENTS

I want to thank my wonderful advisors John Doyle and Adam Wierman, without whom this thesis would not have been possible. John has been incredibly supportive throughout my PhD, letting me freely explore research directions and collaborate across various groups. He pushes me to think outside of the box and become a more independent researcher. Our discussions have always provided me with fresh perspectives and ideas. Adam has been a constant source of wisdom. He offers technical guidance, calls out my shaky arguments, and stands by our results when I am quick to dismiss progress. Most importantly, I thank Adam for showing me the power of kindness and mentorship. In many critical moments of my PhD, Adam's encouragement was the lighthouse in the midst of the uncertainties in life. I would also like to thank my thesis committee members Yisong Yue and Steven Low, who have always made themselves available to provide helpful insights and feedback.

I have been extremely fortunate to have worked with many awesome collaborators. My first paper at Caltech was written with Dimitar Ho, who has continued to be an inspiring friend, mentor, and collaborator ever since. I met Yuh-shyang Wang during my internship at the GE Research Center in my second year of PhD. With his brilliant insights and together with the guidance of James Anderson, we wrote up the first infinite-horizon result for system level synthesis. The collaboration with Florian Schaefer, Clement Gehring, and Anima Anandkumar on competitive mirror descent was the result of many, perhaps too many, kitchen table conversations, where Florian relentlessly discussed his research and answered my never-ending questions. This project would have been a complete flop if not for Clement. With my clueless attempts in writing a semi-functional package and wobbly understanding of object-oriented programming, Clement painstakingly guided me through the entire computational part of this project. Thank you Clement for your incredible patience, rigor, and kindness. I can only hope to become the same kind of collaborator for others as you have been for me. It was a pleasure to work with Christopher Yeh and Yuanyuan Shi on online control for voltage regulation, where I learnt so much about the distribution grid and key challenges in sustainable energy systems. Chris has been such a reliable and fun person to work with. The great synergy in this collaboration was one of the reasons that motivated me to pursue an academic career. I also had a blast collaborating and hanging out with Olle Kjellqvist when he visited Caltech from Lund. I loved our discussions on research and life, especially over

ABSTRACT

From intelligent transportation systems to the smart grid, the next generation of cyber-physical systems (CPS) will substantially transform our society. It is vital that these systems are scalable and robust to uncertainties, with contextual awareness and fast adaptation. This dissertation presents progress towards addressing key challenges arising in the control of large-scale CPS, with a special focus on applications in sustainable energy systems.

Large-scale CPS such as the smart grid often consist of numerous interconnected and heterogeneous subsystems that must coordinate to achieve global objectives by exchanging information over a communication network. Therefore, the first part of this thesis focuses on developing control algorithms that handle crucial design requirements emerging from scalability and communication constraints, such as disturbance localization, communication delay conformation, and distributed implementation.

Sustainable energy systems are crucial for reducing greenhouse gas emissions and mitigating climate change. However, the inherent unpredictability and large uncertainties associated with renewable generation pose significant challenges for maintaining system stability and safety. Traditional control approaches, while robust and effective for known system models, often fall short when faced with the dynamic and uncertain nature of modern power systems. In the second part of the thesis, we address this challenge by integrating machine learning techniques with model-based control methods using uncertainty sets constructed from real-time data. In particular, we will introduce and provide convergence guarantees for a classic uncertainty set estimation method. Building on these uncertainty sets, we combine learning and control techniques to tackle core CPS control problems, such as adversarial stability certification for linear time-varying systems as well as networked systems under communication constraints where the system models are unknown.

The final part of this thesis applies the developed methodologies to address the voltage control problem in power distribution networks with unknown grid topologies. We will combine online learning techniques and a robust predictive controller to achieve provably finite-time convergence to safe voltage limits, despite uncertainties in network topology and load variations. Our case study on a Southern California Edison 56-bus distribution system demonstrates the effectiveness of this approach in nonlinear, partial observation, and partial control settings.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Y. Li*, J. Yu*, L. Conger, T. Kargin, and A. Wierman, "Learning the uncertainty sets of linear control systems via set membership: A non-asymptotic analysis," *Forty-first International Conference on Machine Learning (ICML)*, 2024. [Online]. Available: https://openreview.net/forum?id=n2kq2EOHFE.
J.Y. participated in the conception of the project and the development of the theoretical results, conceived and conducted the simulations, and co-led the writing of the manuscript.

[2] C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Online learning for robust voltage control under uncertain grid topology," *IEEE Transactions on Smart Grid*, 2024. DOI: 10.1109/TSG.2024.3383804.
J.Y. participated in the the conception of the project and development of the theoretical results, conceived and implemented parts of the simulations, and particiated the writing of the manuscript.

[3] J. Yu, V. Gupta, and A. Wierman, "Online adversarial stabilization of unknown linear time-varying systems," *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 8320–8327, 2023. DOI: 10.1109/CDC49753.2023.1038384.
J.Y. led the conception of the project and the development of the theoretical results and algorithms, conceived and conducted the simulations, and led the writing of the manuscript.

[4] J. Yu, D. Ho, and A. Wierman, "Online adversarial stabilization of unknown networked systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems (SIGMETRICS)*, vol. 7, no. 1, pp. 1–43, 2023. DOI: 10.1145/3579452.
J.Y. led the conception of the project and the development of the theoretical results and algorithms, conceived and conducted the simulations, and led the writing of the manuscript.

[5] O. Kjellqvist* and J. Yu*, "On infinite-horizon system level synthesis problems," *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 5238–5244, 2022. DOI: 10.1109/CDC51059.2022.9992443.
J.Y. co-led the conception of the project and the development of the theoretical results and algorithms, conceived and conducted the simulations, and co-led the writing of the manuscript.

[6] C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Robust online voltage control with an unknown grid topology," *Proceedings of the Thirteenth ACM international conference on future energy systems (e-Energy)*, pp. 240–250, 2022. DOI: 10.1145/3538637.3538853.

J.Y. participated in the development of the theoretical results and the writing of the manuscript.

[7] J. Yu, C. Gehring, F. Schäfer, and A. Anandkumar, "Robust reinforcement learning: A constrained game-theoretic approach," *Learning for Dynamics and Control (L4DC)*, pp. 1242–1254, 2021. [Online]. Available: `https://proceedings.mlr.press/v144/yu21a/yu21a.pdf`.
J.Y. participated in the conception of the project and the development of the theoretical results and algorithms, conducted the simulations, and led the writing of the manuscript.

[8] J. Yu, Y.-S. Wang, and J. Anderson, "Localized and distributed $\mathcal{H}_2$ state feedback control," *2021 American Control Conference (ACC)*, pp. 2732–2738, 2021. DOI: `10.23919/ACC50511.2021.9483301`.
J.Y. participated in the conception of the project and the development of the theoretical results and algorithms, conceived and conducted the simulations, and participated in the writing of the manuscript.

[9] J. Yu* and D. Ho*, "Achieving performance and safety in large scale systems with saturation using a nonlinear system level synthesis approach," *2020 American Control Conference (ACC)*, pp. 968–973, 2020. DOI: `10.23919/ACC45564.2020.9147577`.
J.Y. participated in the conception of the project and the development of the theoretical results, conceived and conducted the simulations, and co-led the writing of the manuscript.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

From intelligent transportation systems to the smart grid, the next generation of cyber-physical systems (CPS) will reshape our society. Control theory provides principled frameworks to design robust, distributed, and performant controllers, ensuring worst-case stability and constraint satisfaction guarantees for these safety-critical CPS. However, control methods often rely on accurate model information, which is increasingly difficult to obtain due to the growing complexity, scale, and uncertainty in CPS. For example, smart grid topology can change unpredictably due to renewable energy supply variations or line failures, while in robotics, differing environmental conditions lead to distinct system dynamics.

On the other hand, recent advancements in machine learning (ML) open up exciting opportunities to transform current control approaches to data-driven methods with unprecedented performance, contextual awareness, and flexible adaptation. In particular, there is now a growing body of research that investigates the incorporation of ML techniques in control design, often referred to as learning-based control, to enable efficient algorithms for complex systems whose dynamics are hard to model or even unknown.

Despite recent theoretical developments and successful applications of ML in various domains including games [1], robotics [2], [3], and more recently large language models [4], principled integration of learning in control tasks for CPS remain open due to the *safety-critical*, *large-scale*, *networked*, and *resource-constrained* nature of CPS:

- **Safety-critical**: Safety assumptions in the application of ML in highly-controlled environments, such as bounded system behavior and unconstrained exploration, are no longer practical for CPS due to physical limitations, operational requirements, and the feedback effects of algorithms deployed in the closed loop. For instance, a common assumption in the learning-based control literature is that the learning algorithm has access to a known stabilizing controller. This assumption ensures bounded and, therefore, safe system behavior throughout the learning phase, sidestepping the main technical

hurdle of safe exploration in uncertain environments. Unfortunately, such an assumption is often unrealistic, as stabilization itself, a prerequisite for any other performance objectives, is a significant task even if the dynamics model is available. For example, the voltage stabilization problem remains challenging in inverter-based resources dominated microgrids even when the network is known. It remains a grand challenge to deploy data-driven methods in critical tasks where incorrect decisions can have catastrophic consequences.

- **Large-scale and networked**: Centralized control algorithms are often prohibitively expensive to design and execute for large-scale networked CPS, such as robotic swarms and the smart grid. Moreover, centralized schemes can expose the system to targeted attacks, creating disastrous vulnerabilities. Therefore, significant efforts have been made to develop distributed controllers for such systems, where sub-controllers leverage local information and computational resources from small neighborhoods in the network and coordinate among one another to achieve system-level objectives. On the other hand, most ML algorithms perform centralized data collection and learning. If learning-based algorithms are to be deployed in real-world systems, it is imperative to adapt ML to distributed design.

- **Resource-constrained** A key characteristic of large-scale CPS is that the subsystems in the network often face significant resource constraints, including limited computational power, memory, and communication. Even with distributed algorithms, these limitations persist, leading to further implementation challenges. For instance, communication delays can occur due to restricted communication bandwidth, particularly when substantial volumes of data are exchanged within the network, as seen in data centers. Therefore, to achieve robust and reliable CPS operation, it is essential to consider these resource constraints and the associated implementation challenges during algorithm design and understand the impact of such constraints.

The goal of this thesis is to provide a unified learning and control framework to broadly addresses key challenges arising in the control of large-scale CPS, with a special focus on applications in sustainable energy systems. In the next sections, we highlight important problems, survey related work, and provide an outline for the thesis.

1.1 Major Challenges and Prior Work

Significant progress has been made in algorithm design for large-scale CPS, especially in the area of learning-based control. Despite the many milestones in the recent years, we are still in the early stages of developing reliable, scalable, and practical controllers for future CPS, particularly in the following key problems.

Adversarially Safe Control

Safety in CPS is often modeled by defining a set of permissible states and actions that the system can undertake during operation. A minimum safety criteria is to maintain system stability, where all internal states of the system remain bounded or converge to a desired steady state despite exogenous perturbations. Ensuring safety is of paramount importance to critical CPS infrastructures. However, it has become increasingly challenging due to various uncertainties stemming from the growing complexity, scale, and technological advancements in modern CPS. For instance, intelligent transportation systems have humans in the loop with strategic behaviors that introduce additional variability and unpredictability. The power grid now faces more frequent and volatile changes due to the rising penetration of distributed renewable energy sources. Thus, how to ensure safety despite complex and dynamic uncertainties is a core challenge.

Handling uncertainty is one of the main goals of classical control design, commonly referred to as robust control [5]–[7]. Robust control is often formulated to achieve stability and performance guarantees in the worst-case under all possible uncertainties. A canonical example is to design a simultaneously stabilizing controller for an entire set of uncertain models. Naturally, there does not exist such a controller if the model uncertainty set is too large. Thus, a fundamental challenge of such offline worst-case control design is that the conservativeness limits the feasibility of these problems to only "small" uncertainty sets. To tackle it, adaptive and online control literature incorporates online data to significantly reduce the conservativeness [8]–[11]. Inspired by this line of work, recent progress has been made towards a promising approach that integrates robust control and machine learning-based online adaptation to tackle large uncertainties in the system model [12]–[15].

Another common uncertainty modeling is to introduce disturbances into the system. It is common to model disturbances as stochastic, with simplified distributions that are amenable to theoretical analysis [16]–[21]. However, such stochastic modeling does not fully capture many realistic uncertainties such as discretization errors as

a result of sampling-based digital control, unmodeled nonlinearities, or malicious attacks. Therefore, an alternative approach is to treat disturbances as deterministic and arbitrary from a fixed admissible disturbance set, e.g., [22], [23]. However, these problems do not simultaneously consider large model uncertainties and adversarial disturbances.

Despite much progress, the fundamental problem of adversarial stability guarantee for the simplest setting, i.e., linear dynamical systems with *arbitrarily large model uncertainty* under *adversarial disturbances*, remained elusive. In this thesis, we will address this gap with a unified framework that incorporate ML techniques with robust control design to enable robustness and safety against large uncertainties commonly seen in CPS. Still, many open questions stand, such as heterogeneous and imperfect sensing for decision making with adversarial safety. In particular, there is a gap in our grasp of adversarial robustness in control dynamical systems where uncertainties and attacks in the perception module lead to compounding errors due to feedback effects. Another important question is how we should define and provide adversarial safety for human-in-the-loop multi-agent systems. The methodologies and insights from this thesis will hopefully serve as a step towards a principled understanding in adversarially safe control design.

Learn and Control with Uncertainty Sets

While adversarial guarantees are crucial for safety-critical CPS, it is well known that algorithms designed to be robust against adversarial attacks and worst-case uncertainties can be overly conservative and costly in typical, non-adversarial scenarios. On the other hand, recent advancements in ML have enabled unprecedented performance in various domains, but often overlook rare yet plausible worst-case scenarios, leaving the system open to potentially catastrophic failures. Integrating learning and robust algorithms with simultaneous worst-case safety and average-case performance guarantees is an essential step towards real-world applications.

To enable learning-based control methods with such capabilities, a popular approach is to apply ML techniques for estimation of the system model and parameters that are relevant for the control tasks, then apply model-based control using the estimated model [24]–[28]. It is therefore critical to quantify the uncertainty of the estimations in order for the algorithm to robustly satisfy safety constraints despite these uncertainties [29], [30]. On the one hand, an uncertainty set that is too large gives rise to over-conservative control actions, resulting in degraded performance.

On the other hand, if the uncertainty set is underestimated and fails to contain the true system model, the resulting controller may lead to unsafe behaviors [30], [31]. Therefore, understanding how we leverage sequentially revealed online data to learn non-conservative uncertainty sets that guarantee the containment of the true model is a key step in the integration of ML and control.

Furthermore, given an uncertainty set, it remains unclear how to design controllers based on the uncertainty set, as most control methods are based on a single model rather than a set of models. A common approach is to select a point from the uncertainty set for downstream model-based control design. In particular, the selection must balance between exploitation (select models that are close to the true model) and exploration (select models such that the control actions generated based on the selection reduce the size of the uncertainty set) tradeoff. While a similar tradeoff has been extensively studied in online learning (see [32] and the references therein), little has been explored in the control and dynamical system setting.

Scalable Control and Learning under Communication Constraints
New developments in distributed sensing, communication, and computation technologies, such as Internet of Things and edge computing, create novel opportunities for distributed algorithm design for CPS that are efficient, scalable, reconfigurable, and adaptive [33]–[36]. While many of the algorithms have seen success in high-fidelity simulations and industrial implementation, there is a lack of principled understanding of potential vulnerabilities and limitations of the algorithms through theoretical analysis. In particular, many realistic considerations must be taken into account. One major challenge is that communication constraints, e.g., resulting from communication delay or privacy considerations, impose structural constraints on algorithm design. It is known that even for simpler classes of policies such as linear distributed controllers, such constraints are generally non-convex [37]. For general controllers, structured optimal control design is intractable [38].

Recent work has focused on special cases of distributed control problems, such as those satisfying Quadratic Invariance (QI) [39], where exact convex reformulation can be found. Under QI, substantial progress was made towards control design that can handle constraints such as communication delay and distributed implementation [40]–[43]. However, QI requires global information exchange for strongly connected networks. This imposes major scalability issues on the synthesis procedures based on QI and the implementation of the resulting distributed controllers. In particular, it

was observed in [44] that controllers leveraging QI can be more complex to synthesize than their central counterparts. In order to move beyond QI, researchers have focused on novel controller parameterization that admits convex formulation [45]–[47]. In this thesis, we focus on an important class of structured control design problems with communication delay and localization constraints. We will develop novel optimal and robust controllers with fully distributed synthesis and implementation.

Even with distributed and scalable control algorithms that conform to various communication constraints, it is not straightforward to integrate ML into the design loop for systems to handle large uncertainties, e.g., power grid with heterogeneous energy storage and high renewables penetration. In particular, the sheer scale and communication constraints in CPS, such as privacy concerns, often render it impossible to carry out centralized learning and data collection. However, scalability and communication constraints have only been considered separately; no general approach exists in learning-based control literature. Bridging this gap will provide crucial insights for a unified framework for scalable learning and control algorithms of CPS.

1.2 Thesis Roadmap and Contributions

Motivated by the key challenges outlined above, the central mission statement of this thesis is:

> *To design **scalable** controllers that are capable of **learning and adaptation**, with simultaneous worst-case **safety** and average-case **performance guarantee**, all while conforming to **communication constraints**.*

The remainder of this thesis is structured as follows.

Part I: Distributed Control under Communication Constraints

Large-scale CPS such as the smart grid often consist of numerous interconnected and heterogeneous subsystems that must coordinate to achieve global objectives. One of the key challenges in managing such large-scale CPS is the effective design and implementation of distributed controllers that can operate under various communication constraints while allowing scalable synthesis and implementation.

In this context, the first part of this thesis focuses on developing distributed controllers to address important classes of communication constraints when the system model is

known. Specifically, we will concentrate on three essential design requirements that naturally emerge due to communication constraints in large-scale CPS:

1. *Localization*: It is desirable that the effects of disturbances as well as information exchange among subsystems are limited to a predefined local region without cascading to the global network. By confining the impact of disturbances, we can enhance the robustness and reliability of the entire system. Disturbance localization also facilitates localized communication, where it becomes unnecessary for subsystems that are far away from one another to exchange information.

2. *Communication delay*: Controllers for each subsystem can only access information that may be delayed from other subsystems according to the structure of a given communication network. Handling communication delays is crucial for maintaining system stability and performance, as delays can significantly impact the closed-loop stability.

3. *Distributed implementation*: Controller implementation needs to be distributed, allowing only sparse information to be exchanged between subsystems within a local region to reduce communication and computation burdens on local systems. A distributed approach not only alleviates the load on communication networks but also enhances the scalability and flexibility of the system, enabling it to adapt to varying system topologies.

In Chapter 2, we will present the optimal distributed $\mathcal{H}_2$ linear state-feedback controllers that satisfy the three design requirements for linear systems under a fixed communication structure. In particular, we develop a state-space controller under the framework of System Level Synthesis (SLS), an emerging distributed control framework that enables scalable design of optimal controllers for systems with locality, communication, and other constraints that made pre-SLS methods intractable. All previous SLS-based distributed controllers required a "finite-horizon" approximation. Such an approximation not only resulted in suboptimality but also restricted the applicability of the framework since it is well known that such an approximation is not feasible for many systems of interest such as robotics. In contrast, we derive the first infinite-horizon solution in this chapter. Building upon the optimal state-feedback controller, we leverage the separation principle to develop a suboptimal distributed output-feedback controller that is internally stabilizing and memory-efficient.

In Chapter 3, we focus on linear systems with input and state constraints, which are essential for safety-critical systems. We develop a distributed nonlinear controller that not only adheres to the safety constraints but also prevents large oscillations and ensures graceful performance degradation under input saturation caused by unexpected disturbances. The resulting nonlinear controller reduces the overall control cost by 5-35% compared to state-of-the-art linear methods, while conforming to sparse actuation, disturbance localization, and communication delay constraints.

Part II: Interfacing Learning and Control via Uncertainty Sets

In the second part of this thesis, we will address the key challenge of learning and adaptation with safety and performance guarantees. We investigate a unified framework for learning-based control that uses uncertainty sets as the main building block to interface ML techniques and model-based control approaches like the ones developed in Part I. Such an uncertainty-set based approach enables novel algorithms with first-of-its-kind, rigorous guarantees, such as adversarially safe learning-based control as well as scalable and distributed learning-based control.

Set Membership uncertainty sets. We will start by introducing a classic and popular uncertainty set estimation method, called set membership estimation (SME) [48]. One of the most popular uncertainty set estimation methods in control literature, SME identifies the set of system models that are consistent with observed data.

Despite broad applications and empirical success, even in the basic setting of linear systems under simple stochastic distributions, the convergence rate of SME remains open. Quantifying the convergence rate will unlock rigorous performance and safety analysis of popular control designs based on SME, such as adaptive model predictive control and robust online control [49]–[55]. Therefore, in Chapter 4 we provide the first non-asymptotic convergence rate for SME in linear dynamical systems. Moreover, we alleviate the common yet restrictive assumption that a tight upper bound of disturbances must be known a priori and present a novel upper confidence bound based SME method with explicit convergence rate. An immediate implication of this result is that many state-of-the-art algorithms that were based on alternative uncertainty set estimation methods can now use SME to enable substantially improved performance guarantees.

An uncertainty-set centric learning-based control framework. Motivated by new theoretical insights of SME developed in Chapter 4, we investigate an

uncertainty-set based framework where ML techniques and model-based control design interfaces through the uncertainty sets generated by SME, which was first proposed in [13].

The main ideas of the framework is shown in Figure 1.1. At the beginning of the algorithm, we will have an arbitrarily large uncertainty set for the system model and parameters. As an example, this is illustrated as a dark gray square denoted as $\mathcal{P}_0$. As the algorithm interacts with the unknown or uncertain system and environment during deployment, it will observe the resulting state transitioned from the previous time step as a result of the control actions and any unobservable environmental changes such as disturbances and measurement noise (step 1). Every such new observation will provide some information about the uncertainty set. In particular, the observations will provide new constraints (illustrated as dashed lines) for the models that are consistent with the data, carving out the set of implausible models. This is essentially the SME method for constructing uncertainty sets (line 2). In the pictorial example, the new uncertainty set is denoted as $\mathcal{P}_1$. With this latest SME uncertainty set $\mathcal{P}_1$, we will invoke learning algorithms to select a hypothesis model out of the uncertainty set. The selected hypothesis model is denoted as $\widehat{\Theta}_1 \in \mathcal{P}_1$ (Step 3). Based on the selected model, we will perform model-based control synthesis depending on the control objective (step 4). For example, to guarantee safety constraints satisfaction, one may use model predictive control. If instead only stability is required, then robust control methods can be used. The control actions generated based on the currently selected hypothesis model will then be input into the system dynamics, and the process continues, as illustrated in the third picture under step 2, where a new uncertainty set is denoted as $\mathcal{P}_2$ after a new observation.
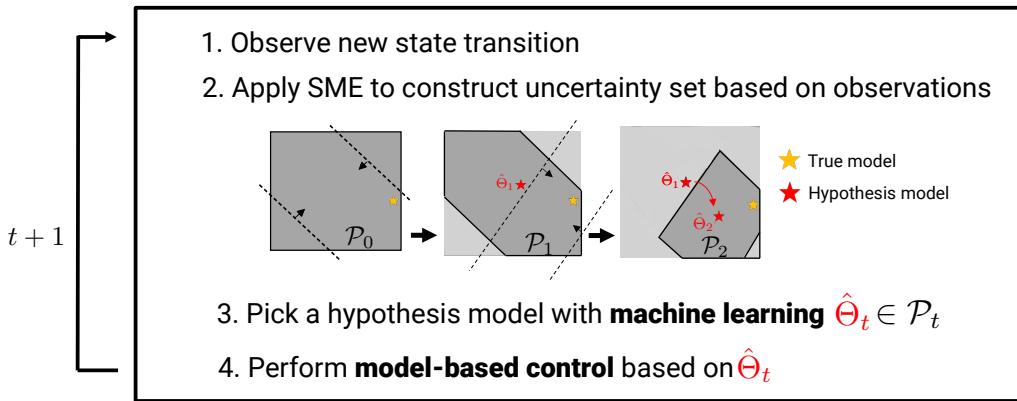
Figure 1.1: Overview of the Uncertainty Set-based Framework

This approach differs from existing adaptation and learning-based control methods

such as Thompson sampling [28], [56], [57], online least squares [58], [59], optimism-based exploration [60]–[62], etc., in that (1) the uncertainty set generated via SME provides a *deterministic uncertainty quantification* of the unknown system that remains valid even with distribution shift of the disturbances, and (2) the selection via online learning induces a novel exploration-exploitation tradeoff that is particularly well-suited for systems that are safety-critical where experiments and data collection can be highly restricted, such as power systems.

Under this general framework, we will illustrate different combinations of ML techniques (step 3) and control methods (step 4) that unlocks novel algorithms with rigorous safety and performance guarantees. In Chapter 5, we will discuss an instantiation of the framework to enable online learning-based control for time-varying linear systems under adversarial disturbances with guaranteed stability. Chapter 6 will demonstrate how this framework can flexibly incorporate control approaches developed in Part I to enable scalable and distributed learning-based control under communication constraints.

Part III: Application to Sustainable Energy Systems
The expansion of sustainable energy systems is essential for decarbonizing the energy infrastructure, but a significant challenge lies in managing the unpredictability and large uncertainties associated with renewable generation as well as the participation of energy storage systems. For example, the renewable energy leads to faster voltage variations in the electricity distribution network, causing voltage deviations from nominal operating limits, which may damage electrical equipment and cause power outages [63], [64]. Yet, an increasing number of distributed energy resources, e.g., photovoltaic and storage devices, are not owned or operated by electric utility companies. This results in a lack of up-to-date information about the grid topology [65], [66], making it challenging for grid operators to design control algorithms which generally require accurate grid topology to respond to the variations from the renewables.

The last part of this thesis focuses on addressing the problem of voltage control in the distribution network where there is large uncertainty in both the network topology as well as load and generation variations. In particular, in Chapter 7, we combine a nested convex body chasing algorithm with a robust predictive controller under the framework presented in Part II to achieve provably finite-time convergence to safe voltage limits. In an online fashion, our algorithm narrows down the set of

possible grid models that are consistent with observations and adjusts reactive power generation accordingly to keep voltages within desired safety limits. Our approach can also incorporate existing partial knowledge of the network to improve voltage control performance. We will demonstrate the effectiveness of our approach in a case study on a Southern California Edison 56-bus distribution system. Our experiments show that in practical settings, the controller is indeed able to narrow the set of consistent topologies quickly enough to make control decisions that ensure stability in both linearized and realistic nonlinear models of the distribution grid.

Part I

Distributed Control under Communication Constraints

Chapter 2

OPTIMAL DISTRIBUTED H2 CONTROL WITH LOCALIZATION AND DELAY

Large-scale CPS such as data centers often face significant scalability issues and communication constraints in algorithm design due to the sheer volume of data exchanged between subsystems and privacy concerns. We will start our investigation of distributed control design under communication constraints with the classical control setting of linear quadratic control, also known as the optimal $\mathcal{H}_2$ control, where the cost is quadratic and the dynamics is linear in states and control inputs. We describe a variant of the optimal $\mathcal{H}_2$ control problem under localization and communication delay constraints and present the optimal state-feedback controller with distributed implementation under this setting. Based on the separation principle of the optimal linear quadratic Gaussian (LQG) control, we will build upon the state-feedback controllers and describe a suboptimal but memory-efficient output-feedback controller under the same communication constraints. This chapter is mainly based on the following papers:

[1]   J. Yu, Y.-S. Wang, and J. Anderson, "Localized and distributed $\mathcal{H}_2$ state feedback control," *2021 American Control Conference (ACC)*, pp. 2732–2738, 2021. DOI: 10.23919/ACC50511.2021.9483301.

[1]   O. Kjellqvist* and J. Yu*, "On infinite-horizon system level synthesis problems," *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 5238–5244, 2022. DOI: 10.1109/CDC51059.2022.9992443.

2.1   Introduction

Large-scale interconnected systems often demand control designs that comply with structural requirements induced by communication constraints, such as sparsity for distributed or localized control [44], [67], and communication delay constraints [44], [68]. These requirements become especially crucial in engineering applications such as power grids [69] and vehicle platoons [70]. Such control design problems are challenging due to the non-convex nature of the problem [39]. Collectively, the challenge of designing controllers subject to these constraints is referred to as

distributed or structured control [71]. It is known that distributed control problems are in general non-convex.

Lately, researchers have focused on novel controller parameterization that admits convex formulation under structural constraints [45]–[47]. These constraints are referred to as the system-level constraints (SLCs) [45]. Examples of SLCs include sparsity constraints on the Youla parameter [72], qudratic invariance (QI) [39] subspace constraints on the feedback controllers, and finite impulse response constraints on the closed-loop responses to disturbances [44], [46]. Among these parameterizations, System Level Synthesis (SLS) emerges as a promising and unified framework for structured controller synthesis that is capable of handling the largest class of structural constraints subsuming all previously convexifiable constraints [73]. A vital feature of the SLS framework is that both the synthesis and the implementation of the structured controller can be done *locally* and *in parallel*, thus scaling favorably with the number of subsystems in a network.

All current SLS-based control methods require both the parameterization and implementation to have finite impulse responses (FIR), with the exception of [74] where continuous-time SLS was formulated, and [75] where suboptimal SLS controllers without structural constraints were presented. This is because optimal controller synthesis, regardless of the choice of convex reparameterization, remains an infinite-dimensional non-convex optimization problem due to system dynamics. The current method of choice to relax SLS problem into a tractable *finite-dimensional* optimization problem is to restrict the optimization variables to having FIR. Such relaxation technique is required for many parameterizations other than SLS [76] and can be used to specify the temporal propagation of the disturbances in the closed loop. Although previous work almost exclusively uses FIR approximations to make SLS tractable, we emphasize that FIR is not a requirement for SLS, but rather a convenient way to use off-the-shelf optimization software. However, FIR approximations result in suboptimal control actions, and more importantly, they lead to deadbeat control, which can cause poorly damped oscillations between discrete sampling intervals and a lack of robustness to model uncertainties due to the high control gains needed to reach the origin in a finite time [75]. Moreover, if the system is only stabilizable but not controllable, FIR approximations can be infeasible.

Contribution.   In this chapter, we lift the FIR constraint of SLS methods to synthesize the *optimal* infinite-horizon solution to a canonical $\mathcal{H}_2$ control problem

under localization and communication delay constraints in the *state-feedback* case. To achieve this, we make a connection between the infinite-horizon state-feedback SLS problems with the Riccati solution for the classic linear quadratic regulator (LQR) problems. Our formulation further relieves several assumptions in existing work, such as the requirement of block diagonal control matrix in the dynamics. The resulting SLS controller has a distributed state-space form that significantly reduces the required memory compared to the FIR SLS controllers and obtains optimal $\mathcal{H}_2$ performance compared to FIR approximations.

In the second half of this chapter, we then generalize the state-feedback result and investigate a class of infinite-horizon *output-feedback* SLS problems. In particular, we study an output-feedback SLS problem that corresponds to a class of LQG problems with localization and communication delay constraints. Our solution leverages an analogous separation principle for SLS parameterization, where the infinite-horizon state-feedback SLS solution is used. A key advantage of this approach is the ability to compute the control gains of each subsystem locally in one swoop using *local* information, without iterations or communications among subsystems, which were required by previous methods, e.g., [77]. Furthermore, we demonstrate an internally stabilizing output-feedback controller that is distributed and localized based on the proposed suboptimal solution. The proposed state-space controller has a fixed, low memory requirement, unlike existing FIR-based SLS controllers where the length of the memory grows linearly with the FIR horizon.

Notation.   Latin letters $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$ present vectors and matrices respectively. $A(i, j)$ refers to the (i, j)th element of the matrix. We use $A(:, j)$ and $A(j, :)$ to refer to the jth column and jth row of A respectively. Bold font $\mathbf{x}$ denotes the signal vector sequence $\mathbf{x} := \{x(t)\}_{t=0}^{\infty}$. Transfer matrices are written as $\mathbf{G}(z) \in \mathbb{C}^{n \times m}$ where $\mathbf{G}(z) = \Sigma_{i=0}^{\infty} z^{-i} G[i]$ with convolution kernels $G[i] \in \mathbb{R}^{n \times m}$. We will omit the dependence on the complex variable z and use $\mathbf{G}$ when there is no ambiguity. The jth standard basis vector is $e_j \in \mathbb{R}^n$. $\text{Sp}(\cdot)$ is the support of a matrix. For two binary matrices $S_1, S_2 \in \{0, 1\}^{m \times n}$, the operation $S_1 \cup S_2$ performs an element-wise OR operation. Given the matrix A, we say $\text{Sp}(A) \subseteq S_1$ if $\text{Sp}(A) \cup S_1 = S_1$. We abbreviate the set $\{1, 2, \ldots, N\}$ as $[N]$ for $N \in \mathbb{N}$. Non-negative integers are denoted as $\mathbb{N}_+$. We write $A > B$ ($A \geq B$) to mean that $A - B$ is a positive (semi)definite matrix. We use $\mathbb{R}\mathcal{H}_\infty$ for the space of all proper and real rational stable transfer matrices and denote $\mathbf{G} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty$ if and only if $z\mathbf{G} \in \mathbb{R}\mathcal{H}_\infty$.
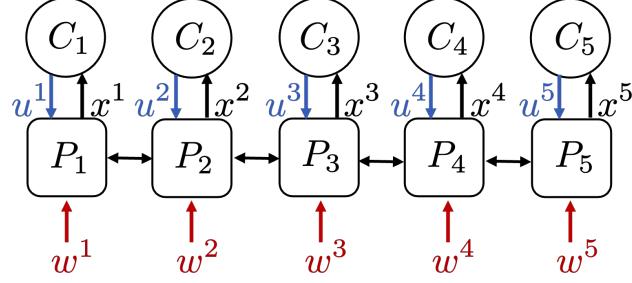
Figure 2.1: Scalar chain network for Example 1, 2, and 3.

2.2 The Localized and Distributed H2 Problem

In this section, we describe the system model, and both the state-feedback and output-feedback $\mathcal{H}_2$ optimal control design problem subject to localization and communication delay constraints.

System Model

We consider interconnected systems consisting of N subsystems. For each subsystem i, let $x^i \in \mathbb{R}^{n_i}$, $u^i \in \mathbb{R}^{m_i}$, $w^i \in \mathbb{R}^{n_i}$ be the local state, control, and disturbance vectors respectively. Each subsystem i has discrete-time dynamics:

$$x^i(t+1) = \sum_{j \in \mathcal{N}_{\text{in}}^1(i)} A^{ij} x^j(t) + B^{ij} u^j(t) + w^i(t),$$

where we denote $j \in \mathcal{N}_{\text{in}}^k(i)$ if the states and control actions of subsystem j affect those of subsystem i in k time steps through the open-loop network dynamics. Analogously, we write $i \in \mathcal{N}_{\text{out}}^k(j)$ when the states of subsystem i are affected by subsystem j via dynamics in k time steps. Stacking the dynamics of all subsystems, we can represent the global network dynamics as

$$x(t+1) = Ax(t) + Bu(t) + w(t), \tag{2.1}$$

with $x(t) \in \mathbb{R}^{N_x}$ and $u(t) \in \mathbb{R}^{N_u}$. We will refer to this system as the *state-feedback* dynamics where the states $x(t)$ can be directly observed and used for feedback control design.

Example 1. *Consider a chain network as shown in Figure 2.1. Each subsystem i has its local plant P_i and controller C_i with scalar state x^i and control action u^i. For each i, $\mathcal{N}_{in}^1(i)$ and $\mathcal{N}_{out}^1(i)$ only contains its nearest neighbors. The stacked network dynamics (2.1) for this system has tri-diagonal state propagation matrix A*

and diagonal B matrix:

$$
A = \begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & * & * & * & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix} \quad B = \begin{bmatrix} * & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & * \end{bmatrix}.
$$

In practice, the states of the system (2.1) may not be directly observable. Therefore, we also consider the *output-feedback* system where in addition to (2.1),

$$
y(t) = Cx(t) + v(t). \tag{2.2}
$$

Here, $y(t) \in \mathbb{R}^{N_y}$ is the observable output of the system and $v(t) \in \mathbb{R}^{N_y}$ is the measurement noise.

In what follows, we define structural constraints arising from locality and communication delays, which belong to the family of SLCs that can be formulated as convex sparsity constraints via the SLS parameteration. We denote any such sparsity constraints as $\mathcal{S} := \{S[k]\}_{k=1}^{\infty}$ where each $S[k]$ are binary matrices specifying the sparsity of the kernels of a transfer matrix. Given a transfer matrix $\mathbf{\Phi}$ and an SLC defined by the sparsity constraint $\mathcal{S}$, we say $\mathbf{\Phi} \in \mathcal{S}$ if $\mathrm{Sp}\,(\Phi[k]) \subseteq S[k]$ for all k.

Localization

It is often desirable to limit the effects of disturbances in (2.1) to a local region for a large network. We can specify the disturbance localization pattern with binary matrices.

Definition 2.2.1 (Localization SLCs). *An SLC $\mathcal{S}^L := \{S^L[k]\}_{k=1}^{\infty}$ is called the localization SLC if $S^L[k]$ for all k are the same binary matrices.*

To specify disturbance localization with localization SLCs, we can impose on control design such that disturbance $\mathbf{w}^j$ propagates to the state $\mathbf{x}^i$ in the closed loop if and only if $\mathcal{S}^L(i, j) \neq 0$. We call the subsystems that can be affected by $\mathbf{w}^i$ the *localized region* of $\mathbf{w}^i$. Subsystems in the localized region of $\mathbf{w}^i$ correspond to the indices of the non-zero elements of the *i*th *column* of $\mathcal{S}^L$. A simple way to ensure disturbance localization per Definition 2.2.1 is that the "boundary" subsystems of each localized region must remain at zero to prevent disturbances from propagating outside of the localized region. To this end, we will formalize the notion of the boundary subsystems.

Definition 2.2.2 (Extended Localization SLCs). *Given a localization SLC $\mathcal{S}^L$, the corresponding extended localization SLC is $\mathcal{S}^{L,e} := \{Sp\left(Sp\left(A\right)S^L[k]\right)\}_{k=1}^{\infty}$.*

Since any (extended) localization SLCs can be uniquely identified by its binary matrix components, we will use $\mathcal{S}^L$ ($\mathcal{S}^{L,e}$) and its binary matrix component interchangeably when there is no ambiguity. Sparsity pattern $\mathcal{S}^{L,e}$ can be interpreted as the propagation of $\mathcal{S}^L$ according to dynamics (2.1) if no action were to be taken to contain the spread of disturbances. We now define the boundary subsystems for a given localization SLC $\mathcal{S}^L$.

Definition 2.2.3 (Boundary Subsystems). *The set of the boundary subsystems for the localized region of $\mathbf{w}^i$ is*

$$\mathcal{B}(i) := \{j \in [N_x] \mid \mathcal{S}^{L,e}(j,i) - \mathcal{S}^L(j,i) \neq 0\}.$$

Intuitively, the set $\mathcal{B}(i)$ for the localized region of $\mathbf{w}^i$ contains the indices of the bordering subsystems that controls the spread of the disturbance from within the localized region to the outside of the region.

Example 2. *We continue with Example 1, where we now specify $\mathcal{S}^L = Sp\left(A\right)$. With Definition 2.2.2 and 2.2.3, we have:*

$$\mathcal{S}^{L,e} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \quad \mathcal{S}^{L,e} - \mathcal{S}^L = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The boundary index set $\mathcal{B}(i)$ thus corresponds to the position of non-zero elements on the ith column of $\mathcal{S}^{L,e} - \mathcal{S}^L$. For instance, $\mathcal{B}(3) = \{1,5\}$ and $\mathcal{B}(1) = \{3\}$.

Communication Delay

In addition to disturbance localization, one may wish to additionally incorporate *communication delay SLCs* to specify how information exchange happens among subsystems while each subsystem tries to localize disturbances. Combined together, such *delayed localization* SLCs correspond to scenarios when communication delays allow disturbances to propagate through dynamics, before subsystems are able to coordinate to completely attenuate and localize them. For ease of exposition, here we consider a communication delay pattern among subsystems that matches the

dynamics. The results in this chapter can be generalized to broader classes of communication patterns.

Definition 2.2.4 (d-Delayed Localization SLCs). *For a fixed integer $d \geq 1$, a localization SLC $\mathcal{S}^L$, a delayed localization SLC $\mathcal{S}^d := \{S^d[k]\}_{k=1}^{\infty}$ is such that $S^d[k] = Sp\left(Sp\,(A)^k\right) for\ k \leq d\ and\ S^d[k] = Sp\left(Sp\,(A)^d\right) \cup S^L[k]\ for\ all\ k \geq d+1.$*

This is sometimes called the (A, d)-sparsity [44] and generalizes the localization SLCs. For the rest of chapter, we consider d-delayed localization SLCs for structured controller synthesis. We assume that any given SLCs are feasible for the control design of the underlying system.

Problem Statement

We now state the localized and distributed optimal $\mathcal{H}_2$ problem both in state-feedback form (P0-SF) and output-feedback form (P0-OF). The objective function is the $\mathcal{H}_2$ performance index of output $\mathbf{z} = Q^{\frac{1}{2}}\mathbf{x} + R^{\frac{1}{2}}\mathbf{u}$ ($Q^{\frac{1}{2}}\mathbf{y} + R^{\frac{1}{2}}\mathbf{u}$) of the state-feedback (output-feedback) closed loop of (2.1) (and (2.2)), with $Q^{\frac{1}{2}}, R^{\frac{1}{2}} > 0$. The disturbances $w(t)$ are assumed to be independently and identically distributed (iid) over different coordinates as well as time, and drawn from zero-mean Gaussian distribution $\mathcal{N}(0, I)$ in the state-feedback case. In the output-feedback case, we assume a more general form where $w(t) \sim \mathcal{N}(0, W)$, and $v(t) \sim \mathcal{N}(0, V)$ iid, with $W, V > 0$. The goal is to synthesize a controller that localizes disturbances and accommodates to the communication delay among subsystems.

State-feedback control. This problem can be represented as the following optimization problem:

$$\underset{\mathbf{K}}{\text{minimize}} \qquad \mathbb{E}_{w(t)\sim\mathcal{N}(0,I)} \left\| \begin{bmatrix} Q^{\frac{1}{2}} & 0 \\ 0 & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \right\|_2^2 \qquad \text{(P0-SF)}$$

subject to   State-feedback dynamics(2.1)

                    $\mathbf{u} = \mathbf{Kx}$, $\mathbf{K}$ internally stabilizing

                    $\mathbf{K}$ localizes disturbances according to $\mathcal{S}^L$ (2.3a)

                    $\mathbf{K}$ conforms to $\mathcal{S}^d$,   (2.3b)

where $\|\mathbf{x}\|_2^2 := \sum_{k=0}^{\infty} \|x(k)\|_2^2$. We assume (A, B) is stabilizable. Problem (P0-SF) has practical application in large-scale cyber-physical systems, especially in power systems [78], [79]. We note that in contrast to all previously formulated SLS

problems, there is no FIR constraint in (P0-SF), rendering it an infinite-dimensional problem.

Output-feedback control. Analogous to the state-feedback problem, the distributed and localized $\mathcal{H}_2$ output-feedback control problem is as follows:

$$
\begin{aligned}
\underset{\mathbf{K}}{\text{minimize}} \qquad & \mathbb{E}_{\substack{w(t)\sim\mathcal{N}(0,W) \\ v(t)\sim\mathcal{N}(0,V)}} \left\| \begin{bmatrix} Q^{\frac{1}{2}}C & 0 \\ 0 & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \right\|_2^2 && \text{(P0-OF)} \\
\text{subject to} \qquad & \text{Output-feedback dynamics}(2.1),\ (2.2) \\
& \mathbf{u} = \mathbf{Ky}, \ \ \mathbf{K} \text{ internally stabilizing} \\
& \mathbf{K} \text{ localizes disturbances according to } \mathcal{S}^L \\
& \mathbf{K} \text{ conforms to } \mathcal{S}^d,
\end{aligned}
$$

where instead of quadratic costs on the states, we penalize quadratic costs on the observable outputs.

2.3   System Level Synthesis

It is well known that structured control design problems such as (P0-SF) and (P0-OF) are infinite-dimensional problems and non-convex in $\mathbf{K}$. Therefore, we leverage the System Level Synthesis (SLS) framework [73]. The SLS theory approaches the constrained state-feedback and output-feedback control problem described above by characterizing all achievable closed-loop mappings (CLMs) from $\mathbf{w}$ (and $\mathbf{v}$) to $\mathbf{x}$, $\mathbf{u}$ under an internally stabilizing controller $\mathbf{K}$. Then, using any achievable CLMs, SLS provides an implementation of the controller $\mathbf{K}$ that realizes the prescribed CLMs.

State-feedback SLS

Consider the closed loop of (2.1) under any linear (potentially dynamic and time-varying) state feedback policy $\mathbf{u} = \mathbf{Kx}$. We denote the closed-loop mappings (CLMs) from disturbance $\mathbf{w}$ to $\mathbf{x}$ and $\mathbf{u}$ by $\boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{uw}$ respectively, i.e., $\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_{xw} \\ \boldsymbol{\Phi}_{uw} \end{bmatrix} \mathbf{w}$. As an example, consider a fixed static controller $K \in \mathbb{R}^{N_u \times N_x}$ such that $u(t) = Kx(t)$. Then the system (2.1) has the following closed loop dynamics,

$$
x(t) = \sum_{k=1}^{t} (A + BK)^{k-1} w(t-k), \quad u(t) = \sum_{k=1}^{t} K(A + BK)^{k-1} w(t-k), \quad (2.5)
$$

where we absorb the initial state $x(0)$ into $w(-1)$ and assume $x(0) = 0$ without loss of generality. Let $\Phi_{xw}[k] := (A + BK)^{k-1}$ and $\Phi_{uw}[k] := K(A + BK)^{k-1}$ be the

convolution kernels. Then (2.5) can be written as $x(t) = \sum_{k=1}^{t} \Phi_{xw}[k]w(t-k)$ and $u(t) = \sum_{k=1}^{t} \Phi_{uw}[k]w(t-k)]$, or equivalently, $\mathbf{x} = \boldsymbol{\Phi}_{xw}\mathbf{w}$ and $\mathbf{u} = \boldsymbol{\Phi}_{uw}\mathbf{w}$.

A main result of SLS is the following characterization of all achievable CLMs under internally stabilizing state-feedback controllers. Crucially, SLS allows re-parameterization of any stabilizing controllers to be expressed and implemented with CLMs.

Theorem 1 ([45]). *For the linear dynamics with $C = I$ and $v(t) = 0$, CLMs $\boldsymbol{\Phi}_{xw}$ and $\boldsymbol{\Phi}_{uw}$ can be achieved by a linear internally stabilizing controller $\mathbf{K}$ if and only if*

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{xw} \\ \boldsymbol{\Phi}_{uw} \end{bmatrix} = I, \quad \boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{uw} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty. \tag{2.6}$$

Moreover, $\mathbf{K} = \boldsymbol{\Phi}_{uw}(\boldsymbol{\Phi}_{xw})^{-1}$ achieves the prescribed CLMs, and can be implemented as

$$u(t) = \sum_{k=1}^{t} \Phi_{uw}[k]\widehat{w}(t-k) \tag{2.7a}$$

$$\widehat{w}(t) = x(t+1) - \sum_{k=1}^{t-1} \Phi_{xw}[k+1]\widehat{w}(t-k), \tag{2.7b}$$

where $\widehat{\mathbf{w}}$ is the internal state of the controller.

Controller (2.7) can be regarded as estimating past disturbances in (2.7b) and acting upon the estimated disturbances according to a specified closed-loop mapping $\boldsymbol{\Phi}_{uw}$ in (2.7a). An important consequence of Theorem 1 is that any structures imposed on the closed-loop responses $\boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{uw}$, such as sparsity constraints on the kernels of $\boldsymbol{\Phi}_{xw}, \boldsymbol{\Phi}_{uw}$, trivially translate into structures on the realizing controllers (2.7).

Output-feedback SLS
Similar to the state-feedback SLS, the CLMs of the closed loop of the output-feedback dynamics (2.1) and (2.2) under an output-feedback linear controller such that $\mathbf{u} = \mathbf{K}\mathbf{y}$ can be fully characterized.

Theorem 2 ([45]). *CLMs $\boldsymbol{\Phi}_{xx}$, $\boldsymbol{\Phi}_{xy}$, $\boldsymbol{\Phi}_{ux}$, $\boldsymbol{\Phi}_{uy} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty$ can be achieved by a linear internally stabilizing output-feedback controller $\mathbf{K}$ if and only if*

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{xx} & \boldsymbol{\Phi}_{xy} \\ \boldsymbol{\Phi}_{ux} & \boldsymbol{\Phi}_{uy} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \tag{2.8a}$$
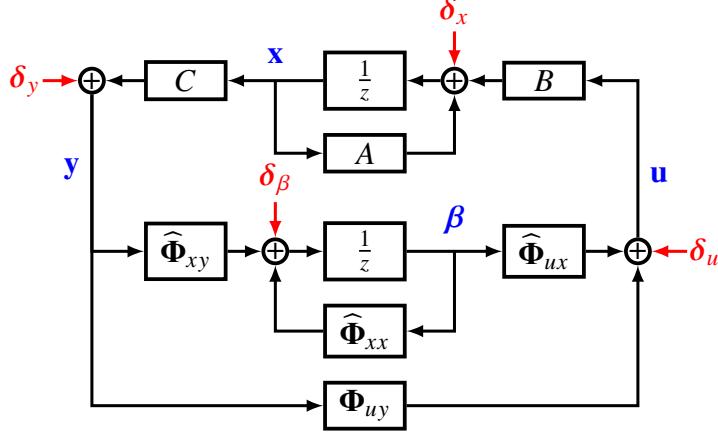
Figure 2.2: Output-feedback controller architecture adapted from [45]. Here, $\widehat{\boldsymbol{\Phi}}_{xx} = z(I - z\boldsymbol{\Phi}_{xx})$, $\widehat{\boldsymbol{\Phi}}_{ux} = z\boldsymbol{\Phi}_{ux}$ and $\widehat{\boldsymbol{\Phi}}_{xy} = -z\boldsymbol{\Phi}_{xy}$. The controller is internally stable; the closed-loop mappings from perturbations $(\boldsymbol{\delta}_x, \boldsymbol{\delta}_y, \boldsymbol{\delta}_u, \boldsymbol{\delta}_\beta)$ to internal signals $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\beta})$ are stable.

$$\begin{bmatrix} \boldsymbol{\Phi}_{xx} & \boldsymbol{\Phi}_{xy} \\ \boldsymbol{\Phi}_{ux} & \boldsymbol{\Phi}_{uy} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}, \tag{2.8b}$$

where $\boldsymbol{\Phi}_{xx}$, $\boldsymbol{\Phi}_{xy}$, $\boldsymbol{\Phi}_{ux}$, $\boldsymbol{\Phi}_{uy}$ maps $\mathbf{w}, \mathbf{v}$ to $\mathbf{x}$, $\mathbf{u}$ under an output-feedback controller $\mathbf{K}$,
i.e., $\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_{xx} & \boldsymbol{\Phi}_{xy} \\ \boldsymbol{\Phi}_{ux} & \boldsymbol{\Phi}_{uy} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix}$. In particular, $\mathbf{K}$ can be implemented as the following,
which is illustrated in Figure 2.2:

$$\begin{aligned} z\boldsymbol{\beta} &= \widehat{\boldsymbol{\Phi}}_{xx}\boldsymbol{\beta} + \widehat{\boldsymbol{\Phi}}_{xy}\mathbf{y} \\ \mathbf{u} &= \widehat{\boldsymbol{\Phi}}_{ux}\boldsymbol{\beta} + \boldsymbol{\Phi}_{uy}\mathbf{y}, \end{aligned} \tag{2.9}$$

where $\widehat{\boldsymbol{\Phi}}_{xx} = z(I - z\boldsymbol{\Phi}_{xx})$, $\widehat{\boldsymbol{\Phi}}_{ux} = z\boldsymbol{\Phi}_{ux}$, $\widehat{\boldsymbol{\Phi}}_{xy} = -z\boldsymbol{\Phi}_{xy}$, and $\boldsymbol{\beta}$ is the controller internal state.

Further, it was shown in [45] that (2.8) is equivalent to stabilizability and detectability of (2.1) and (2.2). Therefore, (2.9) parameterizes all internally stabilizing linear controller $\mathbf{K}$ for (2.1) and (2.2). We also note that Theorem 1 is a special case of Theorem 2.

2.4 Main Results

With the SLS framework introduced in Section 2.3, SLCs in (P0-SF) and (P0-OF) can be equivalently expressed in terms of the CLMs of the closed loop of (2.1) and (2.2).

$$\min_{\Phi_{xw}, \Phi_{uw} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_{xw} \\ \Phi_{uw} \end{bmatrix} \right\|_{\mathcal{H}_2}^2$$

$$\text{s.t.} \quad \begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \Phi_{xw} \\ \Phi_{uw} \end{bmatrix} = I, \quad \text{(SF-SLS)}$$

$$\Phi_{xw} \in \mathcal{S}^L, \quad \Phi_{uw} \in \mathcal{S}^d,$$

and

$$\min_{\Phi \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \left\| \begin{bmatrix} Q^{1/2}C & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{ux} & \Phi_{uy} \end{bmatrix} \begin{bmatrix} W^{1/2} & 0 \\ 0 & V^{1/2} \end{bmatrix} \right\|_{\mathcal{H}_2}^2$$

$$\text{s.t.} \quad \begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{ux} & \Phi_{uy} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}$$

$$\begin{bmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{ux} & \Phi_{uy} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad \text{(OF-SLS)}$$

$$\Phi_{xx}, \Phi_{xy} \in \mathcal{S}^L \quad \text{and} \quad \Phi_{ux}, \Phi_{uy} \in \mathcal{S}^d,$$

where we used Φ to collectively refer to the tuple $\left(\Phi_{xx}, \Phi_{ux}, \Phi_{xy}, \Phi_{uy}\right)$ to reduce notation.

In what follows, we will first derive the infinite-horizon optimal solution to the state-feedback Problem (SF-SLS) with the corresponding distributed controller description. Then we build upon the state-feedback solution and leverage separation principle to synthesize a suboptimal output-feedback solution to (OF-SLS) with distributed state-space controller implementation.

The State-feedback Solution

We will derive the optimal solution of (SF-SLS) in two steps. First, we will derive the solution to (SF-SLS) disregarding delay constraints, i.e., we will restrict our attention to the special case that $\Phi_{xw}, \Phi_{uw} \in \mathcal{S}^L$. We will decribe an explicit controller that implements the CLMs that enables distributed computation with only local information. The solution to this special problem will be used to further solve the full problem of (SF-SLS) with delay constraints.

Column-wise decomposition

It is known that (SF-SLS) is column-wise separable [73], [78], where the optimization problem can be decomposed into subproblems involving columns of the CLMs.

Therefore, we will synthesize the CLMs one column at a time where each subsystem synthesizes the columns corresponds to its local states, in a parallel fashion. Such parallel synthesis scales favorably with the number of subsystems in large networks. From here on, everything will be seen by the jth subsystem. To reduce notation, we assume each subsystem has scalar dynamics where $x^i(t)$, $u^i(t)$, $w^i(t) \in \mathbb{R}$[1]. Let $\boldsymbol{\varphi_x} := \boldsymbol{\Phi}_{xw}(:, j)$ and $\boldsymbol{\varphi_u} := \boldsymbol{\Phi}_{uw}(:, j)$ with kernels $\varphi_x[k]$ and $\varphi_u[k]$ for $k \in \mathbb{N}_+$, respectively corresponding to the ith column of $\boldsymbol{\Phi}_{xw}$ and $\boldsymbol{\Phi}_{uw}$. Furthermore we use $s_x[k]$ and $s_u[k]$ to denote the jth column of $S^L[k]$ and $S^d[k]$ respectively. Each corresponding column problem to be solved locally by subsystem j becomes

$$\min_{\boldsymbol{\varphi_x}, \boldsymbol{\varphi_u}} \quad \sum_{k=1}^{\infty} \varphi_x[k]^\top Q \varphi_x[k] + \varphi_u[k]^\top R \varphi_u[k] \tag{P1}$$

$$\text{s.t.} \quad \varphi_x[k+1] = A\varphi_x[k] + B\varphi_u[k] \tag{2.10a}$$

$$\varphi_x[0] = 0, \quad \varphi_x[1] = e_j$$

$$\varphi_x[k] \in s_x[k], \ \varphi_u[k] \in s_u[k], \ \forall k \in \mathbb{N}_+. \tag{2.10b}$$

This new problem is a constrained linear quadratic optimal control problem, and would be a standard infinite-horizon LQR problem if not for the sparsity constraints (2.10b).

The solution to this column problem has two steps. First, we will solve a version of the problem that only considers localization constraints. Then we will transform the original column problem to a finite-horizon LQR problem with time-varying dynamics. A dynamic programming based solution is proposed based on the solutions with only localization constraints.

Solution with localization SLCs only

In this section, we will consider the column problem (P1) with only localization constraints, where (2.10b) is replaced by $\varphi_x[k] \in S^L[k]$, and $\varphi_u[k] \in S^L[k]$, for all $k \in \mathbb{N}_+$. Therefore, every kernel of $\boldsymbol{\varphi_x}$ and $\boldsymbol{\varphi_u}$ will now have the same localization sparsity constraints for all the kernel matrices. Therefore, we can reduce (P1) by removing zero entries in $\boldsymbol{\varphi_x}$ and $\boldsymbol{\varphi_u}$, other than those associated with the indices in $\mathcal{B}(j)$. We denote the reduced column vectors that contains the entries associated with $\mathcal{B}(j)$ as $\widetilde{\boldsymbol{\varphi}}_x$ and $\widetilde{\boldsymbol{\varphi}}_u$. Similarly, the problem parameters A, B, Q, R can be reduced by selecting submatrices $A^{(j)}$, $B^{(j)}$, $Q^{(j)}$, and $R^{(j)}$ consisting of columns and rows associated with the boundary entries and non-zero entries of $\boldsymbol{\varphi_x}$ and $\boldsymbol{\varphi_u}$. Note these sub-matrices now contain only dynamics information from subsystems

---

[1]One can alleviate this assumption by running the algorithm for multiple columns per subsystem.

that are allowed to transmit information to the subsystems to which $\mathbf{x}^j$ belong. We further rearrange the reduced vectors and matrices in (2.10a) by grouping the entries associated with boundary subsystems as follows:

$$\underbrace{\begin{bmatrix} \widetilde{\varphi}_{x,n} \\ \widetilde{\varphi}_{x,b} \end{bmatrix}[k+1]}_{\widetilde{\varphi}_x[k+1]} = \underbrace{\begin{bmatrix} A_{nn}^{(j)} & A_{nb}^{(j)} \\ A_{bn}^{(j)} & A_{bb}^{(j)} \end{bmatrix}}_{A^{(j)}} \begin{bmatrix} \widetilde{\varphi}_{x,n} \\ \widetilde{\varphi}_{x,b} \end{bmatrix}[k] + \underbrace{\begin{bmatrix} B_n^{(j)} \\ B_b^{(j)} \end{bmatrix}}_{B^{(j)}} \widetilde{\varphi}_u[k], \qquad (2.11)$$

where $\widetilde{\varphi}_{x,b}$ denotes the entries on column vector $\widetilde{\varphi}_x$ that are associated with $\mathcal{B}(j)$ and $\widetilde{\varphi}_{x,n}$ represents the entries of $\widetilde{\varphi}_x$ that are not associated with boundary subsystems. Here, $A^{(j)}$ and $B^{(j)}$ are partitioned accordingly. With abuse of notation, we overload $\widetilde{\varphi}_u$ to denote the rearranged and reduced vector of φ_u.

Example 3. *Consider the scalar chain example in Figure 2.1 for the local problem with $j = 4$, i.e., the subproblem (P1) corresponding to the fourth column of Φ_{xw}, Φ_{uw}. We have the constraint $\varphi_x = [0, 0, \Phi_{xw}(3,4), \Phi_{xw}(4,4), \Phi_{xw}(5,4)]^\top$ according to the fourth column of localization pattern $\mathcal{S}^L = Sp(A)$. According to Example 2, we have $\widetilde{\varphi}_{x,b} = [\Phi_{xw}(2,4)]$ defined in Definition 2.2.3 and $\widetilde{\varphi}_{x,n} = [\Phi_{xw}(3,4), \Phi_{xw}(4,4), \Phi_{xw}(5,4)]^\top$. Therefore, the rearranged and reduced vector is $\widetilde{\varphi}_x = [\Phi_{xw}(3,4), \Phi_{xw}(4,4), \Phi_{xw}(5,4), \Phi_{xw}(2,4)]^\top$.*

Note that constraint (2.10b) now becomes equivalent to the requirement that $\widetilde{\varphi}_{x,b}$ remains at the origin at all time for the localized region of $\mathbf{w}^j$. By keeping the entries associated with boundary subsystems at zero, we implicitly impose that for all k, $Sp(A\varphi_x[k] + B\varphi_u[k]) \subseteq \mathcal{S}^L(:, j)$, which is necessary and sufficient to ensure $\varphi_x \in \mathcal{S}^L(:, j)$. Therefore, the local problem (P1) after rearrangement becomes

$$\min_{\widetilde{\varphi}_x, \widetilde{\varphi}_u \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \quad \sum_{k=0}^{\infty} \widetilde{\varphi}_x[k]^\top Q^{(j)} \widetilde{\varphi}_x[k] + \widetilde{\varphi}_u[k]^\top R^{(j)} \widetilde{\varphi}_u[k] \qquad \text{(P2)}$$

$$\text{subject to} \qquad \widetilde{\varphi}_x[0] = 0, \quad \widetilde{\varphi}_x[1] = e_{j_j}, \quad (2.11) \qquad (2.12a)$$

$$\widetilde{\varphi}_{x,b}[k] = 0, \forall k \in \mathbb{N}_+ \qquad (2.12b)$$

where j_i denotes the new position of element $\Phi_{xw}(j, i)$ in the rearranged and reduced vector $\widetilde{\varphi}_x$. Vectors e_{j_i} have the same dimension as $\widetilde{\varphi}_x$. We differentiate the position of element $\Phi_{xw}(j, i)$ in $\widetilde{\varphi}_{x,n}$ with the notation $\widetilde{j}_i$. Vectors $e_{\widetilde{j}_i}$ has the same dimension as $\widetilde{\varphi}_{x,n}$.

Example 4. *Continuing Example 3 where $i, j = 4$, then $\Phi_{xw}(4,4)$ is in the second position in rearranged and reduced vector $\widetilde{\varphi}_x$. Thus, $j_4 = 2$, $e_{j_4} = [0, 1, 0, 0]^\top$,*

and $\widetilde{j}_4 = 2$ *with* $e_{\widetilde{j}_4} = [0,\, 1,\, 0]^\top$. *Consider instead* $j = 4$ *and* $i = 5$, *then* $\mathbf{\Phi}_{xw}(4, 5)$ *is in the first position in* $\widetilde{\boldsymbol{\varphi}}_x = [\mathbf{\Phi}_{xw}(4,5),\, \mathbf{\Phi}_{xw}(5,5),\, \mathbf{\Phi}_{xw}(3,5)]^\top$ *while it is also in the first position in* $\widetilde{\boldsymbol{\varphi}}_{x,n} = [\mathbf{\Phi}_{xw}(4,5),\, \mathbf{\Phi}_{xw}(5,5)]^\top$. *We then have* $j_5 = 1$ *with* $e_{j_5} = [1,\, 0,\, 0]^\top$ *and* $\widetilde{j}_5 = 1$ *with* $e_{\widetilde{j}_5} = [1,\, 0]^\top$.

We are now in a position to de-constrain (P2) by characterizing CLMs that satisfy (2.12b). We first substitute (2.12b) into (2.11) in (P2) and conclude that (2.12b) is equivalent to requiring

$$-B_b^{(j)}\widetilde{\boldsymbol{\varphi}}_u = A_{bn}^{(j)}\widetilde{\boldsymbol{\varphi}}_{x,n}. \tag{2.13}$$

Due to the equality constraint (2.12a) and (2.11), the free optimization variable is $\widetilde{\boldsymbol{\varphi}}_u$ in (P2). It is clear (2.13) has solutions $\widetilde{\boldsymbol{\varphi}}_u$ if and only if the following assumption holds:

Assumption 1 (existence of solution). $B_b^{(j)} B_b^{(j)\dagger} = I$ *for all* $j \in [N_x]$.

Recall that constraint (2.12b) is sufficient and necessary for the CLMs to comply to the localization pattern $\mathcal{S}^L$. This means assumption 1 is the minimum requirement for the each local problems (P2) to be feasible. Further, per Definition 2.2.3, the number of boundary subsystems can generally be less than the total dimension of control actions, *i.e.*, $B_b^{(j)}$ can be a wide matrix.

Lemma 1. *Under Assumption 1, the parametrization*

$$\widetilde{\varphi}_u[k] = -B_b^{(j)\dagger} A_{bn}^{(j)} \widetilde{\varphi}_{x,n}[k] + \left(I - B_b^{(j)\dagger} B_b^{(j)}\right)\mu[k] \tag{2.14}$$

with $\mu[k]$ *a free vector variable characterizes all* $\widetilde{\varphi}_u[k]$ *that satisfies* (2.12b).

Proof. Under Assumption 1, (2.13) has solutions of the form (2.14). This can be checked by confirming that $\text{Range}\left(I - B_b^{(j)\dagger} B_b^{(j)}\right) = \text{Kernel}\left(B_b^{(j)}\right)$. Substituting (2.14) in (2.11), one can verify that $\widetilde{\varphi}_{x,b}[k] = 0, \forall k \in \mathbb{N}_+$. $\qquad\square$

The reparametrization with variable $\mu[k]$ enables an equivalent local optimization problem without (2.12b). Substitute (2.14) into (P2), we end up with:

$$\min_{\widetilde{\boldsymbol{\varphi}}_{x,n},\mu \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \quad \sum_{k=0}^\infty \widetilde{\varphi}_{x,n}[k]^\top \widetilde{Q}^{(j)} \widetilde{\varphi}_{x,n}[k] + \mu^\top[k]\widetilde{R}^{(j)}\mu[k]$$

$$\text{subject to} \qquad \widetilde{\varphi}_{x,n}[0] = 0, \quad \widetilde{\varphi}_{x,n}[1] = e_{\widetilde{j}_j} \tag{P3}$$

$$\widetilde{\varphi}_{x,n}[k+1] = \widetilde{A}^{(j)}\widetilde{\varphi}_{x,n}[k] + \widetilde{B}^{(j)}\mu[k],$$

where

$$\widetilde{R}^{(j)} = \left(\left(R^{(j)} \right)^{\frac{1}{2}} \left(I - B_b^{(j)\dagger} B_b^{(j)} \right) \right)^{\top}$$
$$\left(\left(R^{(j)} \right)^{\frac{1}{2}} \left(I - B_b^{(j)\dagger} B_b^{(j)} \right) \right)$$
$$\widetilde{Q}^{(j)} = \left((Q^{(j)})^{\frac{1}{2}} - (R^{(j)})^{\frac{1}{2}} B_b^{(j)\dagger} A_{bn}^{(j)} \right)^{\top} \quad (2.16)$$
$$\left((Q^{(j)})^{\frac{1}{2}} - (R^{(j)})^{\frac{1}{2}} B_b^{(j)\dagger} A_{bn}^{(j)} \right)$$
$$\widetilde{A}^{(j)} = A_{nn}^{(j)} - B_n^{(j)} B_b^{(j)\dagger} A_{bn}^{(j)}$$
$$\widetilde{B}^{(j)} = B_n^{(j)} \left(I - B_b^{(j)\dagger} B_b^{(j)} \right).$$

For each column j with $j \in [N_x]$, problem (P3) can be treated as an infinite horizon LQR problem with which an optimal "control policy" $\widetilde{K}_\star^{(j)}$ can be computed in closed form via discrete-time algebraic Riccati equation (DARE):

$$\widetilde{K}_\star^{(j)} = -\left(\widetilde{R}^{(j)} + \widetilde{B}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{B}^{(j)} \right)^{-1} \widetilde{B}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{A}^{(j)}, \quad (2.17)$$

where $\widetilde{X}_\star^{(j)}$ is the Riccati solution to the DARE:

$$\widetilde{X}_\star^{(j)} = \widetilde{Q}^{(j)} + \widetilde{A}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{A}^{(j)} - \widetilde{A}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{B}^{(j)}$$
$$\left(\widetilde{R}^{(j)} + \widetilde{B}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{B}^{(j)} \right)^{-1} \widetilde{B}^{(j)T} \widetilde{X}_\star^{(j)} \widetilde{A}^{(j)}. \quad (2.18)$$

With the optimal solution $\mu[k] = \widetilde{K}_\star^{(j)} \widetilde{\varphi}_{x,n}[k]$ to (P3), the solution to (P2) can be recovered via (2.14) as

$$\widetilde{\varphi}_{x,n}[1] = e_{\widetilde{j}_j} \quad (2.19a)$$
$$\widetilde{\varphi}_u[k] = \left(-B_b^{(j)\dagger} A_{bn}^{(j)} + \left(I - B_b^{(j)\dagger} B_b^{(j)} \right) \widetilde{K}_\star^{(j)} \right) \widetilde{\varphi}_{x,n}[k] \quad (2.19b)$$
$$\widetilde{\varphi}_{x,n}[k] = \left(\widetilde{A}^{(j)} + \widetilde{B}^{(j)} \widetilde{K}_\star^{(j)} \right) \widetilde{\varphi}_{x,n}[k-1]. \quad (2.19c)$$

Note the optimal solution to (P3) via the Riccati equation is stable, so μ and $\widetilde{\varphi}_{x,n}$ construct stable and strictly proper operators.

In summary, we went through a series of transformations and decompositions from the original localized and distributed state feedback $\mathcal{H}_2$ problem (P0-SF) with only localization SLCs to (P3). Given solutions to the local problems (P3), solutions to (SF-SLS) can be recovered. In particular, we define embedding operator $E_x(\cdot)$ and $E_u(\cdot)$ that apply padding of zero's to the reduced vectors $\widetilde{\varphi}_{x,n}[k]$ and $\widetilde{\varphi}_u[k]$ by assigning entries of $\widetilde{\varphi}_{x,n}[k]$ and $\widetilde{\varphi}_u[k]$ to the positions of non-zero elements of

$\Phi_{xw}[k](:,j)$ and $\Phi_{uw}[k](:,j)$ such that $E_x\left(\widetilde{\varphi}_{x,n}[k]\right) \in \mathbb{R}^{N_x}$ and $E_u\left(\widetilde{\varphi}_u[k]\right) \in \mathbb{R}^{N_u}$. We define the application of the embedding operator on the transfer matrices to be the application of the embedding operator on the kernel matrices.

Theorem 3. *Let $\mathbf{\Phi}_{xw}^{\star}$ be the column-wise concatenation of $E_x\left(\widetilde{\varphi}_{x,n}\right)$ and let $\mathbf{\Phi}_{uw}^{\star}$ be the column-wise concatenation of $E_u\left(\widetilde{\varphi}_u\right)$, where $\widetilde{\varphi}_{x,n}$'s and $\widetilde{\varphi}_u$'s are solutions to* (P3) *via* (2.19). *Then $\mathbf{\Phi}_{xw}^{\star}$ and $\mathbf{\Phi}_{uw}^{\star}$ are the unique optimal solution to* (SF-SLS) *with localization SLCs only.*

Proof. It is straightforward to check that optimization (SF-SLS) is an instance of *column-wise separable problem* (see Section III, [78]) where both the objective function and constraints are column-wise separable and can be partitioned and solved in columns as in (P1) in parallel. Therefore, solutions to subproblem (P1) can be concatenated to recover the solution to (SF-SLS). Note that by construction, $E_x\left(\widetilde{\varphi}_{x,n}\right) = \mathbf{\Phi}_{xw}(:,j)$ and $E_u\left(\widetilde{\varphi}_u\right) = \mathbf{\Phi}_{uw}(:,j)$ for all $j \in [N_x]$. Therefore, they comprise the optimal solution to (P1) for each j. Concatenate $E_x\left(\widetilde{\varphi}_{x,n}\right)$'s and $E_u\left(\widetilde{\varphi}_u\right)$'s in a column-wise fashion and the resulting matrices are solutions to (SF-SLS) when only localization SLCs are considered. The uniqueness of the optimal solution is given by the fact that the objective function is strongly convex. $\square$

Distributed controller implementation

Given the column-wise state-space description of the infinite-horizon CLMs $\mathbf{\Phi}_{xw}(:,i)$ and $\mathbf{\Phi}_{xw}(:,i)$ in (2.19), we now describe a state-space agent-level controller that implements these CLMs.

By Theorem 1, we can directly conclude that *theoretically*, $\mathbf{K} = \mathbf{\Phi}_{uw}\left(\mathbf{\Phi}_{xw}\right)^{-1}$ with implementation (2.7) achieves any given CLMs $\mathbf{\Phi}_{xw}, \mathbf{\Phi}_{uw}$ and conforms to the localization and delay constraints. This is because the inheritance of sparsity structures of the controller implementation from CLMs in Theorem 1. Interested readers are referred to [80] for in-depth discussion on implementation of SLS controllers for cyber-physical systems. However, due to the state-space form of the solution with explicit kernel description, *practical* implementation of a controller that achieves the *theoretical* global CLMs remains elusive.

We will decompose the global SLS controller (2.7) into N_x sub-controllers using the solution to (SF-SLS). The global control action $u[t]$ can be accordingly decomposed into N_x "sub-control actions". These sub-control actions are then assembled together to recover the global control action. Importantly, the computation of each sub-control

action conforms to the localization SLC $\mathcal{S}^L$. We now make precise of this high-level description.

To ease notation, we denote $x_\ell(t) \in \mathbb{R}$ and $w_\ell(t) \in \mathbb{R}$, for $\ell \in [N_x]$ as the ℓth position in the state $x(t)$ and disturbance vector $w(t)$ in the global dynamics (2.1), respectively. Furthermore, for general networks with N non-scalar subsystem with local state vector $x^j \in \mathbb{R}^{n_j}$ and $u^j \in \mathbb{R}^{m_j}$, we define the indices associated with x^j of subsystem $j \in [N]$ as $\mathcal{X}(j) := \{\ell \in [N_x] \mid x_\ell \in x^j\}$. Thus, $\mathcal{X}(j)$ partitions the global state vector $x(t)$ in (2.1) into N sets containing the coordinates associated with the N subsystems. Conversely, we use $\mathcal{X}^{-1}(\ell)$ to denote the subsystem index to which state x_ℓ belongs.

For each $\ell \in [N_x]$, we compute the sub-control action vector $u_\ell(t)$, which is a vector with the same dimension as the total number of non-zero elements in the ℓth column of $\Phi_{uw}[t]$, as:

$$\widehat{w}_\ell(t) = x_\ell(t) - \sum_{i \in \mathcal{N}^w(\ell)} \xi_i(t) \left(\widetilde{\ell}_i\right) \tag{2.20a}$$

$$\xi_\ell(t+1) = A_K^\ell \xi_\ell(t) + B_K^\ell \widehat{w}_\ell(t) \tag{2.20b}$$

$$u_\ell(t) = C_K^\ell \xi_\ell(t) + D_K^\ell \widehat{w}_\ell(t), \tag{2.20c}$$

where $\widehat{w}_\ell(t) \in \mathbb{R}$ can be considered as an estimate of the ℓth element of the true disturbance vector $w(t)$. The internal state $\xi_\ell(t)$ of each sub-controller has the same dimension as the total number of non-zero elements in the ℓth column of $\Phi_{xw}[t]$ and $\xi_i(t) \left(\widetilde{\ell}_i\right)$ denotes the $\widetilde{\ell}_i$th element in the internal state vectors ξ_i. Note that controller internal variables have initial condition $\widehat{w}_\ell(0) = x_\ell(0)$ and $\xi_\ell(0) = 0$. We also define the set $\mathcal{N}^w(\ell)$ as $\mathcal{N}^w(\ell) := \left\{i \in [N_x] \mid S^L\left(\mathcal{X}^{-1}(\ell), \mathcal{X}^{-1}(i)\right) \neq 0\right\}$. In particular, the set $\mathcal{N}^w(\ell)$ contains global indices $i \in [N_x]$ such that x_i is a state that is allowed to communicate its information to the subsystem that contains state x_ℓ, conforming to the localization communication pattern $\mathcal{S}^L$. Equation (2.20b) and (2.20c) are the sub-controller internal dynamics specified by $\left(A_K^\ell, B_K^\ell, C_K^\ell, D_K^\ell\right)$ that takes in estimated disturbance $\widehat{w}_\ell$ and output decomposed control actions u_ℓ. The internal dynamics for the lth sub-controller are:

$$A_K^\ell = \widetilde{A}^{(\ell)} + \widetilde{B}^{(\ell)} \widetilde{K}_\star^{(\ell)}, \quad B_K^\ell = \left(\widetilde{A}^{(\ell)} + \widetilde{B}^{(\ell)} \widetilde{K}_\star^{(\ell)}\right) e_{\widetilde{\ell}_\ell}$$

$$C_K^\ell = -B_b^{(\ell)\dagger} A_{bn}^{(\ell)} + \left(I - B_b^{(\ell)\dagger} B_b^{(\ell)}\right) \widetilde{K}_\star^{(\ell)}$$

$$D_K^\ell = \left(-B_b^{(\ell)\dagger} A_{bn}^{(\ell)} + \left(I - B_b^{(\ell)\dagger} B_b^{(\ell)}\right) \widetilde{K}_\star^{(\ell)}\right) e_{\widetilde{\ell}_\ell},$$

where all matrices in the equations above are defined in (2.16) and (2.17). It is straightforward to see that (2.20) is indeed the state space realization of each decomposed SLS controller implementing the reduced ℓth column of $\boldsymbol{\Phi}_{xw}(:, \ell)$ and $\boldsymbol{\Phi}_{uw}(:, \ell)$ synthesized from (P2). In particular, (2.20) implements a transfer function mapping from scalar signal $\mathbf{x}_\ell$ to vector signal $\mathbf{u}_\ell$. Further, each sub-controller is stable since A_K^ℓ is Hurwitz. The block diagram of this transfer function is shown in Figure 2.3, where:

$$\boldsymbol{\Psi}_x^\ell = \left[\begin{array}{c|c} A_K^\ell & B_K^\ell \\ \hline I & 0 \end{array} \right], \quad \boldsymbol{\Psi}_u^\ell = \left[\begin{array}{c|c} A_K^\ell & B_K^\ell \\ \hline C_K^\ell & D_K^\ell \end{array} \right]. \tag{2.21}$$

For each state ℓth state $\mathbf{x}_\ell$ deviating from the origin due to disturbance $\mathbf{w}_\ell$, it triggers subsystems $j \in \mathcal{N}^w(\ell)$ to transmit information among each other in order to generate a *collaborative* sub-control action $\mathbf{u}_\ell$ from these subsystems. Moreover, internal dynamics (2.20b), (2.20c) of each ℓ sub-controller involves only the global dynamics associated with subsystems $j \in \mathcal{N}^w(\ell)$. Therefore, by definition of $\mathcal{N}^w(\ell)$, we conclude that each sub-controller's implementation conforms to the communication pattern specified by $\mathcal{S}^L$. By the superposition property of the input-output behaviors of linear systems, we can sum over all the sub-control actions induced by each $\mathbf{w}_\ell$ and the global control action $u(t) \in \mathbb{R}^{N_u}$ is:

$$u(t) = \sum_{i=1}^{N_x} E_u(u_\ell(t)), \tag{2.22}$$

where each sub-control action $\mathbf{u}_\ell$, which has the same vector dimension as $\widetilde{\Phi}_u^\ell$ can be appropriately padded with zeros using the linear operator $E_u(\cdot)$ to recover a vector dimension in $\mathbb{R}^{N_u}$.
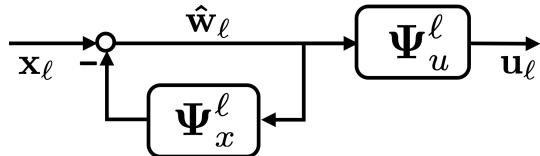
Figure 2.3: Column-wise sub-controller implementation for global controller $\mathbf{K} = \boldsymbol{\Phi}_{uw}(\boldsymbol{\Phi}_{xw})^{-1}$ where the CLMs are computed column-wise in (P2). $\mathbf{x}_\ell$ is the ℓth state, $\widehat{\mathbf{w}}_\ell$ is the estimated ℓth disturbance, and $\mathbf{u}_\ell$ is the sub-control actions induced by ℓth state's deviation from origin.

The following result confirms that collectively, the sub-controllers indeed achieve the prescribed global behaviors.

Theorem 4. *controller implemented* (2.20) *and* (2.22) *defined by solutions to* (P2) *is internally stabilizing for* (2.1) *and achieves the closed-loop mappings* $\mathbf{\Phi}_{xw}$ *and* $\mathbf{\Phi}_{uw}$ *constructed by stacking in a column-wise fashion the solutions to* (P2).

Proof. Recall Theorem 1, where an internally stabilizing controller that realizes $\mathbf{\Phi}_{xw}$ and $\mathbf{\Phi}_{uw}$ has centralized implementation (2.7). Therefore, we establish the equivalence between global control action $u(t)$ generated from (2.7) and $u(t)$ generated from (2.22). Consider (2.7b) where the controller's internal state $\widehat{\mathbf{w}}$ has dynamics

$$\widehat{w}(t) = x(t) - \sum_{k=1}^{t} \Phi_{xw}[k]\widehat{w}(t-k)$$

$$= x(t) - \sum_{i=1}^{N_x} \sum_{k=1}^{t} \Phi_{xw}^{i}[k]\widehat{w}_i(t-k).$$

For each ℓth position in $\widehat{w}(t)$, due to the localization sparsity pattern $\mathcal{S}^L$ imposed on $\mathbf{\Phi}_{xw}$, the scalar dynamics is

$$\widehat{w}_\ell(t) = x_\ell(t) - \sum_{i \in \mathcal{N}^w(\ell)} \sum_{k=1}^{t} \Phi_{xw}(\ell, i)[k]\, \widehat{w}_i(t-k).$$

Since $\mathbf{\Phi}_{xw}^{i}$ for all $i \in [N_x]$ are recovered from (2.19) via the linear operators $E_x(\cdot)$, it is straightforward to verify that

$$\sum_{k=1}^{t} \Phi_{xw}(\ell, i)[k]\, \widehat{w}_i(t-k) = \xi_\ell(t)(\widetilde{\ell}_i), \quad \text{for } t = 1, 2, \ldots.$$

We therefore conclude that (2.7b) and (2.20a),(2.20b) are equivalent. Similarly, re-write (2.7a) as

$$u(t) = \sum_{i=1}^{N_x} \sum_{k=0}^{t} \Phi_{uw}^{i}[k]\widehat{w}_i(t-k).$$

According to (2.19), one can check that $\sum_{k=0}^{t} \Phi_{uw}^{i}[k]\widehat{w}_i(t-k) = E_u(u_\ell(t))$, thus verifying the equivalence between (2.7a) and (2.20b), (2.20c), and (2.22). □

The intuition behind sub-controllers is that at every time step, the global controller actions are decomposed into ℓth sub-control actions that only attenuate the ℓth disturbance, *i.e.*, $\mathbf{w}_\ell$. Therefore, whenever $\mathbf{w}_\ell$ enters the system, only subsystems in the localized region of this disturbance reacts, computing the sub-control actions using only local information available according to $\mathcal{S}^L$.

Solution with delayed localization SLCs

Using the optimal solution to (SF-SLS) with only localization SLCs, we will now present the solution to the original problem (SF-SLS) with *delayed localization* SLCs. In particular, we will follow a similar procedure as in the localization only problem, then invoke the solution to the localization only problem to derive the full solution.

Recall $s_x[k]$ and $s_u[k]$ denote the jth column of $S^L[k]$ and $S^d[k]$ respectively. Let $n_x[k]$ be the number of non-zero elements in the N_x-dimensional vector $s_x[k]$ and $n_u[k]$ be the number of non-zero elements in the N_u-dimensional vector $s_u[k]$. Then there exists a surjective matrix $M_x[k] \in \mathbb{R}^{(N_x - n_x[k]) \times N_x}$ and an injective matrix $M_u[k] \in \mathbb{R}^{N_u \times n_u[k]}$ such that (2.10b) is equivalent to

$$M_x[k]\varphi_x[k] = 0, \quad \varphi_u[k] = M_u[k]q[k], \tag{2.23}$$

where $q[k] \in \mathbb{R}^{n_u[k]}$ becomes the new variable. In particular, one can construct $M_x[k]$ by horizontally stacking standard basis vectors with non-zero positions corresponding to the positions that are zero in $\varphi_x[k]$. On the other hand, $M_u[k]$ can be obtained similarly but with basis vectors corresponding to the non-zero positions in $\varphi_u[k]$. Since $\varphi_x[k+1]$ is uniquely determined by $\varphi_x[k]$ and $\varphi_u[k]$, substitution of (2.23) into (2.10a) yields

$$M_x[k+1]A\varphi_x[k] + \underbrace{M_x[k+1]BM_u[k]}_{F[k]} q[k] = 0. \tag{2.24}$$

The solutions to (2.24) can be expressed as

$$q[k] = F[k]^\dagger M_x[k+1]A\varphi_x[k] + N_F[k]r[k], \tag{2.25}$$

where $N_F[k] \in \mathbb{R}^{n_u[k] \times n_r[k]}$ is a bijection onto the nullspace of $F[k]$. The vector $r[k] \in \mathbb{R}^{n_r[k]}$ is now our new unconstrained optimization variable. Substituting $\varphi_u[k] = M_u[k]q[k]$ and (2.25) into (P1) we get the equivalent time-varying LQR problem

$$\min_{r[k] \in \mathbb{R}^{n_r[k]}} \quad \sum_{k=1}^{\infty} \Big(\varphi_x[k]^\top \widetilde{Q}[k]\varphi_x[k] +$$

$$2r[k]^\top \widetilde{Z}\varphi_x[k] + r[k]^\top \widetilde{R}[k]r[k] \Big)$$

$$\text{s.t.} \quad \varphi_x[k+1] = \widetilde{A}\varphi_x[k] + \widetilde{B}[k]r[k] \tag{2.26}$$

$$\varphi_x[0] = 0, \quad \varphi_x[1] = e_j,$$

where

$$\begin{aligned}
\kappa[k] &= M_u[k]F[k]^\dagger M_x[k+1]A \\
\widetilde{Z}[k] &= N_F[k]^\top M_u[k]^\top R\kappa[k], \quad \widetilde{Q}[k] = Q + \kappa[k]^\top R\kappa[k] \\
\widetilde{R}[k] &= (M_u[k]N_F[k])^\top R[k]M_u[k]N_F[k] \\
\widetilde{A}[k] &= A - B\kappa[k], \quad \widetilde{B}[k] = BM_u[k]N_F[k].
\end{aligned}$$

(2.27)

Finally we note that for $k \geq d+1$, the localization patterns are constant, implying that the dynamics matrices of the transformed problem are static for $k \geq d+1$. Standard dynamic programming arguments allow us to first solve the Riccati equation for the time-invariant problem for $k \geq d+1$ to get the positive definite solution $\widetilde{X}_\star$ and the feedback gain $\widetilde{K}_\star$, and then to solve a finite-horizon time-varying problem by replacing the cost function of each column problem (2.26) with equivalent cost function

$$J = \sum_{k=1}^{d} \left(\varphi_x[k]^\top \widetilde{Q}[k]\varphi_x[k] + 2r[k]^\top \widetilde{Z}\varphi_x[k] + r[k]^\top \widetilde{R}[k]r[k] \right)$$
$$+ \varphi_x[d+1]^\top \widetilde{X}_\star \varphi_x[d+1].$$

(2.28)

The matrices $\widetilde{X}_\star$ and $\widetilde{K}_\star$ are defined to be (2.17) and (2.18) for each column problem, which is the optimal solution to the time-invariant problem with only static localization SLCs. Finally, the solution to the time-varying finite-horizon problem (2.26) with cost (2.28) is given by the Riccati iteration with $\widetilde{X}[d+1] = \widetilde{X}_\star$, and for $k = 1, \ldots, d$,

$$\begin{aligned}
\widetilde{X}[k] &= \widetilde{Q}[k] + \widetilde{A}[k]^\top \widetilde{X}[k+1]\widetilde{A}[k] - \left(\widetilde{A}[k]^\top \widetilde{X}[k+1]\widetilde{B}[k] + \widetilde{Z}[k] \right) \\
&\quad \cdot \left(\widetilde{R}[k] + \widetilde{B}[k]^\top \widetilde{X}[k+1]B[k] \right)^{-1} \left(\widetilde{B}[k]^\top \widetilde{X}[k+1]\widetilde{A}[k] \right) \\
\widetilde{K}[k] &= \left(\widetilde{R}[k] + \widetilde{B}[k]\widetilde{X}[k+1]\widetilde{B}[k] \right)^{-1} \cdot \left(\widetilde{B}[k]^\top \widetilde{X}[k+1]\widetilde{A}[k] + \widetilde{Z}[k]^\top \right).
\end{aligned}$$

(2.29)

Substituting $r[k] = \widetilde{K}[k]\varphi_x[k]$ into (2.25) and further into (2.24), one can obtain the solution to the original problem (P1). We formally state the optimality of the proposed solution.

Theorem 5. *The optimal solution to the infinite-horizon state-feedback SLS problem in* (SF-SLS) *is given, in a column-wise fashion, by*

$$\mathbf{\Phi}_{xw}^{\star}(:,i) = \left[\begin{array}{c|c} A_{SF}^i & B_{SF}^i \\ \hline C_{SF}^i & 0 \end{array}\right], \quad \mathbf{\Phi}_{uw}^{\star}(:,i) = \left[\begin{array}{c|c} A_{SF}^i & B_{SF}^i \\ \hline K_{SF}^i & 0 \end{array}\right], \qquad (2.30)$$

where

$$A_{SF}^i = \begin{bmatrix} 0 & \dots & & & & 0 \\ \widetilde{A}_{CL,i}[1] & 0 & \dots & & & 0 \\ 0 & \widetilde{A}_{CL,i}[2] & 0 & \dots & & 0 \\ 0 & 0 & \ddots & & & \vdots \\ & & & \widetilde{A}_{CL}[d] & \widetilde{A}_{CL,\star}^{(i)} \end{bmatrix}$$

$$B_{SF}^i = \begin{bmatrix} e_i^\top & 0\dots0 \end{bmatrix}^\top \quad C_{SF}^i = \begin{bmatrix} I & I \dots & I \end{bmatrix}$$

$$K_{SF}^i = \begin{bmatrix} \widetilde{K}_i[1] & \widetilde{K}_i[2] & \dots & \widetilde{K}_i[d] & \widetilde{K}_\star^{(i)} \end{bmatrix}, \qquad (2.31)$$

with $\widetilde{K}_i[k]$ *and* $\widetilde{A}_{CL,i}[k] := \widetilde{A}[k] - \widetilde{B}[k]\widetilde{K}[k]$ *computed using* (2.27) *and* (2.29) *for the ith column problem. Matrix* $\widetilde{K}_\star^{(i)}$ *and* $\widetilde{A}_{CL,\star}^{(i)} := \left(\widetilde{A}^{(j)} + \widetilde{B}^{(j)}\widetilde{K}_\star^{(j)}\right)$ *are given by* (2.17) *and* (2.19c) *respectively, where with slight abuse of notation, we overload* $\widetilde{K}_\star^{(i)}$ *and* $\widetilde{A}_{CL,\star}^{(i)}$ *to mean both the original matrix and the matrix with appropriately padded zeros such that* $\widetilde{K}_\star^{(i)} \in \mathbb{R}^{N_u \times N_x}$ *and* $\widetilde{A}_{CL,\star}^{(i)} \in \mathbb{R}^{N_x \times N_x}$ *by reversing the reduction procedure carried out in* (2.11).

Proof. The optimality follows directly from the column separable property of (SF-SLS), and the equivalent transformations between (2.10) and (2.26). The finite-horizon LQR problem with cost (2.28) is equivalent to (2.26) by Bellman's optimality principle. It is straightforward to verify that (2.30) is a state-space realization of the solution to (2.26) by substituting the optimal solution $r[k]$ via (2.29) into (2.25). □

Compared to the solutions with only localization SLCs, the state-space realization of the optimal CLMs here has a higher order because of the first d-delay pattern. Given the column-wise state-space description of the optimal CLMs $\mathbf{\Phi}_{xw}^{\star}(:,i)$ and $\mathbf{\Phi}_{uw}^{\star}(:,i)$, we can adopt the same state-space agent-level controller described in Section 2.4 by simply replacing the state-space implementation of the CLMs with the state space solution (2.30).

Structured Kalman filter design

Theorem 5 can also be used to solve the dual problem of optimal structured Kalman filter design with delayed localization SLCs for (2.1) [81]. In particular, the optimal structured infinite-horizon CLMs that map **w** and **v** to state estimation error **e** under a linear observer **L** with respect to the mean estimation error is given by the solution to the dual problem of (SF-SLS) as shown below:

$$\min_{\mathbf{\Phi}_{ew}, \mathbf{\Phi}_{ev} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \quad \left\| \begin{bmatrix} W^{1/2} & 0 \\ 0 & V^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{ew}^\top \\ \mathbf{\Phi}_{ev}^\top \end{bmatrix} \right\|_{\mathcal{H}_2} \quad \text{(KF-SLS)}$$

$$\text{subject to} \quad \begin{bmatrix} \mathbf{\Phi}_{ew} & \mathbf{\Phi}_{ev} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I \quad (2.32)$$

$$\mathbf{\Phi}_{ew} \in \mathcal{S}^d, \quad \mathbf{\Phi}_{ev} \in \mathcal{S}^d.$$

Readers are referred to [81] for detailed derivation. We highlight the resemblance between constraints (2.32) and that of (SF-SLS) as well as (OF-SLS). In what follows, we will use the optimal solutions from the state-feedback and Kalman-filter SLS problem to construct a suboptimal solution to the output-feedback SLS problem.

The Output-feedback Solution

It is well known that for a linear system, observer-based feedback is always stabilizing if the observer error dynamics are stable and the feedback gain stabilizes the state-feedback case. In [45], the authors pointed out that a similar property holds for CLMs from state-feedback and Kalman-filter SLS problems described above.

Theorem 6 ([73]). *Assume there exist stable and strictly proper transfer matrices* $\mathbf{\Phi}^{SF} = (\mathbf{\Phi}_{xw}^{SF}, \mathbf{\Phi}_{uw}^{SF})$ *and* $\mathbf{\Phi}^{KF} = (\mathbf{\Phi}_{ew}^{KF}, \mathbf{\Phi}_{ev}^{KF})$ *satisfying*

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{xw}^{SF} \\ \mathbf{\Phi}_{uw}^{SF} \end{bmatrix} = I,$$

$$\begin{bmatrix} \mathbf{\Phi}_{ew}^{KF} & \mathbf{\Phi}_{ev}^{KF} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I.$$

The transfer functions

$$\begin{aligned}
\mathbf{\Phi}_{xx} &= \mathbf{\Phi}_{xw}^{SF} + \mathbf{\Phi}_{ew}^{KF} - \mathbf{\Phi}_{xw}^{SF}(zI - A)\mathbf{\Phi}_{ew}^{KF} \\
\mathbf{\Phi}_{ux} &= \mathbf{\Phi}_{uw}^{SF} - \mathbf{\Phi}_{uw}^{SF}(zI - A)\mathbf{\Phi}_{ew}^{KF} \\
\mathbf{\Phi}_{xy} &= \mathbf{\Phi}_{ev}^{KF} - \mathbf{\Phi}_{xw}^{SF}(zI - A)\mathbf{\Phi}_{ev}^{KF} \\
\mathbf{\Phi}_{uy} &= -\mathbf{\Phi}_{uw}^{SF}(zI - A)\mathbf{\Phi}_{ev}^{KF}
\end{aligned} \quad (2.33)$$

are strictly proper and satisfy (2.8).
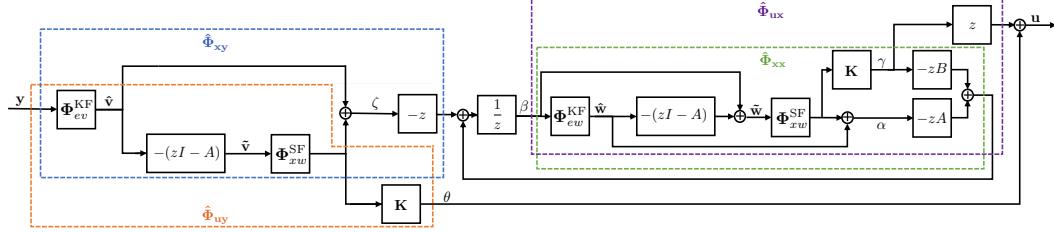
Figure 2.4: Controller implementation of Figure 2.2 after plugging (2.33) in the output-feedback SLS controller (2.9).

Then, a feasible output-feedback SLS controller satisfying the constraints in (OF-SLS) can be constructed using (2.33) with CLMs from (SF-SLS) and (KF-SLS) to stabilize the output-feedback system (2.1) and (2.2) while respecting the prescribed localization and communication delay constraints.

Local controller implementation

This section describes Algorithm 1, which summarizes the local implementation of the global controller (2.9) using CLMs in Figure 2.2 using the localized state-feedback controllers and Kalman filters of Section 2.4.

Globally, the controller after plugging (2.33) in the controller (2.9) is shown in Figure 2.4. Consider the intermediate signals in Figure 2.4,

$$
\begin{aligned}
\zeta &= -\mathbf{\Phi}_{xw}^{\mathrm{SF}} \overbrace{(zI - A)\mathbf{\Phi}_{ev}^{\mathrm{KF}}\mathbf{y}}^{\widetilde{\mathbf{v}}} + \overbrace{\mathbf{\Phi}_{ev}^{\mathrm{KF}}\mathbf{y}}^{\widehat{\mathbf{v}}} \\
\theta &= -\mathbf{\Phi}_{uw}^{\mathrm{SF}}(zI - A)\mathbf{\Phi}_{ev}^{\mathrm{KF}}\mathbf{y} \\
\alpha &= \mathbf{\Phi}_{xw}^{\mathrm{SF}} \overbrace{(\boldsymbol{\beta} - (zI - A)\mathbf{\Phi}_{ew}^{\mathrm{KF}}\boldsymbol{\beta})}^{\widetilde{\mathbf{w}}} + \overbrace{\mathbf{\Phi}_{ew}^{\mathrm{KF}}\boldsymbol{\beta}}^{\widehat{\mathbf{w}}}, \\
\boldsymbol{\gamma} &= \mathbf{\Phi}_{uw}^{\mathrm{SF}}(\boldsymbol{\beta} - (zI - A)\mathbf{\Phi}_{ew}^{\mathrm{KF}}\boldsymbol{\beta}).
\end{aligned}
\tag{2.34}
$$

With these intermediate signals, we can compute the controller internal state $\boldsymbol{\beta}$ and the control signal $\mathbf{u}$ in Figure 2.2 from $z\boldsymbol{\beta} = -z(A\boldsymbol{\alpha} + B_2\boldsymbol{\gamma}) - z\zeta$, and $\mathbf{u} = z\boldsymbol{\gamma} + \boldsymbol{\theta}$.

Locally, due to the communication constraints specified in Section 2.2, one can not carry out the computation described above in a centralized way. In particular, the local computation of each signal in (2.34) involves delayed and locally available information. We now describe the information exchange among subsystems and how they compute (2.34). Recall that the state-feedback solution $\mathbf{\Phi}^{\mathrm{SF}} := (\mathbf{\Phi}_{xw}^{\mathrm{SF}}, \mathbf{\Phi}_{uw}^{\mathrm{SF}})$ and Kalman-filters $\mathbf{\Phi}^{\mathrm{KF}} := (\mathbf{\Phi}_{ew}^{\mathrm{KF}}, \mathbf{\Phi}_{ev}^{\mathrm{KF}})$ synthesized using (2.30) are enforced to respect

the communication and localization constraints expressed as d-delayed localization SLCs. Denote

$$\mathbf{\Phi}_w^{\text{SF}}(:,i) = \left[\begin{array}{c|c} A_{\text{SF}}^i & B_{\text{SF}}^i \\ \hline I & 0 \end{array}\right]. \tag{2.35}$$

Then $\mathbf{\Phi}_{xw}^{\text{SF}}(:,i) = C_{\text{SF}}^i \mathbf{\Phi}_w^{\text{SF}}(:,i)$ and $\mathbf{\Phi}_{uw}^{\text{SF}}(:,i) = K_{\text{SF}}^i \mathbf{\Phi}_w^{\text{SF}}(:,i)$ where C_{SF}^i and K_{SF}^i are from (2.31). Computing the local components of α and β requires only one realization of $\mathbf{\Phi}_w^{\text{SF}}$ as they can share the same copy of the states within each subsystem. An analogous statement holds true for ζ and θ. Denote the two realizations of (2.35) as $\mathbf{\Phi}_{w,\alpha}^{\text{SF}}$ and $\mathbf{\Phi}_{w,\zeta}^{\text{SF}}$. During each time step t, every node observes its local output $y_i(t)$ and goes through four stages of computation and communication with its neighbors leading to an update to the internal controller states and the application of the actuator signal $u_i(t)$. This is summarized in Algorithm 1 with subroutines 2.1–2.4 describing these computations in detail.

Algorithm 1: Local computation of controller signals

for *Each node $i = 1, \ldots, N$* **do**
  Input: $\mathbf{\Phi}_{w,\alpha}^{\text{SF}}(:,i), \mathbf{\Phi}_{w,\zeta}^{\text{SF}}(:,i), \mathbf{\Phi}_{ew}^{\text{KF}}(:,i), \mathbf{\Phi}_{ev}^{\text{KF}}(:,i)$
  Initialize : $\beta_i(0) \leftarrow 0, w_i(t) \leftarrow 0, v_i(t) \leftarrow 0$
end
for $t = 0, 1, \ldots$ **do**
  for *each node $i = \{1, \ldots, N\}$* **do**   // parallel
1 | Observe $y_i(t)$ Receive $\beta_j(t)$ and $y_j(t)$ from $j \in \mathcal{N}_{\text{in}}^d(i)$
2 | subroutine1()
3 | Receive $\widehat{w}_j(t)$ and $\widehat{v}_j(t)$ from $j \in \mathcal{N}_{\text{in}}^1(i)$
4 | subroutine2()
5 | Receive $\widehat{\alpha}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1), \widehat{\zeta}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1), \widehat{\gamma}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1)$ and
    $\widehat{\theta}^{(\mathcal{N}_{\text{out}}^d(j))}(t)$ from $j \in \mathcal{N}_{\text{in}}^d(i)$
6 | subroutine3()
7 | Receive $\alpha_j(t+1)$ and $\gamma_j(t+1)$ from $j \in \mathcal{N}_{\text{in}}^1(i)$
8 | subroutine4()
9 | Apply $u_i(t)$
10 | **end**
end

Control signal computation at subsystem i begins by receiving the measurements from neighbors j at most d steps away (line 1) and computing the ith element of the internal signals $\widehat{v}(t+1)$ and $\widehat{w}(t+1)$ via Subroutine 2.1, where the function
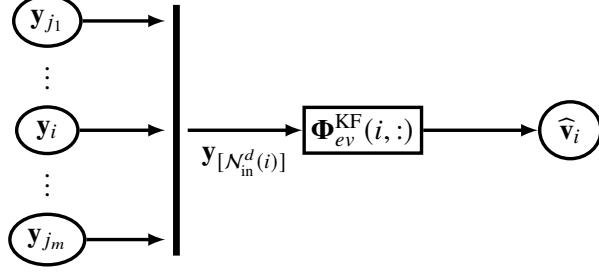
Figure 2.5: Illustration of subroutine 1. The computation of $\widehat{w}$ is similar.
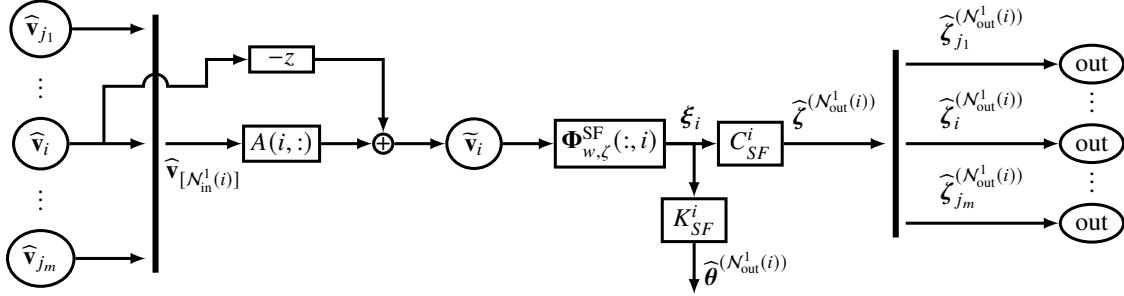
Figure 2.6: Illustration of subroutine 2.

$\text{step}(G, u)$ means that the internal dynamics of the system G is propagated one time-step with the input u.

Subroutine 2.1: Compute $\widehat{w}_i(t+1)$ and $\widehat{v}_i(t+1)$

Receive $\beta_j(t)$ and $y_j(t)$ from $j \in \mathcal{N}_{\text{in}}^d(i)$

$\beta_{[\mathcal{N}_{\text{in}}^d(i)]}(t) \leftarrow \text{vec}(\beta_{j_1}(t), \ldots, \beta_{j_m}(t))$

$y_{[\mathcal{N}_{\text{in}}^d(i)]}(t) \leftarrow \text{vec}(y_{j_1}(t), \ldots, y_{j_m}(t))$

$\widehat{w}_i(t+1) \leftarrow \text{step}(z\mathbf{\Phi}_{ew}^{\text{KF}}(i,:), \beta_{[\mathcal{N}_{\text{in}}^d(i)]}(t))$

$\widehat{v}_i(t+1) \leftarrow \text{step}(z\mathbf{\Phi}_{ev}^{\text{KF}}(i,:), y_{[\mathcal{N}_{\text{in}}^d(i)]}(t))$

In the second stage, the node receives $\widehat{v}_j(t)$ and $\widehat{w}_j(t)$ from its closest neighbors (line 3) and computes the outgoing components of (2.34). The computations are outlined in Subroutine 2.2 and illustrated in Figure 2.6.

In the third stage, which is demonstrated in Figure 2.7, the node receives the components pertaining to its element of the signals in (2.33) from other nodes a distance at most d steps away with delayed information (line 5) and sums them to compute the ith element of each signal in (2.34). This step is described in Subroutine 2.3.

In the final stage (Subroutine 2.4) the node receives $\alpha_j(t+1)$ and $\gamma_j(t+1)$ from its

---

Subroutine 2.2: Compute the outgoing components of (2.34)

$\widehat{w}_{[\mathcal{N}_{\text{in}}^1(i)]}(t) \leftarrow \text{vec}(\widehat{w}_{j_1}(t), \ldots, \widehat{w}_{j_m}(t))$

$\widehat{v}_{[\mathcal{N}_{\text{in}}^1(i)]}(t) \leftarrow \text{vec}(\widehat{v}_{j_1}(t), \ldots, \widehat{v}_{j_m}(t))$

$\widetilde{w}_i(t) = \beta_i(t) + A\widehat{w}_{[\mathcal{N}_{\text{in}}^1(i)]}(t) - \widehat{w}_i(t+1)$

$\widetilde{v}_i(t) = A\widehat{v}_{[\mathcal{N}_{\text{in}}^1(i)]}(t) - \widehat{v}_i(t+1)$

$\lambda_i(t+1) \leftarrow \text{step}(\boldsymbol{\Phi}_{w,\alpha}^{\text{SF}}(:,i), e_i \widetilde{w}_i(t))$

$\xi_i(t+1) \leftarrow \text{step}(\boldsymbol{\Phi}_{w,\zeta}^{\text{SF}}(:,i), e_i \widetilde{v}_i(t))$

$\widehat{\alpha}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1) \leftarrow C_{SF}^i \lambda_i(t+1)$

$\widehat{\zeta}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1) \leftarrow C_{SF}^i \xi_i(t+1)$

$\widehat{\gamma}^{(\mathcal{N}_{\text{out}}^d(j))}(t+1) \leftarrow K_{SF}^i \lambda_i(t+1)$

$\widehat{\theta}^{(\mathcal{N}_{\text{out}}^d(j))}(t) \leftarrow K_{SF}^i \xi_i(t+1)$
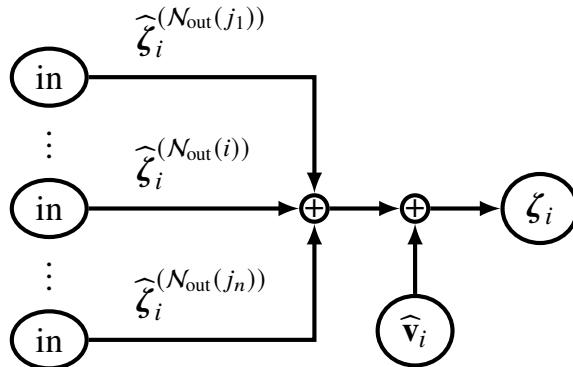
Figure 2.7: Illustration of subroutine 3.

Subroutine 2.3: Compute the local elements of (2.34)

$\alpha_i(t+1) \leftarrow \widehat{w}_i(t+1) + \sum_j \widehat{\alpha}_i^{(\mathcal{N}_{\text{out}}^1(j))}(t+1)$

$\gamma_i(t+1) \leftarrow \sum_j \widehat{\gamma}_i^{(\mathcal{N}_{\text{out}}^1(j))}(t+1)$

$\zeta_i(t+1) \leftarrow \widehat{v}_i(t+1) + \sum_j \widehat{\zeta}_i^{(\mathcal{N}_{\text{out}}^1(j))}(t+1)$

$\theta_i(t) \leftarrow \sum_j \widehat{\theta}_i^{(\mathcal{N}_{\text{out}}^1(j))}(t)$

---

closest neighbors (line 7) and computes $\beta_i(t+1)$ and $u_i(t)$. We conclude that the node has now received information from nodes at most $2d + 2$ steps away.

Subroutine 2.4: Compute $u_i(t)$ and $\beta_i(t+1)$

$\alpha_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) \leftarrow \text{vec}\left(\alpha_{j_1}(t+1), \ldots, \widehat{\alpha}_{j_m}(t+1)\right)$

$\gamma_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) \leftarrow \text{vec}\left(\gamma_{j_1}(t+1), \ldots, \widehat{\gamma}_{j_m}(t+1)\right)$

$\beta_i(t+1) \leftarrow -A(i,:)\alpha_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) - B(i,:)\gamma_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) - \zeta_i(t+1)$

$u_i(t) \leftarrow \gamma_i(t+1) + \theta_i(t)$

We summarize the stability properties of Algorithm 1 in the following theorem:

Theorem 7. *Algorithm 1 with $\mathbf{\Phi}^{SF}_w$ as in* (2.35)*, $\mathbf{\Phi}^{KF}_{ew}$ and $\mathbf{\Phi}^{KF}_{ev}$ solved from* (KF-SLS) *internally stabilizes system* (2.1) *and* (2.2)*. Moreover if $\mathbf{\Phi}^{SF}_w$, $\mathbf{\Phi}^{KF}_{ew}$ and $\mathbf{\Phi}^{KF}_{ev}$ are d-localized, the closed loop is at most $2d + 2$-localized.*

Proof. By Theorem 6 the closed-loop mappings satisfy (2.8a) and (2.8b). Concatenating $\mathbf{y}_i$, $\boldsymbol{\beta}_i$ and $\mathbf{u}_i$ we get precisely the signals in Figure 2.2 which is internally stable [73], we need to show that the closed-loop is internally stable for perturbations entering in the intermediate steps outlined in Subroutines 2.1–2.4. Note that a perturbation entering at any of the intermediate signals can be modeled as a disturbance entering as $\boldsymbol{\delta}_x$, $\boldsymbol{\delta}_y$ or $\boldsymbol{\delta}_\beta$ pre-filtered through a stable linear system. Similarly, probing any of the internal signals can be represented as probing $\mathbf{y}$, $\mathbf{u}$ or $\boldsymbol{\beta}$ post-filtered through a stable system. We conclude Algorithm 1 is internally stable in feedback with the system (2.1), (2.2). Finally, as d-localization is closed under addition, and composition of a d- and a k-localized operator is at most $d + k$-localized, (2.33) implies that the closed loop is at most $2d + 2$-localized. $\qquad \square$

2.5 Simulation

Consider a bi-directional scalar chain network parameterized by α and ρ:

$$x^i(t+1) = \rho(1 - 2\alpha)x^i(t) + \rho\alpha \sum_{j \in \{i \pm 1\}} x^j(t) + u^i(t) + w^i(t),$$

where α is a coupling constant and ρ is the spectral radius of the global state-transition matrix A, with $\rho \geq 1$ being unstable. We first verify the optimality of the infinite-horizon state-feedback solution given in Section 2.4. In this simulation, we choose the number of scalar subsystems to be 15, $\alpha = 0.6$ and $\rho = 1$. For the quadratic cost matrices, we let $Q = I$ and $R = 300 \cdot I$. For SLCs, we let the delayed localization
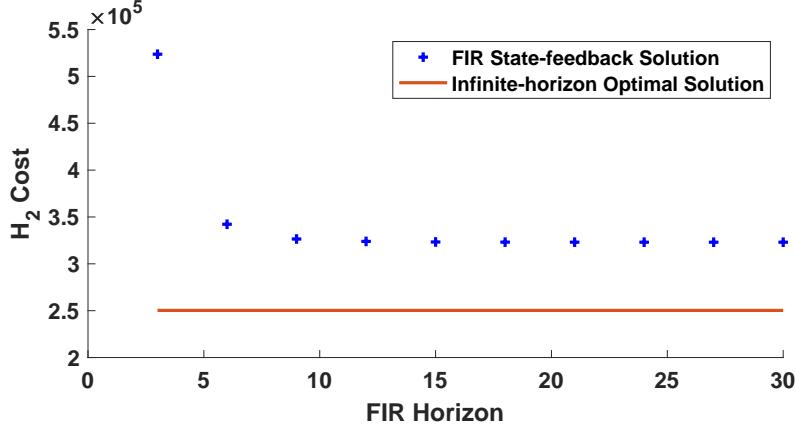
Figure 2.8: The infinite-horizon SLS solution achieves the optimal cost.

parameter be $d = 3$. The result is shown in Figure 2.8, where the optimality of our approach is clear. Due to the high penalty on the control actions, the performance degradation under FIR approximation can be significant.

Next, we investigate the optimality gap between the suboptimal infinite-horizon output-feedback solution proposed in this work against the FIR output-feedback solution computed numerically with a fixed FIR horizon of 20. We let all other parameters remain the same as before, and change the number of subsystems to 10 in this simulation. First, we study how the d delayed localization parameter influence the optimality gap. This is illustrated in Figure 2.9. As expected, the more localized the output-feedback problem is, the bigger the optimality gap is between the constructed solution using separation principle and the direct FIR output-feedback solution. As the delayed localization pattern becomes more global, the proposed output-feedback solution becomes more optimal. When the delayed localization SLCs become non-binding (for $d \geq 6$), we see that the proposed infinite-horizon output-feedback solution actually becomes optimal and achieves lower cost than the FIR solution. This is due to the separation principle of centralized LQG.

Next, we investigate how the optimality gap grows with the number of subsystems in the network. Here we set $C = I$ and fix the delayed localization parameter to be $d = 3$. As can be seen in Figure 2.10, we observe that the optimality gap grows apparently linearly in the number of subsystems. However, we highlight the numerical efficiency and stability of our approach despite the suboptimality. When the number of subsystems exceeds 12 with FIR horizon of 20, the FIR solution solved in MATLAB using CVX renders NaN due to numerical instability (total of 11520 variables).
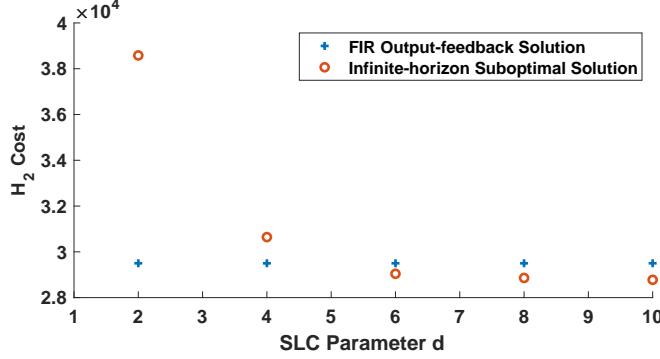
Figure 2.9: The proposed infinite-horizon suboptimal solution to the output-feedback SLS problem versus the FIR output-feedback solution numerically computed for (OF-SLS) for varying SLC delayed localization parameter d.
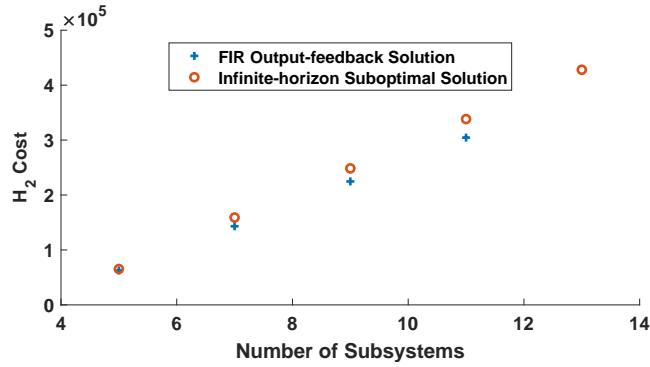
Figure 2.10: The proposed suboptimal solution to the output-feedback SLS problem versus the FIR output-feedback solution numerically computed for (OF-SLS) for varying number of subsystems in the network.

2.6   Conclusion

In this chapter, we propose and derive the optimal solution to a structured $\mathcal{H}_2$ state-feedback control problem under localization and communication delay constraints. In particular, our controller is synthesized directly in infinite-dimensional space, without the finite-impulse response relaxation common in related work. Our method can also be used to construct optimal distributed Kalman filters with limited information exchange. We combine the distributed Kalman filter with state-feedback control to perform localized LQG control with communication constraints. We provide agent-level implementation details for the resulting output-feedback state-space controller. The distributed Riccati solutions presented in this work can be used as quadratic terminal cost in distributed model predictive control to enable significantly better performance, as seen in [82].

Chapter 3

NONLINEAR DISTRIBUTED CONTROL FOR LARGE-SCALE SYSTEMS WITH CONSTRAINTS

In Chapter 2, we presented the optimal distributed linear controller for linear systems with quadratic cost under communication constraints. However, it is known that even the optimal linear controllers can be outperformed by nonlinear controllers for control design under communication constraints [38]. Moreover, the optimal linear distributed controllers do not handle safety constraints, such as when states are constrained to be inside a safe set. Such safety constraints are critical for CPS such as the power grid.

Therefore, in this chapter, we propose a novel nonlinear distributed control technique for linear discrete-time systems with input saturation and state constraints that allows us to simultaneously satisfy performance and safety objectives defined for small- and large-disturbance regimes respectively. Previous methods for this class of systems trade off between the two objectives whereas our method allows the controller to perform well in both regimes by incorporating online information about disturbances during operation.

[1]   J. Yu* and D. Ho*, "Achieving performance and safety in large scale systems with saturation using a nonlinear system level synthesis approach," *2020 American Control Conference (ACC)*, pp. 968–973, 2020. DOI: 10.23919/ ACC45564.2020.9147577.

3.1   Introduction

Linear systems with input saturation, where control inputs are limited to a certain range, and state constraints, which require the system states to remain within predefined bounds, are an important class of nonlinear models for many practical applications, such as robotics, power systems, and automotive systems, where physical limitations on actuators and safety requirements must be strictly adhered to. Moreover, even without the state constraints, the input saturation alone can result in large and sometimes diverging oscillations in the closed loop.

There are generally two approaches to handle the challenge. The first approach is to include these constraints in the overall controller design. Methods following

such principle include robust Model Predictive Control (MPC) [83]–[85] and constrained System Level Synthesis (SLS) [86]. However, such methods are often characteristically conservative. In particular, MPC-based methods may suffer from recursive feasibility issues in addition to challenges in dealing with disturbances. The second approach attempts to separately design a desired nominal controller for small signals and a "compensator" that compensates for saturation-induced performance degradation or even instability. Generally, the design goal is to have the saturated closed loop stay as close as possible to the small-signal nominal behavior. Anti-windup Schemes [87], Reference Governor [88], and other Lyapunov invariant set-based compensators[89] all belong to this approach. Despite vast literature on the topics, decentralized schemes that accommodate large-scale systems are few.

Contribution. In this chapter, we propose a novel offline distributed nonlinear controller synthesis procedure that outperforms any optimal linear distributed controller for constrained LQR problem [83]–[86]. In addition, the same controller inherently prevents windup-instabilities in saturated linear systems which are traditionally mitigated via additional anti-windup design [88], [90], [91]. Another significant advantage of the approach, is that despite being a nonlinear synthesis method it naturally enjoys the same benefits as the linear system level approach introduced in Chapter 2, which makes it scalable in the large-system setting.

[92] describes the system-level characterization of closed loops of general nonlinear discrete-time systems and introduces a simple universal control structure, called a system level controller, that has the capacity to stabilize any nonlinear system if parametrized with the according closed loop maps. In this chapter, we will show that just using a very special case of the framework presented in [92] provides new promising tools for control design. In particular, we will illustrate how a simple projection nonlinearity can become a powerful tool for solving the problems described above.

The remainder of the chapter starts with preliminaries and notations in Section 3.2, followed by a review on nonlinear System Level Synthesis (NLSLS) in Section 3.3. The proposed nonlinear controller is introduced in Section 3.4, followed by Section 3.5 where the constrained distributed LQR problem is discussed. We show in Section 3.6 that the proposed nonlinear controller can be augmented for natural anti-windup properties and therefore allow for large-scale distributed anti-windup design. Numerical simulation in Section 3.7 corroborates the presented theory.

3.2 Preliminaries

We will define ℓ^n to be the space of sequences of vectors in $\mathbb{R}^n$. Sequences of vectors will be denoted by small bold letters $\mathbf{x} := (x_t)_{t=0}^\infty$ unless otherwise specified. Occasionally, we will define sequences explicitly with the tuple notation $\mathbf{x} = (x_0, x_1, \dots)$ and x_t^j denotes the jth element of vector x_t. We use the $x_{i:j}$ to refer to the truncation of a sequence $\mathbf{x}$ to the tuple $(x_i, x_{i+1}, \dots, x_j)$. Furthermore, we will adopt $|x|$ and $|A|$ for $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ as the vector ∞-norm and induced ∞-norm on $\mathbb{R}^n$, respectively. We reserve $\|\cdot\|_p$ to refer to the norm and induced norm over vector sequence space ℓ_p:

$$\|\mathbf{x}\|_p := \left(\sum_{k=0}^\infty |x_k|^p \right)^{\frac{1}{p}} \qquad \|\mathbf{x}\|_\infty := \sup_{k \geq 0} |x_k|.$$

Operators

Operators that maps between sequence spaces will be denoted in bold capital letters $\mathbf{T} : \ell^n \to \ell^k$. Similar to the sequence of vectors, we write $\mathbf{T} := \{T_t\}_{t=0}^\infty$ with its component functions $T_t : \ell^n \to \mathbb{R}^k$. An operator $\mathbf{T}$ will be called *causal* if for any pair of input $\mathbf{x}$ and its corresponding output $\mathbf{y} = \mathbf{T}(\mathbf{x})$, the output y_t does not depend on future input sequence $x_{t+k}, k \geq 1$. More precisely, a causal operator $\mathbf{T}$ is fully characterized by its component functions $T_t : \mathbb{R}^{n \times (t+1)} \to \mathbb{R}^k$ such that

$$\mathbf{T}(\mathbf{x}) = (T_0(x_0), T_1(x_1, x_0), T_2(x_2, x_1, x_0), \dots).$$

Note that every component function T_t of a causal operator $\mathbf{T}$ has $t + 1$ arguments which are populated in reverse-chronological order. If in addition , component functions T_t satisfy $T_t(x_{t:0}) = T_t(0, x_{t-1:0})$, then $\mathbf{T}$ will be called *strictly causal*.

We define the space of all causal and strictly causal operators that maps $\ell^n \to \ell^p$ as $C(\ell^n, \ell^p)$ and $C_s(\ell^n, \ell^p)$, respectively. Moreover, let the space of all linear causal and strictly causal operators be denoted as $\mathcal{LC}(\ell^n, \ell^p) \subset C(\ell^n, \ell^p)$ and $\mathcal{LC}_s(\ell^n, \ell^p) \subset C_s(\ell^n, \ell^p)$. Occasionally, for two operators with matching domains such as $\mathbf{A} \in C(\ell^n, \ell^p)$ and $\mathbf{B} \in C(\ell^n, \ell^q)$, we denote the composite operator $(\mathbf{A}, \mathbf{B}) \in C(\ell^n, \ell^p \times \ell^q)$ as $(\mathbf{A}, \mathbf{B}) : \mathbf{x} \mapsto (\mathbf{A}(\mathbf{x}), \mathbf{B}(\mathbf{x}))$.

ℓ_p Stability

Let the vector sequence space $\ell_p^n \subset \ell^n$ be defined as

$$\ell_p^n := \{\mathbf{x} \in \ell^n \mid \|\mathbf{x}\|_p < \infty\}.$$

We define stability for causal operators as follows:

Definition 3.2.1 (ℓ_p Stability)**.** *An operator* $\mathbf{T} \in C(\ell^n, \ell^m)$ *is said to be* ℓ_p*-stable, if*

$$\mathbf{T}(\mathbf{x}) \in \ell_p^m \quad \text{for all } \mathbf{x} \in \ell_p^n.$$

Further, if there exist two scalars $\gamma, \beta \geq 0$ *such that for all* $\mathbf{x} \in \ell_p^n$*, we have*

$$\|\mathbf{T}(\mathbf{x})\|_p \leq \gamma \|\mathbf{x}\|_p + \beta,$$

then $\mathbf{T}$ *is finite gain* ℓ_p*-stable.*

3.3 An Overview of the Nonlinear System Level Approach

This section will focus on introducing the notion of closed loop maps as causal operators with respect to a general nonlinear causal system. Moreover, we summarize necessary and sufficient conditions for operators to be closed loop maps and how they can be realized by a dynamic controller.

Closed Loop Maps of Nonlinear Systems

Consider a discrete-time nonlinear system with additive disturbances

$$x_t = f(x_{t-1}, u_{t-1}) + w_t, \tag{3.1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $w_t \in \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ with $f(0,0) = 0$ and $x_0 = w_0$. Let $\mathbf{F}(\mathbf{x}, \mathbf{u}) : \ell^n \times \ell^m \to \ell^n$ be the strictly causal operator representation of the function f such that $\mathbf{F}(\mathbf{x}, \mathbf{u}) := (0, f(x_0, u_0), f(x_1, u_1), \dots)$. Assume that w_t can not be measured and that u_t is generated by some causal controller $\mathbf{K} \in C(\ell^n, \ell^m)$ such that $u_t = K_t(x_{t:0})$. An equivalent operator form of the dynamics (3.1) is

$$\mathbf{x} = \mathbf{F}(\mathbf{x}, \mathbf{u}) + \mathbf{w} \tag{3.2a}$$

$$\mathbf{u} = \mathbf{K}(\mathbf{x}). \tag{3.2b}$$

For a fixed disturbance sequence $\mathbf{w}$, the dynamics (3.1) produces unique closed loop trajectories for state $\mathbf{x}$ and input $\mathbf{u}$. Therefore, given a fixed $\mathbf{K}$, the dynamics induce a causal map from $\mathbf{w}$ to $(\mathbf{x}, \mathbf{u})$ and we will call the corresponding operators disturbance-to-state and disturbance-to-input *closed loop map*, respectively.

Definition 3.3.1 (Closed Loop Maps)**.** *Define* $\mathbf{\Phi}[\mathbf{F}, \mathbf{K}] \in C(\ell^n, \ell^n \times \ell^m)$ *as the operator that maps* $\mathbf{w}$ *to the corresponding response* $(\mathbf{x}, \mathbf{u})$ *according to the closed-loop dynamics* (3.2). *We call* $\mathbf{\Phi}[\mathbf{F}, \mathbf{K}]$ *the closed loop maps (CLMs) of* (3.2). *Moreover we will refer to the partial maps* $\mathbf{w} \to \mathbf{x}$ *and* $\mathbf{w} \to \mathbf{u}$ *with* $\mathbf{\Phi}^{\mathbf{x}}[\mathbf{F}, \mathbf{K}]$ *and* $\mathbf{\Phi}^{\mathbf{u}}[\mathbf{F}, \mathbf{K}]$*, respectively.*

Without specifying a controller $\mathbf{K}$, one could alternatively consider the *realizable* CLMs of (3.2a) for *some* causal controller $\mathbf{K}'$. We call a composite operator $\mathbf{\Psi} = (\mathbf{\Psi^x}, \mathbf{\Psi^u}) \in C(\ell^n, \ell^n \times \ell^m)$ that maps $\mathbf{w} \mapsto (\mathbf{\Psi^x}(\mathbf{w}), \mathbf{\Psi^u}(\mathbf{w}))$ *realizable CLMs* for open-loop dynamic (3.2a) if there exists a so-called *realizing* controller $\mathbf{K}'$ such that $\mathbf{\Psi} = \mathbf{\Phi}[\mathbf{F}, \mathbf{K}']$. With this notion of realizable CLMs of an open-loop dynamics, we define the space of all realizable CLMs:

Definition 3.3.2 (Space of Realizable CLMs). *Given an open-loop dynamics* (3.2a), *the set of all feasible closed loop maps* $\overline{\mathbf{\Phi}}[\mathbf{F}] \in C(\ell^n, \ell^n \times \ell^m)$ *for open loop* (3.2a) *is defined as*

$$\overline{\mathbf{\Phi}}[\mathbf{F}] := \{\mathbf{\Psi} : \exists \mathbf{K} \in C(\ell^n, \ell^m) \text{ such that } \mathbf{\Psi} = \mathbf{\Phi}[\mathbf{F}, \mathbf{K}]\}.$$

The following theorem characterizes the space of realizable CLMs for a given open loop:

Theorem 8 (Characterization of CLMs [92]). *A composite operator* $\mathbf{\Psi} = (\mathbf{\Psi^x}, \mathbf{\Psi^u}) \in C(\ell^n, \ell^n \times \ell^m)$ *are realizable CLMs of the open loop* (3.2a) *if and only if they satisfy the operator equation*

$$\mathbf{\Psi^x} = \mathbf{F}(\mathbf{\Psi}) + \mathbf{I}. \tag{3.3}$$

Moreover, for any operators $\mathbf{\Psi}$ *satisfying* (3.3), *the inverse* $(\mathbf{\Psi^x})^{-1}$ *exists, is a causal operator, and* $\mathbf{K} = \mathbf{\Psi^u}(\mathbf{\Psi^x})^{-1}$ *is a realizing controller for CLMs* $\mathbf{\Psi}$. *If* $\mathbf{\Psi^u}$ *is surjective, then* $\mathbf{K}$ *is unique.*

System Level Implementations

Aside from the technical assumption on the codomain of $\mathbf{\Psi^u}$, Theorem 8 states that there is a one-to-one relation between CLMs $(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ and their realizing controllers $\mathbf{K} = \mathbf{\Psi^u}(\mathbf{\Psi^x})^{-1}$. Nevertheless, different implementations of $\mathbf{K}$ need to be distinguished: despite realizing the same CLMs with respect to the trajectory $(\mathbf{w}, \mathbf{x}, \mathbf{u})$, they do not give the same closed loop behavior once we add additional perturbations to the system. We will denote the following realization of $\mathbf{K} = \mathbf{\Psi^u}(\mathbf{\Psi^x})^{-1}$ as the *System Level* (SL)-implementation of $\mathbf{K}$:

Definition 3.3.3 (SL Implementation). *Given a composite operator* $\mathbf{\Psi} = (\mathbf{\Psi^x}, \mathbf{\Psi^u}) \in C(\ell^n, \ell^n \times \ell^m)$ *satisfying* (3.3), *the realizing controller* $\mathbf{K} = \mathbf{\Psi^u}(\mathbf{\Psi^x})^{-1}$ *can be implemented as follows*

$$u_t = \Psi_t^u(\widehat{w}_{t:0}) \tag{3.4a}$$

$$\widehat{w}_{t+1} = x_{t+1} - \Psi^x_{t+1}(0, \widehat{w}_{t:0}) \tag{3.4b}$$

for $t = 0, 1, \ldots$, where $\widehat{w}$ denotes the internal state of the controller with initial condition $\widehat{w}_0 = x_0$. We will write $\mathbf{K} = \mathrm{SL}(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ to underscore that the controller $\mathbf{K} = \mathbf{\Psi^u}(\mathbf{\Psi^x})^{-1}$ is implemented in this fashion.

Now consider the closed loop of (3.1) and controller $\mathbf{K} = \mathrm{SL}(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ perturbed by additional noise $\mathbf{v}$ and input disturbance $\mathbf{d}$ such that

$$x_t = f(x_{t-1}, u_{t-1}) + w_t \tag{3.5a}$$

$$u_t = \Psi^u_t(\widehat{w}_{t:0}) + d_t \tag{3.5b}$$

$$\widehat{w}_t = x_t - \Psi^x_t(0, \widehat{w}_{t-1:1}) + v_t. \tag{3.5c}$$

The following result characterizes the closed-loop stability under the realizing controller.

Theorem 9 (Internal Stability of Closed Loop [92])**.** *If f is uniformly continuous and the operator $(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ is ℓ_p^{n+m}-stable (or ℓ_p^{n+m} finite gain-stable) CLMs of (3.1), then the closed loop dynamics (3.5) are ℓ_p^{n+m}-stable (or finite gain ℓ_p^{n+m}-stable) with respect to the perturbation $(\mathbf{w}, \mathbf{d}, \mathbf{v})$.*

Relation to Linear System Level Approach

If we restrict the previous analysis to linear time-invariant (LTI) systems and controllers, we recover the results of [73] for the state-feedback case. If the open-loop dynamics now is $x_t = Ax_{t-1} + Bu_{t-1} + w_t$ and $\mathbf{K}$ is an LTI operator, then the corresponding linear CLMs are LTI as well. In this case, we recover the CLMs introduced in Section 2.3, where the component functions of the CLMs can be written as

$$\Psi^x_t(w_{t:0}) = \sum_{k=1}^{t+1} R_k w_{t+1-k} \tag{3.6a}$$

$$\Psi^u_t(w_{t:0}) = \sum_{k=1}^{t+1} M_k w_{t+1-k}, \tag{3.6b}$$

where $R_k \in \mathbb{R}^{n \times n}$ are the kernels of the state-feedback linear Sytem Level Synthesis (SLS) CLM from $\mathbf{w}$ to $\mathbf{x}$, i.e., $\mathbf{\Phi}_{xw}$, and $M_k \in \mathbb{R}^{m \times n}$ are the kernels of $\mathbf{\Phi}_{uw}$[1] (c.f. Section 2.3). Moreover, the corresponding CLMs characterization condition (3.3)

---

[1] In this chapter, we adopted the convention of causal relationship between $\mathbf{x}$ and $\mathbf{u}$ in (3.1), where x_t is influenced by w_t, in order to simplify arguments regarding operators. This causes a one-step time

reduces to the affine constraint on the matrices R_k, M_k which coincides with the linear SLS feasibility conditions (2.6). In particular, if we further restrict $\boldsymbol{\Psi}$ to have Finite Impulse Response (FIR) with horizon T, *i.e.*, component functions Ψ_t^x and Ψ_t^u only depend on the past $\min\{T, t+1\}$ inputs, then (3.6) becomes

$$\Psi_t^x(w_{t:0}) = \sum_{k=1}^{\min\{t+1,T\}} R_k w_{t+1-k} \tag{3.7a}$$

$$\Psi_t^u(w_{t:0}) = \sum_{k=1}^{\min\{t+1,T\}} M_k w_{t+1-k} . \tag{3.7b}$$

The CLMs characterization (3.3) in the FIR LTI case reduces to the following conditions on the kernel matrices R_k, M_k for $k = 1, \ldots, T - 1$:

$$R_1 = I \tag{3.8a}$$

$$R_{k+1} = AR_k + BM_k \tag{3.8b}$$

$$AR_T + BM_T = 0 . \tag{3.8c}$$

Further, $\mathbf{K} = \mathrm{SL}(\boldsymbol{\Psi^x}, \boldsymbol{\Psi^u})$ results in the implementation below, which also coincides with the linear SLS controller (2.7):

$$u_t = \sum_{k=1}^{\min\{t+1,T\}} M_k \widehat{w}_{t+1-k}$$

$$\widehat{w}_{t+1} = x_{t+1} - \sum_{k=2}^{\min\{t+2,T\}} R_k \widehat{w}_{t+2-k},$$

for all $k = 0, 1, \ldots$ with $\widehat{w}_0 = x_0$.

3.4 Nonlinear Blending of Linear System Level Controllers

As introduced in the previous section, system level controllers defined in Definition 3.3.3 can implement arbitrary CLMs for nonlinear systems of the form (3.1). The results in [92] motivate a new approach for nonlinear control synthesis: Searching for stable operators $\boldsymbol{\Psi^x}$, $\boldsymbol{\Psi^u}$ that satisfy (3.3) and constructing a corresponding system level controller $\mathrm{SL}(\boldsymbol{\Psi^x}, \boldsymbol{\Psi^u})$ by Definition 3.3.3.

It is conceivable that the generality of this approach could lead to an entirely new direction of nonlinear dynamic control methods. Serving as a first step towards

index shift from the classic convention where x_t is influenced by w_{t-1}, which is used in Chapter 2. Therefore, the NLSLS CLMs are defined to be causal, while the linear SLS CLMs are strictly causal. However, different indexing conventions do not affect the equivalence between the two since one can simply re-label the indices of the disturbances and the kernel matrices in one convention without loss of generality to match the other convention.

exploring the potential of this new perspective, the remainder of this chapter focuses on a subset of nonlinear system level controllers $SL(\Psi^x, \Psi^u)$ that proves particularly useful for controlling large-scale linear systems subject to state/input constraints and input saturation.

In particular we will restrict ourselves to the class of controllers $SL(\Psi^x, \Psi^u)$ where Ψ^x and Ψ^u are structured as

$$\Psi_t^x(\cdot) = \sum_{i=1}^{N} \sum_{k=1}^{\min\{T,t+1\}} R_k^{(i)} (P_{\eta_i} - P_{\eta_{i-1}})(w_{t+1-k})$$

$$\Psi_t^u(\cdot) = \sum_{i=1}^{N} \sum_{k=1}^{\min\{T,t+1\}} M_k^{(i)} (P_{\eta_i} - P_{\eta_{i-1}})(w_{t+1-k}). \tag{3.9}$$

We choose $\eta_N \geq \eta_{N-1} \geq \cdots \geq \eta_0 = 0$ and the operator $P_{\eta_i}(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ as any nonlinear function with a projection-like property defined for parameter η_i. $R_k^{(i)} \in \mathbb{R}^{n \times n}$, $M_k^{(i)} \in \mathbb{R}^{m \times n}$ are the kernel matrices associated with linear FIR CLMs $\Psi^{x,i}, \Psi^{u,i}$, $i \in [N]$ with FIR horizon T for a linear system of interests:

$$x_t = Ax_{t-1} + Bu_{t-1} + w_t, \tag{3.10}$$

with $x_t \in \mathbb{R}^n, w_t \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ such that for each $i \in \{N\}$, $\Psi^{x,i}, \Psi^{u,i}$ satisfies (3.8). Concretely, we consider two specific nonlinear projections:

Definition 3.4.1 (Saturation Projection)**.** *Let vector $w = [w^1, \ldots, w^n]^\top \in \mathbb{R}^n$. The saturation projection is an element-wise projection:*

$$P_\eta(w) := \begin{bmatrix} \text{sat}(w^1, \eta) \\ \vdots \\ \text{sat}(w^n, \eta) \end{bmatrix} \tag{3.11}$$

where $\text{sat}(w, \eta) = sign(w) \max\{|w|, \eta\}$.

Definition 3.4.2 (Radial Projection)**.** *The radial projection is defined as*

$$P_\eta(w) := \frac{\text{sat}(|w|/\eta, 1)}{|w|/\eta} w \tag{3.12}$$

Unless otherwise specified, the results derived in the rest of the chapter hold for both projections.

Remark 1. *For $n = 1$, radial projection and saturation projection coincide with each other. The radial and saturation projection operator act as the identity whenever $|w| \leq \eta$. Otherwise, the radial projection re-scales w such that $|P_\eta(w)| = \eta$ whereas the saturation projection performs element-wise radial projection.*

The proposed nonlinear controller $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ can be thought of as a *nonlinear blend* of the *linear* FIR controllers $SL(\mathbf{\Psi^{x,i}}, \mathbf{\Psi^{u,i}})$, $i \in [N]$. Although the nonlinear operator $\mathbf{\Psi^x}$, $\mathbf{\Psi^u}$ differs from its linear components $\mathbf{\Psi^{x,i}}$, $\mathbf{\Psi^{u,i}}$ only by the static nonlinear function $P_{\eta_i}(w)$, the upcoming sections will demonstrate that this simple additional nonlinearity proves surprisingly useful. In particular, η_i's separate any disturbance w_t into N zones such that for each ith linear controller $SL(\mathbf{\Psi^{x,i}}, \mathbf{\Psi^{u,i}})$, only the portion of w_t that "falls" between η_i and η_{i-1} is acted upon. Intuitively, one could choose different behaviors for various portions of the disturbance signal, specifying either performance or safety properties. The explicit expression of the dynamic controller $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ with CLMs defined in (3.9) is

$$u_t = \sum_{i=1}^{N} \sum_{k=1}^{\min\{T,t+1\}} M_k^{(i)} (P_{\eta_i} - P_{\eta_{i-1}})(\widehat{w}_{t+1-k})$$

$$\widehat{w}_{t+1} = x_{t+1} - \sum_{i=1}^{N} \sum_{k=2}^{\min\{T,t+2\}} R_k^{(i)} (P_{\eta_i} - P_{\eta_{i-1}})(\widehat{w}_{t+2-k}),$$

with $k = 0, 1, \ldots$, and $\widehat{w}_0 = x_0$.

For ease of exposition, we focus on the two-zone case of the proposed controller $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ though all the analysis extends naturally to the N-zone case. Thus, (3.9) simplifies to

$$\Psi_t^x(w_{t:0}) = \sum_{k=1}^{\min\{T,t+1\}} R_k^{(1)} P_{\eta_1}(w_{t+1-k}) +$$

$$R_k^{(2)} (P_{\eta_2}(w_{t+1-k}) - P_{\eta_1}(w_{t+1-k}))$$

$$\Psi_t^u(w_{t:0}) = \sum_{k=1}^{\min\{T,t+1\}} M_k^{(1)} P_{\eta_1}(w_{t+1-k}) +$$

$$M_k^{(2)} (P_{\eta_2}(w_{t+1-k}) - P_{\eta_1}(w_{t+1-k})). \tag{3.13}$$

Note that system level controller $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ of the two-zone CLM is internally stabilizing and achieves the two-zone CLM behavior for (3.10) as long as $\|\mathbf{w}\|_\infty \leq \eta_2$.

In the remainder of this chapter we will explore the consequence of this blending technique for distributed control design with respect to input saturation and state constraints in linear systems. we show that the simple nonlinearity in (3.13) offers a variety of advantages over linear controllers, including the ones presented in Chapter 2.

3.5 A General Framework for Constrained LQR

We present a novel synthesis procedure for a class of constrained LQR problems using the proposed SL controller with CLMs (3.13). In particular, we will show that the synthesized nonlinear blending system level controller is guaranteed to outperform any linear controller for the class of constrained LQR problems to be discussed. Additionally, we comment on how structural constraints for large-scale systems such as delay, actuation sparsity, and localization can be easily accommodated.

Consider a control problem where we wish to minimize an *average* LQR cost, but also want that the closed loop meets certain safety guarantees against a set of rare yet possible *worst-case* disturbances. Ideally, we would like to synthesize a controller that can guarantee the necessary safety constraints without too much loss in performance compared to the unconstrained LQR controller. We will phrase this design goal as the following constrained LQR problem:

$$\min_{\mathbf{K}} \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{w_t^i \sim p(w)}[\mathcal{J}(x_t, u_t)] \tag{3.14a}$$

$$s.t. \qquad x_t = A x_{t-1} + B u_{t-1} + w_t \tag{3.14b}$$

$$u_t = K_t(x_{t:0}) \tag{3.14c}$$

$$\forall \mathbf{w} : ||\mathbf{w}||_\infty \le \eta_{max} : \tag{3.14d}$$

$$\sup_k |x_k| \le x_{max} \qquad \sup_k |u_k| \le u_{max},$$

where $\mathcal{J}$ abbreviates the quadratic stage cost $\mathcal{J}(x, u) = x^\top Q x + u P u$ with $Q, P > 0$. We will assume that the disturbance is stochastic but bounded such that $||\mathbf{w}||_\infty \le \eta_{max}$ with known distribution which satisfies the following

Assumption 2. *Disturbance w_t^i are i.i.d. drawn from the scalar centered distribution $p(w)$ and uncorrelated in time t and coordinate i.*

We can equivalently phrase the optimal control problem (3.14) in terms of closed-loop mappings as defined in Section 3.3. Recalling Definition 3.3.1, the optimal control problem (3.14) can be described as an optimization over the set of feasible CLMs $(\mathbf{\Psi^x}, \mathbf{\Psi^u}) \in \overline{\mathbf{\Phi}}(\mathbf{Ax + Bu})$ and by using the characterization in Theorem 8:

$$\min_{\mathbf{\Psi^x}, \mathbf{\Psi^u}} \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\mathcal{J}(\Psi_t^x(w_{t:0}), \Psi_t^u(w_{t:0}))] \tag{3.15a}$$

$$s.t. \qquad \Psi_t^x(w_{t:0}) = \Psi_t^x(0, w_{t-1:0}) + w_t \tag{3.15b}$$

$$\Psi_{t+1}^x(0, w_{t:0}) = A\Psi_t^x(w_{t:0}) + B\Psi_t^u(w_{t:0})$$

$$\forall t, |w_t| \le \eta_{max} : \quad |\Psi_t^x(w_{t:0})| \le x_{max} \tag{3.15c}$$

$$\forall t, |w_t| \le \eta_{max} : \quad |\Psi_t^u(w_{t:0})| \le u_{max} . \tag{3.15d}$$

As in the linear SLS case [73], we do not need to have the controller $\mathbf{K}$ be a decision variable, since we can always realize the optimal solution $(\mathbf{\Psi}^{\mathbf{x},\star}, \mathbf{\Psi}^{\mathbf{u},\star})$ to (3.15) with a system level controller SL($\mathbf{\Psi}^{\mathbf{x},\star}, \mathbf{\Psi}^{\mathbf{u},\star}$).

Conservativeness of Linear Solutions

We will first discuss properties of solutions to our original problem (3.14), if we restrict ourselves to only LTI controllers $\mathbf{K}$. Consider the equivalent problem formulation (3.15) with the CLMs ($\mathbf{\Psi}^{\mathbf{x}}, \mathbf{\Psi}^{\mathbf{u}}$) restricted to be linear. This poses a convex problem and as shown in [86], it can be approximately solved by searching over FIR CLMs ($\mathbf{\Psi}^{\mathbf{x}}, \mathbf{\Psi}^{\mathbf{u}}$). Yet, the corresponding linear CLMs $\{\mathbf{\Psi}^{\mathbf{x},\mathbf{lin},\star}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin},\star}\}$ come with undesirable restrictions:

- $\{\mathbf{\Psi}^{\mathbf{x},\mathbf{lin},\star}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin},\star}\}$ impose stricter safety constraints than the required constraints (3.15c) and (3.15d).

- $\{\mathbf{\Psi}^{\mathbf{x},\mathbf{lin},\star}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin},\star}\}$ do not depend on the disturbance distribution $p(w)$.

To see the first point, we have the following result as a consequence of linearity:

Lemma 2. *For any linear* $\{\mathbf{\Psi}^{\mathbf{x},\mathbf{lin}}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin}}\}$*, the constraint* (3.15c),(3.15d) *is equivalent to*

$$\sup_t |\Psi_t^{x,lin}(w_{t:0})| \le \sup_t \frac{x_{max}}{\eta_{max}}|w_t| \tag{3.16a}$$

$$\sup_t |\Psi_t^{u,lin}(w_{t:0})| \le \sup_t \frac{u_{max}}{\eta_{max}}|w_t|. \tag{3.16b}$$

Proof. Clearly, (3.16) implies (3.15c),(3.15d). The reverse implication follows by the assumed linearity of $\{\mathbf{\Psi}^{\mathbf{x},\mathbf{lin}}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin}}\}$ and homogeneity of norms. □

Lemma 2 shows that the restriction of linearity in CLMs imposes stricter safety conditions (3.16) than (3.15c),(3.15d). To elaborate on the second point, notice that for linear CLMs $\mathbf{\Psi}^{\mathbf{x},\mathbf{lin}}, \mathbf{\Psi}^{\mathbf{u},\mathbf{lin}}$ our objective function (3.15a) can be expressed equivalently as

$$(3.15a) = \sigma^2 \left\| \begin{matrix} Q^{1/2}\mathbf{\Psi}^{\mathbf{x}} \\ P^{1/2}\mathbf{\Psi}^{\mathbf{u}} \end{matrix} \right\|_{\mathcal{H}_2}^2, \quad \sigma^2 := \mathbb{E}_{w\sim p(w)}[w^2] \tag{3.17}$$

where σ^2 denotes the variance of the scalar distribution $p(w)$ and $\|.\|_{\mathcal{H}_2}$ denotes the $\mathcal{H}_2$ norm for linear operators. Since the objective function only gets scaled by a constant factor σ^2 for different distributions $p(w)$, this shows that for linear CLMs, the solutions $\{\mathbf{\Psi}^{\mathbf{x,lin}}, \mathbf{\Psi}^{\mathbf{u,lin}}\}$ to (3.15) are independent of the distribution $p(w)$.

A NLSLS Take

Consider the general problem (3.15), where now we search over CLMs $(\mathbf{\Psi}^{\mathbf{x}}, \mathbf{\Psi}^{\mathbf{u}})$ of the form presented in (3.13) with the choice of $\eta_2 = \eta_{max}$, some $\eta_1 < \eta_2$, and an FIR horizon T. Recall that $(\mathbf{\Psi}^{\mathbf{x}}, \mathbf{\Psi}^{\mathbf{u}})$ is a blending of two linear CLMs and has the form (3.13). Restricting ourselves to this form of CLMs allows us to derive the following convex problem which is a relaxation of the general problem (3.15):

$$\min_{R^{(i)}, M^{(i)}} \left\| \begin{bmatrix} Q & 0 \\ 0 & P \end{bmatrix}^{1/2} \begin{bmatrix} R^{(1)} & R^{(2)} \\ M^{(1)} & M^{(2)} \end{bmatrix} \Sigma_w^{1/2} \right\|_F^2 \tag{3.18a}$$

$$s.t. \quad \eta_1 |R^{(1)}| + (\eta_2 - \eta_1)|R^{(2)}| \leq x_{max} \tag{3.18b}$$

$$\eta_1 |M^{(1)}| + (\eta_2 - \eta_1)|M^{(2)}| \leq u_{max} \tag{3.18c}$$

$$R_{k+1}^{(i)} = AR_k^{(i)} + BM_k^{(i)} \tag{3.18d}$$

$$R_1^{(i)} = I, \quad R_T^{(i)} = 0,$$

where

$$\Sigma_w = \begin{bmatrix} \alpha_1 I & \alpha_2 I \\ \alpha_2 I & \alpha_3 I \end{bmatrix}$$

with $\alpha_1 = \mathbb{E}[P_{\eta_1}(w)^2]$, $\alpha_2 = \mathbb{E}[P_{\eta_1}(w)(P_{\eta_2}(w) - P_{\eta_1}(w))]$, and $\alpha_3 = \mathbb{E}[(P_{\eta_2}(w) - P_{\eta_1}(w))^2]$, where $w \sim p(w)$ and $\|\mathbf{w}\|_\infty \leq \eta_{max}$. Moreover $R^{(i)}$ and $M^{(i)}$ are abbreviations for the horizontal concatenation of the kernel matrices associated with the linear CLMs $\mathbf{\Psi}^{\mathbf{x,i}}, \mathbf{\Psi}^{\mathbf{u,i}}$, *i.e*, $R^{(i)} = [R_T^{(i)}, R_{T-1}^{(i)}, \ldots, R_1^{(i)}]$, $M^{(i)} = [M_T^{(i)}, M_{T-1}^{(i)}, \ldots, M_1^{(i)}]$. Hereby, only constraints (3.18b), (3.18c) are sufficient condition of the constraint (3.15c), (3.15d) via norm multiplicativity. All other equations in the above optimization are equivalent to the original problem (3.15) restricting the search over CLMs of the form (3.13). Finally, solving the convex problem (3.18) gives the suboptimal nonlinear CLMs $\{\mathbf{\Psi}^{\star\mathbf{x}}, \mathbf{\Psi}^{\star\mathbf{u}}\}$ for the system dynamics (3.14b), which can be realized with the internally stabilizing system level controller SL($\mathbf{\Psi}^{\star\mathbf{x}}, \mathbf{\Psi}^{\star\mathbf{u}}$). The next theorem states a main result of this chapter:

Theorem 10. *For all* $\eta_1 \in [0, \eta_2]$, *the nonlinear system level controller* $\mathrm{SL}(\mathbf{\Psi^{\star x}}, \mathbf{\Psi^{\star u}})$ *synthesized from* (3.18) *achieves lower optimal LQR cost for* (3.14) *than any linear solutions.*

Proof. First, recall that restricting **K** to be linear in problem (3.14) is equivalent to restricting $\mathbf{\Psi^x}$ and $\mathbf{\Psi^u}$ to be linear in the equivalent formulation (3.15). Furthermore, notice that under the restriction of linear $(\mathbf{\Psi^x}, \mathbf{\Psi^u})$, problem (3.15) is equivalent to (3.18) with the added constraint $R^{(1)} = R^{(2)}$, $M^{(1)} = M^{(2)}$, which shows that any solution $(\mathbf{\Psi^{\star x}}, \mathbf{\Psi^{\star u}})$ of problem (3.18) achieve smaller cost than a linear solution $(\mathbf{\Psi^{x,lin,\star}}, \mathbf{\Psi^{u,lin,\star}})$ of (3.15). □

The above argument also extends directly to the N-blend case.

Distributed Controller for Constrained LQR

Thanks to the particular form of (3.9), when the projection is chosen to be the saturation projection Definition 3.4.1, structural constraints of controller such as sparsity and delay constraints can be added in a convex way to the synthesis procedure described in Section 3.5. This is because imposing structural constraints on the nonlinear controller (3.9) is equivalent to imposing them on the linear CLM components of (3.9). As discussed in Chapter 2, localization of disturbances, as well as communication and actuation delays are all convex constraints in terms of linear CLMs in the linear SLS framework. Specifically, all mentioned constraints could be cast as a convex system level constraints (SLCs) for linear CLMs $\mathbf{\Psi^{x,i}}, \mathbf{\Psi^{u,i}}$, $i \in [N]$. The corresponding system level controller $\mathrm{SL}(\mathbf{\Psi^u}, \mathbf{\Psi^x})$ can then be implemented in a localized fashion conforming to the SLCs on $\mathbf{\Psi^{x,i}}, \mathbf{\Psi^{u,i}}$. Therefore, the nonlinear controller synthesis in Section 3.5 naturally inherits all capabilities of the linear system level controllers in terms of distributed controller synthesis and implementation.

3.6 Distributed Anti-windup Controller for Saturated Systems

Now consider a linear input saturated system where the disturbances and initial condition are *not* necessarily constrained to have a known norm bound η_{max}. The control actions are projected via saturation projector:

$$x_t = Ax_{t-1} + BP_{u_{max}}(u_{t-1}) + w_t, \tag{3.19}$$

In this scenario, controller $\mathrm{SL}(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ previously constructed with (3.13) *no longer* realizes the designed closed-loop response (3.13) for (3.19). Nevertheless, we would like the saturated system to degrade gracefully and preserve stability. Such property

is traditionally achieved via anti-windup design[91]. Here, we show that the proposed nonlinear controller achieves natural anti-windup property with little modification.

Anti-windup Controller

Inspired by internal model control (IMC) [93], we modify $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ and consider an augmented controller $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ where the operator $\mathbf{\Psi^{x,a}}$ is constructed from (3.13) with augmentation:

$$
\mathbf{\Psi}_t^{\mathbf{x,a}}(w_{t:0}) = \sum_{i=1}^{N} \left(\sum_{k=1}^{\min\{T,t+1\}} R_k^{(i)} (P_{\eta_i} - P_{\eta_{i-1}})(w_{t+1-k}) \right)
$$
$$
+ \sum_{k=1}^{\tau+1} A^{k-1} \left(w_{t+1-k} - P_{\eta_N}(w_{t-k+1}) \right), \qquad (3.20)
$$

where τ is a positive integer. Recall that by design, we have chosen $\eta_N = \eta_{max}$, the expected norm bound on disturbances. Compared to (3.13) for the two-zone case ($N = 2$) and (3.9) for the N-zone case, we note that (3.20) has the additional "open-loop" dynamics term. This extra term accounts for the residual disturbances that are not attenuated by the original controller $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ because the disturbances are larger than expected by the projection mapping, *i.e.*, $|w_t| > \eta_{max}$. Therefore, $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ considers the τ-step propagation of the unaccounted disturbances from $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$. Note that when the disturbances satisfy the assumption $\|\mathbf{w}\|_\infty \leq \eta_{max}$, augmented controller $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ is identical to $SL(\mathbf{\Psi^x}, \mathbf{\Psi^u})$ constructed from (3.9) and (3.13).

The IMC-like structure of the augmented controller $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ helps the saturated system to degrade gracefully and preserve stability even when $\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u}$ are not the exact CLMs for the closed-loop system. The closed-loop dynamics of (3.19) under augmented controller $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ from (3.20) can be checked to be

$$
\widehat{w}_t = A^{\tau+1}(\widehat{w}_{t-\tau} - P_{\eta_{max}}(\widehat{w}_{t-\tau})) + w_t. \qquad (3.21)
$$

As shown in [92], the stability of the overall closed loop is equivalent to the stability of (3.21). We now certify the anti-windup property of $SL(\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u})$ with the following result.

Lemma 3. *If τ satisfies $|A^{\tau+1}| < 1$, then internal dynamics (3.21) is globally finite-gain ℓ_∞-stable where for all $\mathbf{w} \in \ell_\infty^n$,*

$$
\|\widehat{\mathbf{w}}\|_\infty \leq \frac{1}{1 - |A^{\tau+1}|} \|\mathbf{w}\|_\infty .
$$

Proof. We first present an operator small-gain theorem.

Theorem 11 (Small-gain Theorem[92])**.** *Let* $\Delta \in C_s(\ell^n, \ell^n)$. *If for all* $\mathbf{x} \in \ell_p^n$, $\|\Delta(\mathbf{x})\|_p \leq \gamma \|\mathbf{x}\|_p + \beta$ *with* $0 < \gamma < 1$, $\beta \geq 0$, $p = 1, 2, \ldots, \infty$, *then for all* $\mathbf{w} \in \ell_p^n$, $\|\widehat{\mathbf{w}}\|_p \leq \frac{1}{1-\gamma}(\|\mathbf{w}\|_p + \beta)$ *where* $\widehat{\mathbf{w}} = (I - \Delta)^{-1}\mathbf{w}$.

Note that the inverse exists because $\Delta \in C(\ell^n, \ell^n)$ [92]. We are now in a position to prove the lemma. We can write (3.21) in the operator form as

$$\widehat{\mathbf{w}} = (I - \Delta)^{-1}\mathbf{w}, \tag{3.22}$$

where Δ is a strictly causal operator with component function $\Delta_t(\widehat{w}_{t:0}) := A^{\tau+1}(\widehat{w}_{t-\tau} - P_{\eta_{max}}(\widehat{w}_{t-\tau}))$. For all $\widehat{\mathbf{w}} \in \ell_\infty^n$, $\|\Delta(\widehat{\mathbf{w}})\|_\infty \leq |A^{\tau+1}| \|\widehat{\mathbf{w}}\|_\infty$ where we have chosen τ such that $|A^{\tau+1}| < 1$. Therefore, invoking Theorem 11 gives the desired result in Lemma 3. $\qquad\square$

In particular, if A is schur, then there exists $k \in \mathbb{N}$ such that $\|A^k\| < 1$ for any norm. Therefore, if (3.19) is open-loop stable, SL($\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u}$) guarantees graceful degradation when the closed loop is saturated.

Localized Implementation

Similar to the large-scale constrained LQR case in Section 3.5, since the anti-windup controller SL($\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u}$) for the saturated linear system (3.19) is composed of linear CLMs synthesized from (3.18) with locality constraints, localization can be easily imposed as a convex subspace constraint on the composing linear CLMs. When the the information structure of the controllers are constrained to the state propagation pattern according to open-loop dynamics *i.e.*, the sparsity of A, the anti-windup controller SL($\mathbf{\Psi^{x,a}}, \mathbf{\Psi^u}$) can be implemented in a localized fashion where information is exchanged and disturbance is contained in a local controller patch [44]. As will be illustrated in Section 3.7, this allows for distributed anti-windup controller design for large-scale saturated systems.

3.7   Simulation

Constrained LQR

To corroborate the results presented in the previous sections, we demonstrate the performance of a four-zone nonlinear blending controller with radial projection compared against the optimal linear controller for the constrained LQR problem of

an open-loop unstable system:

$$x_{k+1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_k + w_k \tag{3.23}$$

with $u_{max} = 40$, $x_{max} = 15$, $\eta_{max} = 1$, $Q = I_3$, $P = 10$. The disturbances w_k are chosen to be a truncated i.i.d. gaussian random variables with variance σ^2. Figure 3.1 shows the optimal cost improvement of the presented nonlinear approach over the optimal linear controller for different choices of variance σ^2. Figure 3.1 showcases that the proposed controller can exploit the knowledge of the disturbance distribution to achieve performance improvement over the linear optimal linear controller: for small σ the proposed controller gains more than 30% cost reduction over safe controller. On the other hand, with increasing σ, large disturbances in the system become more likely, and therefore the opportunity to improve upon the linear optimal controller is reduced.
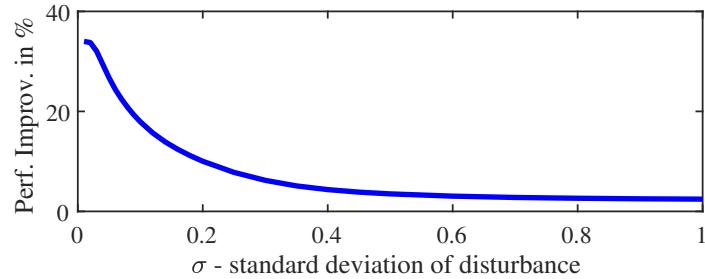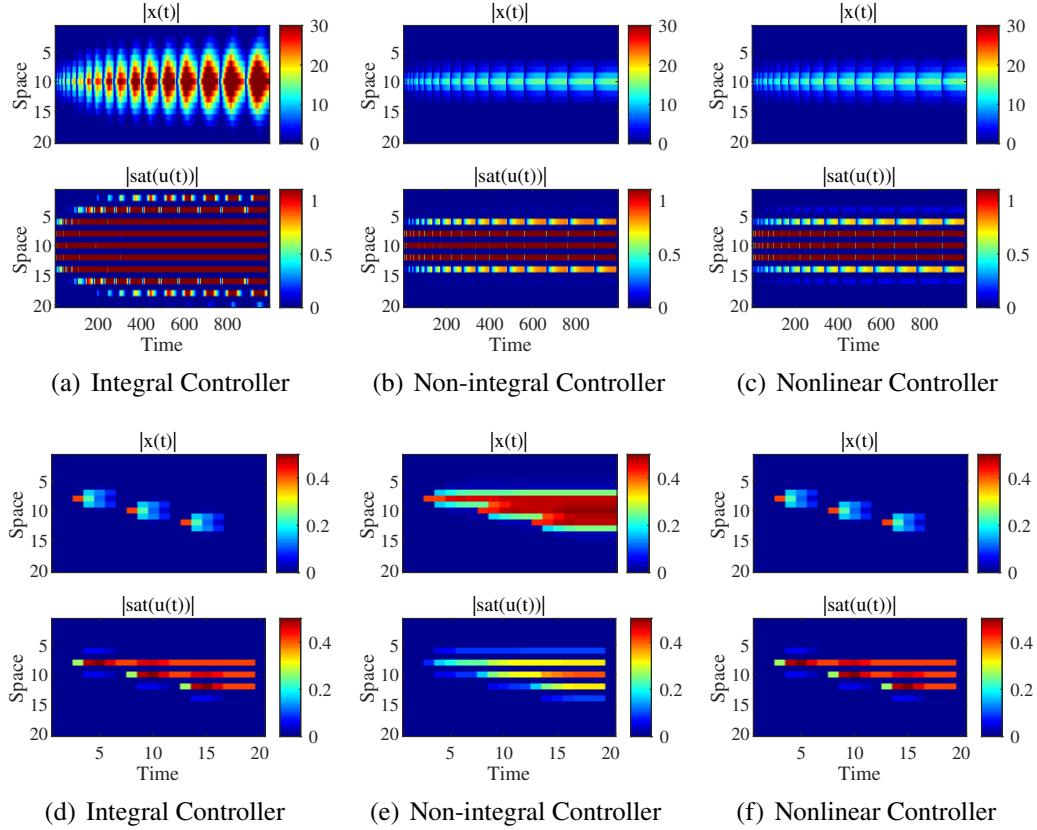
Figure 3.1: Performance improvement of optimal nonlinear controller SL($\Psi^{\star x}$, $\Psi^{\star u}$) over optimal linear controller SL($\Psi^{x,lin,\star}$, $\Psi^{u,lin,\star}$) for different variances σ^2 of the non-truncated disturbance. The nonlinear blending controller synthesizes over 4 linear controllers w.r.t. to the projection parameters $\eta_1 = 0.05, \eta_2 = 0.1, \eta_3 = 0.2, \eta_4 = \eta_{max} = 1$.

Localized Anti-Windup Controller

Consider a bi-directional chain system with ith node's dynamics being

$$x_{t+1}^i = (1 - 0.4|\mathcal{N}(i)|)x_t^i + 0.4 \sum_{j \in \mathcal{N}_i} x_t^j + \text{sat}(u_t^i, u_{max}) + w_t^i,$$

where $\mathcal{N}(i)$ denotes the set of vertices that has an edge connected to ith vertex and w_t^i is the ith coordinate of disturbance vector at time t. In particular, $\|\mathbf{w}\|_\infty \leq 1$ and $x_0 = 0$. One can check that the overall chain system is open-loop marginally stable.

Figure 3.2: ***Worst-case (top row) and Staggered Step Input (bottom row) Response*** *under Saturation for a marginally stable 20-node Chain System with Sparse Actuation*. *Top Row*: The heatmaps show how a worst-case disturbance is propagated through space-time for the saturated chain system. The integral controller becomes unstable due to saturation and the naive blending controller possesses has the anti-windup property of the non-integral controller. In addition to anti-windup, the proposed controller is localized and accommodates sparse actuation, communication delay, and controller sparsity constraints. Here every other node has a control input (50% actuation) with 1 time step actuation delay and 1 time step communication delay between nodes, while enforcing a controller sparsity that conforms to the communication pattern of dynamics matrix *A*. *Bottom Row*: Response to small step disturbances at node 8,10,12 entering at time 2,6,10, respectively. As in the scalar case, the proposed blending controller not only stabilizes under saturation but also recovers the performance objective of rejecting small step disturbances. This contrasts against the non-integral controller, which sacrifices small-signal performance for stability.

We illustrate the anti-windup property of the nonlinear controller (3.9) in the decentralized setting with additional sparsity, locality, and delay constraints. First, a nominal integral controller for this system is designed and dubbed the *Integral Controller*. Due to its integral structure, the *Integral Controller* for the unconstrained closed loop guarantees convergence of the state to the origin under persistent disturbance, *i.e*, step rejection. In comparison, a second linear controller synthesized from standard constrained LQR problem that guarantees stability for all admissible **w** under saturation is generated. We refer to this linear controller as the *Non-integral Controller* since the states only stay bounded under persistent admissable disturbance.

The nonlinear controller with saturation projection here is chosen to be a two-zone blending controller consisted of CLMs of the form (3.13). The simulation shows the anti-windup property as well as preservation of step rejection in both large- and small-disturbance schemes of the proposed method. Figure 3.2 shows that the blending controller stabilizes the system while integral controller becomes unstable under worst-case bounded disturbance. On the other hand, the proposed blending controller preserves performance of step rejection while the linear Non-integral Controllers forfeits the performance objective in order to preserve stability in the saturated closed loop. In this chain example, we allow 1 time step communication delay between nodes and actuation delay with 50% control authority. The localization pattern imposed on the system response allows SL($\mathbf{\Psi^x}, \mathbf{\Psi^u}$) to be implemented in local patches, therefore making the controller distributed.

3.8 Conclusion

Int his chapter, we propose a tractable nonlinear distributed control synthesis method that outperforms any optimal linear controllers for the distributed and localized LQR problems under input saturation and state constraints. It was further shown that such controller naturally possesses anti-windup property for linear systems with input saturation. A key highlight is that the presented approach enjoys the same compatibility with locality and communication delay constraints and distributed implementation, as the linear system level synthesis approach presented in Chapter 2.

Part II

Interfacing Learning and Control via Uncertainty Sets

C h a p t e r   4

NON-ASYMPTOTIC LEARNING OF UNCERTAINTY SETS VIA SET MEMBERSHIP

In this chapter, we present the uncertainty set estimation method that will serve as a link between machine learning techniques and model-based control methods as introduced in Chapter 1. Uncertainty sets are crucial for the quality of robust learning and control since they directly influence the conservativeness and the safety of the algorithm.

Departing from the confidence region analysis of least squares estimation, which is one of the most common estimation methods in learning-based control literature, this chapter investigates properties of the set membership estimation (SME) method commonly seen in robust adaptive control. Though good numerical performance have attracted applications of SME in many domains including energy systems and robotics, the non-asymptotic convergence rate of SME for linear systems remains an open question. In this chapter, we will provide the first convergence rate bounds for SME and propose novel variations of SME under relaxed assumptions that improve the practicality and performance of SME. Interestingly, in some settings, SME breaks through the theoretical lower bound on sample complexity of the estimation task of linear dynamical systems previously shown for the popular ordinary least squares estimation, enabling improved performance for a range of safety-critical tasks over previous approaches.

[1]   Y. Li*, J. Yu*, L. Conger, T. Kargin, and A. Wierman, "Learning the uncertainty sets of linear control systems via set membership: A non-asymptotic analysis," *Forty-first International Conference on Machine Learning (ICML)*, 2024. [Online]. Available: https://openreview.net/forum?id=n2kq2EOHFE.

4.1   Introduction

The problem of estimating unknown linear dynamical systems of the form $x_{t+1} = A^\star x_t + B^\star u_t + w_t$ with unknown parameters $(A^\star, B^\star)$ has seen considerable progress recently [26], [58]–[60], [94]–[96]. Most literature focuses on the analysis of the least squares estimator (LSE) and its variants, where sharp bounds on the convergence rates for subGaussian disturbances w_t have been obtained [59], [96]. Building on

this, there is a rapidly growing body of literature on "learning to control" unknown linear systems that leverages LSE to achieve various control objectives, such as stability and regret [17], [24], [25], [28], [59], [97].
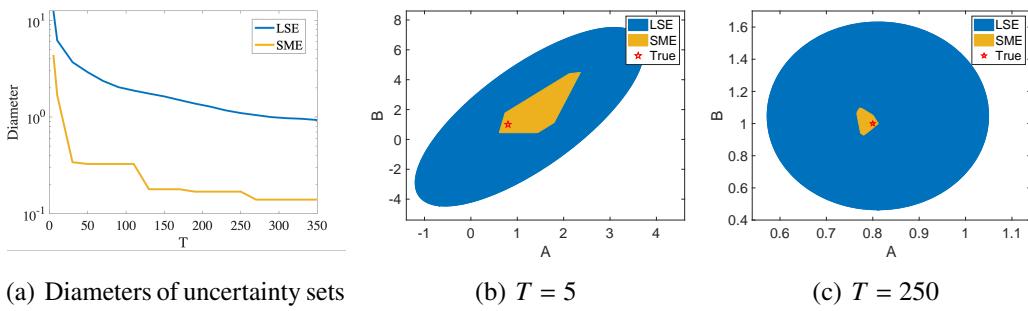
However, for successful application of learning-based control methods to safety-critical applications, it is crucial to quantify the uncertainties of the estimated system and to robustly satisfy safety constraints and stability despite these uncertainties [29], [30]. A promising framework for achieving this is to estimate the uncertainty set of the unknown system parameters and to utilize robust controllers to satisfy the robust constraints under any parameters in the uncertainty set [30], [98]. Uncertainty set estimation is crucial for the success of robust control: on the one hand, too large of an uncertainty set gives rise to over-conservative control actions, resulting in degraded performance; on the other hand, if the uncertainty set is underestimated and fails to contain the true system, the resulting controller may lead to unsafe behaviors [30], [31].

To estimate uncertainty sets, a popular method is to construct LSE's confidence regions [24], [59]. However, this approach yields a confidence region for a point estimate rather than directly estimating the uncertainty set of the model. Further, the confidence regions are usually derived from concentration inequalities, which allows convergence rate analysis but may suffer conservative constant factors [31], [59].

In this chapter, we instead focus on a direct uncertainty set estimation method: set membership estimation (SME), which estimates the uncertainty set without relying on the concentration inequalities underlying the approaches based on LSE. SME has a long history in the control community [48], [49], [99]–[102]. SME has primarily been proposed for scenarios with bounded disturbances, which is common in safety-critical systems, e.g. power systems [103], unmanned aerial vehicles (UAV) [104], [105], and building control [106]. Further, the bounded disturbance is a standard assumption in control when certain safety requirements are desired, such as robust (adaptive) constrained control [24], [49], [107], online (constrained) control [14], [108]–[110].

Consequently, SME has been widely adopted in the robust (adaptive) constrained control literature [49]–[54] and the online control literature [13], [55], [99], [111]. Figure 4.1 provides a toy example illustrating SME's promising performance under bounded disturbances.

On the theory side, the convergence analysis of SME generally considers a simple

(a) Diameters of uncertainty sets   (b) $T = 5$ (c) $T = 250$

Figure 4.1: A visualized toy example of uncertainty set comparison between SME in (4.2) and LSE confidence regions in [59], [60] for a one-dimensional system $x_{t+1} = A^\star x_t + B^\star u_t + w_t$, with $w_t, u_t \in [-1, 1]$ generated i.i.d. from a truncated Gaussian distribution. Detailed experiment settings are in Section 4.I. **Figure (a)** compares the diameters of the uncertainty sets from SME and LSE 90% confidence bounds. **Figure (b) and (c)** visualize the the uncertainty sets after $T = 5$ and $T = 250$ data points.

regression problem: $y_t = \theta^\star x_t + w_t$ with a deterministic sequence of x_t and bounded i.i.d. disturbances w_t [112]–[116]. This regression problem does not capture the correlation between x_t and the history $w_{t-1}, \dots, w_0$ in the dynamical systems. This issue was largely overlooked in the vast literature of empirical algorithm design related to SME (for example, see [49], [117], etc.). It was not until recently that [118] provided the first *asymptotic convergence* guarantees for SME in linear systems. However, the *non-asymptotic convergence rate* still remains open for SME in linear dynamical systems.

Contributions. This chapter tackles the open question above by providing non-asymptotic bounds on the convergence rates of SME for linear systems. To the best of our knowledge, this is the first convergence rate analysis of SME for dynamical systems in the literature.

We consider two scenarios in our analysis. Firstly, when a *tight* bound $\mathcal{W}$ on the support of w_t is known, we provide an instance-dependent convergence rate for SME. Interestingly, for several common distributions of w_t, SME enjoys a convergence rate $\widetilde{O}(n_x^{1.5}(n_x + n_u)^2/T)$, which is faster than the LSE's error bound $O(\frac{\sqrt{n_x+n_u}}{\sqrt{T}})$ in terms of the number of samples T but is worse in terms of the dependence on state and control dimensions n_x, n_u. The improved convergence rate of SME with respect to T is enabled by leveraging the additional boundedness property of w_t,

which is a common assumption in robust constrained control but is not utilized in LSE's analysis. Secondly, when a tight bound of w_t is *unknown*, we introduce a UCB-SME algorithm that learns conservative upper bounds of w_t from data and constructs uncertainty sets based on the conservative upper bounds. We also provide a convergence rate of UCB-SME, which has the same dependence on T but has worse dependence on n_x by a factor of $\sqrt{n_x}$ compared with the convergence rate with a known tight bound.

Our estimation error bound relies on a novel construction of an event sequence based on designing a sequence of stopping times. This construction, together with the BMSB condition in [96], addresses the challenge caused by the correlation between x_t, u_t, and the history disturbances (see the proof of Theorem 12 for more details).

Moreover, our results lay a foundation for future non-asymptotic analysis of control designs based on SME. To illustrate this, we apply our results to robust-adaptive model predictive control and robust SLS and discuss the novel non-asymptotic guarantees enabled by our convergence rates of SME.

Finally, we conduct extensive simulations to compare the numerical behaviors of SME, UCB-SME, and LSE's confidence regions, which demonstrates the promising performance of SME and UCB-SME.

4.2 Problem Formulation and Preliminaries

This chapter focuses on the identification of uncertainty sets of unknown system parameters in the linear dynamical system:

$$x_{t+1} = A^\star x_t + B^\star u_t + w_t, \tag{4.1}$$

where $A^\star, B^\star$ are the unknown system parameters, $x_t \in \mathbb{R}^{n_x}, u_t \in \mathbb{R}^{n_u}$. For notational simplicity, we define $\theta^\star = (A^\star, B^\star)$ by matrix concatenation and $z_t = (x_t^\top, u_t^\top)^\top \in \mathbb{R}^{n_z}$ by vector concatenation, where $n_z = n_x + n_u$. Accordingly, the system (4.1) can be written as $x_{t+1} = \theta^\star z_t + w_t$.

The goal of the uncertainty set identification problem is to determine a set Θ_T that contains the true parameters $\theta^\star = (A^\star, B^\star)$ based on a sequence of data $\{x_t, u_t, x_{t+1}\}_{t=0}^{T-1}$. Set Θ_T is called an uncertainty set since it captures the remaining uncertainty on the system model after the revelation of the data sequence $\{x_t, u_t, x_{t+1}\}_{t=0}^{T-1}$.

Uncertainty sets play an important role in robust control, where one aims to achieve robust constraint satisfaction [49], [107], robust objective optimization [119], and/or

robust stability [120] for any model in the uncertainty set.[1] Therefore, the diameter of the uncertainty sets heavily influences the conservativeness of robust controllers and thus the control performance. Formally, we define the diameter as follows.

Definition 4.2.1 (Diameter of a set of matrices)**.** *Consider a set $\mathbb{S}$ of matrices $\theta \in \mathbb{R}^{n_x \times n_z}$. We define the diameter of $\mathbb{S}$ in Frobenius norm as* $\operatorname{diam}(\mathbb{S}) = \sup_{\theta, \theta' \in \mathbb{S}} \|\theta - \theta'\|_F$.

Set Membership Estimation (SME)

In this section, we review set membership estimation (SME), which is an uncertainty set identification method that has been studied in the control literature for decades [102], [107]. SME primarily focuses on systems with *bounded* disturbances, i.e. $w_t \in \mathcal{W}$ for some bounded $\mathcal{W}$ for all $t \geq 0$. When $\mathcal{W}$ is known, SME computes an uncertainty/membership set by

$$\Theta_T = \bigcap_{t=0}^{T-1} \{\widehat{\theta} : x_{t+1} - \widehat{\theta} z_t \in \mathcal{W}\}. \tag{4.2}$$

It is straightforward to see that $\theta^\star \in \Theta_T$ when $w_t \in \mathcal{W}$.

The bounded disturbance assumption may seem restrictive, considering that the uncertainty set identification based on the confidence region of LSE only requires subGaussian disturbances [96]. However, in many control applications, it is reasonable and common to assume bounded w_t. For example, bounded disturbances is a standard assumption in the robust constrained control literature, such as robust constrained LQR [24], [49], [107], [118], and online constrained control of linear systems [108], [110]. This is different from unconstrained control, where unbounded subGaussian disturbances are usually considered [120]. The difference in the disturbance formulation is largely motivated by the applications: constrained control is mostly applied to safety-critical applications, where the disturbances are usually bounded. For example, in UAV and flight control, the disturbances are mostly caused by wind gusts, and wind disturbances are bounded in practice [104], [105]. Similarly, in building thermal control, the disturbances are caused by external heat exchanges, which are also bounded [106].

Ideally, one hopes that Θ_T converges to the singleton of the true model $\{\theta^\star\}$ or at least a small neighborhood of $\theta^\star$. This usually calls for additional assumptions, such

---

[1]In addition to model uncertainties, robust control may also consider other system uncertainties, e.g., disturbances, measurement noises, etc.

as the persistent excitation property on the observed data and additional stochastic properties on w_t. In this chapter, we consider the following assumptions to establish convergence rate bounds on the diameter of Θ_T, which, to the best of our knowledge, is the first non-asymptotic guarantee of SME for linear dynamical systems.

The first assumption formalizes the bounded disturbance assumption discussed above and introduces stochastic properties of w_t for analytical purposes.

Assumption 3 (Bounded i.i.d. disturbances). *The disturbances are box-constrained, $w_t \in \mathcal{W} := \{w \in \mathbb{R}^{n_x} : \|w\|_\infty \leq w_{\max}\}$ for all $t \geq 0$. Further, w_t is i.i.d., has zero mean and positive definite covariance matrix Σ_w.*

Assumption 3 is common in SME literature, e.g. [113], [116], [118]. In terms of generality, boundedness is essential for SME. The stochastic properties, such as i.i.d., zero mean, positive definite covariance, are standard in the recent learning-based control literature and allow the use of statistical tools utilized and developed in the recent literature for non-asymptotic analysis [96], [121]. Besides, it is worth mentioning that SME still works in non-stochastic settings. In particular, as long as $w_t \in \mathcal{W}$, even without the stochastic properties in Assumption 3, the SME algorithm (4.2) still generates a valid uncertainty set that contains $\theta^\star$. It is an interesting future direction to study the convergence rate of SME without assuming stochastic disturbances.

Next, we introduce the assumptions on u_t, which relies on the block-martingale small-ball (BMSB) condition proposed in [96]. It can be shown that the BMSB guarantees persistent excitation (PE) with high probability under proper conditions (see Proposition 2.5 in [96] and Lemma 4). The PE condition requires that z_t explores all directions, which is essential for system identification [122].

Definition 4.2.2 (Persistent excitation). *There exists $\alpha > 0$ and $m \in \mathbb{N}_+$, such that for any $t_0 \geq 0$,*

$$\frac{1}{m} \sum_{t=t_0}^{t_0+m-1} \begin{pmatrix} x_t \\ u_t \end{pmatrix} (x_t^\top, u_t^\top) \succeq \alpha^2 I_{n_x+n_u}.$$

Definition 4.2.3 (BMSB [96]). *Consider a filtration $\{\mathcal{F}_t\}_{t\geq 1}$ and an $\{\mathcal{F}_t\}_{t\geq 1}$-adapted random process $\{Z_t\}_{t\geq 1}$ in $\mathbb{R}^d$. $\{Z_t\}_{t\geq 1}$ satisfies the (k, Γ_{sb}, p)-block martingale small-ball (BMSB) condition for $k > 0$, a positive definite Γ_{sb}, and $0 \leq p \leq 1$, if the following holds: for any fixed $\lambda \in \mathbb{R}^d$ with $\|\lambda\|_2 = 1$, we have $\frac{1}{k} \sum_{i=1}^k \mathbb{P}(|\lambda^\top Z_{t+i}| \geq \sqrt{\lambda^\top \Gamma_{sb} \lambda} \mid \mathcal{F}_t) \geq p$ for all $t \geq 1$.*

The following is the assumption on u_t.

Assumption 4 (BMSB and boundedness). *With filtration $\mathcal{F}_t = \mathcal{F}(w_0, \ldots, w_{t-1}, z_0, \ldots, z_t)$, the $\mathcal{F}_t$-adapted stochastic process $\{z_t\}_{t\geq 0}$ satisfies $(1, \sigma_z^2 I_{n_z}, p_z)$-BMSB for some $\sigma_z, p_z > 0$. Besides, there exists $b_z \geq 0$ such that $\|z_t\|_2 \leq b_z$ almost surely for all $t \geq 0$.*

Assumption 4 requires u_t to guarantee both BMSB and bounded z_t. This can be satisfied by several robust (adaptive) constrained control policies, such as robust (adaptive) model predictive control [49], [107], [118], system level synthesis [24], and control barrier functions [123]. In the following, we briefly discuss robust (adaptive) MPC as an example. The other approaches can be similarly shown to satisfy Assumption 4.

Example 5 (Robust (adaptive) MPC). *Robust MPC is a popular method for the robust constrained control [124], which aims to optimize the control objective while satisfying robust safety constraints,*

$$z_t \in \mathbb{Z}_{\text{safe}}, \text{ where } x_{t+1} = \theta z_t + w_t, \forall \, \theta \in \Theta_0, w_t \in \mathcal{W}, \tag{4.3}$$

where Θ_0 is an initial uncertainty set known a priori, and the safety constraint $\mathbb{Z}_{\text{safe}}$ is usually bounded. The robust MPC policy, denoted by $u_t = \pi_{\text{RMPC}}(x_t; \Theta_0, \mathcal{W})$, satisfies the constraints (4.3) for any $\theta \in \Theta_0$. Therefore, it naturally guarantees bounded z_t under the true $\theta^\star$. Further, as shown in [125], BMSB can be achieved by adding a random disturbance, i.e. $u_t = \pi_{\text{RMPC}}(x_t; \Theta_0, \mathcal{W}) + \eta_t$, where η_t is i.i.d., bounded, and has positive definite covariance. Therefore, the randomly perturbed robust MPC can satisfy Assumption 4. Robust adaptive MPC is based on the same control design, $u_t = \pi_{\text{RMPC}}(x_t; \Theta_t, \mathcal{W})$, but utilizes adaptively updated uncertainty sets Θ_t. Notice that Θ_t is usually updated by SME in the literature of robust adaptive MPC [49], [107], [117].

We also note that BMSB and bounded z_t with high probability are assumed in LSE literature (Theorem 2.4 [96]), and bounded z_t with high probability under subGaussian disturbances corresponds to bounded z_t under bounded disturbances for linear systems (see bounded-input-bounded-output stability in Sec. 9 of [126]).

Finally, we assume that the bound $w_{\max}$ on w_t is tight in all directions, which is common in the literature on SME analysis [112], [113], [118].

Assumption 5 (Tight bound on w_t). *For any $\epsilon > 0$, there exists $q_w(\epsilon) > 0$, such that for any $1 \leq j \leq n$, we have*

$$\min(\mathbb{P}(w_t^j \leq \epsilon - w_{\max}), \mathbb{P}(w_t^j \geq w_{\max} - \epsilon)) \geq q_w(\epsilon),$$

where w_t^j denotes the jth entry of vector w_t. Without loss of generality, we can further assume $q_w(\epsilon)$ to be non-decreasing with ϵ and $q_w(2w_{\max}) = 1$. [2]

In essence, Assumption 5 requires that a *hyper-cubic* $\mathcal{W} = \{w : \|w\|_\infty \leq w_{\max}\}$ should be *tight* on the support of w_t in all coordinate directions, that is, there exists a positive probability $q_w(\epsilon)$ such that w_t visits an ϵ-neighborhood of $w_{\max}$ and $-w_{\max}$, respectively, on all coordinates.

When the support of w_t is indeed $\mathcal{W} = \{w : \|w\|_\infty \leq w_{\max}\}$, many common distributions enjoy $q_w(\epsilon) \geq \Omega(\epsilon)$.[3] For example, for the uniform distribution on $\mathcal{W}$, we have $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}}$; for the truncated Gaussian distribution with zero mean, $\sigma_w^2 I_n$ covariance, and truncated region $\mathcal{W}$, we have $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}\sigma_w} \exp(\frac{-w_{\max}^2}{2\sigma_w^2})$; and for the uniform distribution on the boundary of $\mathcal{W}$ (a generalization of Rademacher distribution), we have $q_w(\epsilon) \geq \frac{1}{2n_x} \geq \Omega(\epsilon)$ (see Section 4.C for more details).

However, knowing a tight bound on the support of w_t can be challenging in practice. Therefore, we will discuss how to relax this assumption and learn a tight bound from data in Section 4.3.

Further, the requirement of a hyper-cubic $\mathcal{W}$ can be restrictive because different entries of disturbances may have different magnitudes, resulting in a hyper-rectangular support that violates Assumption 5. Our follow-up work [127] relaxes this assumption and generalizes the results in this chapter.

4.3 Set Membership Convergence Analysis

Convergence Rate of SME with Known $w_{\max}$

We now present the main result (Theorem 12) of this chapter, which is a non-asymptotic bound on the estimation error of SME given bounded i.i.d. stochastic disturbances.

---

[2]This is because $\mathbb{P}(w_t^j \leq \epsilon - w_{\max})$ and $\mathbb{P}(w_t^j \geq w_{\max} - \epsilon)$ are non-decreasing with ϵ, and $\mathbb{P}(w_t^j \geq -w_{\max}) = \mathbb{P}(w_t^j \leq w_{\max}) = 1$ by Assumption 3.

[3]The $\Omega(\cdot)$ notation is the lower bound version of $O(\cdot)$.

Theorem 12 (Convergence rate of SME). *For any $m > 0$ any $\delta > 0$, when $T > m$, we have*

$$\mathbb{P}(\operatorname{diam}(\Theta_T) > \delta) \leq \underbrace{\frac{T}{m}\widetilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m)}_{\mathbb{T}_1}$$

$$+ \underbrace{\widetilde{O}((n_x n_z)^{2.5})a_4^{n_x n_z}\left(1 - q_w\left(\frac{a_1 \delta}{4\sqrt{n_x}}\right)\right)^{\lceil T/m \rceil}}_{\mathbb{T}_2(\delta)}$$

where $a_1 = \frac{\sigma_z p_z}{4}$, $a_2 = \frac{64 b_z^2}{\sigma_z^2 p_z^2}$, $a_3 = \frac{p_z^2}{8}$, $a_4 = \frac{4 b_z \sqrt{n_x}}{a_1}$, p_z, σ_z, b_z are defined in Assumption 4, $\lceil \cdot \rceil$ denotes the ceiling function, and $\operatorname{diam}(\cdot)$ is defined in Definition 4.2.1, the factors hidden in $\widetilde{O}(\cdot)$ are provided in Appendix 4.D.

Theorem 12 provides an upper bound on the "failure" probability of SME, i.e., the probability that the diameter of the uncertainty set is larger than δ. In this bound, $\mathbb{T}_1$ decays exponentially with m, so for any small $\epsilon > 0$, m can be chosen such that $\mathbb{T}_1 \leq \epsilon$, which indicates $m \geq O(n_z + \log T + \log(1/\epsilon))$. For any $\delta > 0$, $\mathbb{T}_2(\delta)$ decays exponentially with the number of data points T and involves a distribution-dependent function $q_w(\cdot)$, which characterizes how likely it is for w_t to visit the boundary of $\mathcal{W}$ as defined in Assumption 5. To ensure the probability upper bound in Theorem 12 to be less than 1, one can choose $m = O(\log T)$ and a large enough T such that $T \geq O(m) = O(\log(T))$. If w_t is more likely to visit the boundary, (a larger $q_w(\cdot)$), then SME is less likely to generate an uncertainty set with a diameter bigger than δ.

Estimation error bounds when $q_w(\epsilon) = \Omega(\epsilon)$. To provide intuition for $\mathbb{T}_2(\delta)$ and discuss the estimation error bound in Theorem 12 more explicitly, we consider distributions satisfying $q_w(\epsilon) = \Omega(\epsilon)$ for all $\epsilon > 0$. Notice that several common distributions satisfy this additional requirement, such as uniform distribution and truncated Gaussian distribution as discussed after Assumption 5.

Corollary 12.1 (Estimation error bound when $q_w(\epsilon) = \Omega(\epsilon)$). *For any $\epsilon > 0$, let*

$$m \geq O(n_z + \log T + \log(1/\epsilon))$$

in the following. [4] If w_t is generated i.i.d. by a distribution satisfying $q_w(\epsilon) = \Omega(\epsilon)$ for all $\epsilon > 0$, then with probability at least $1 - 2\epsilon$, for any $\widehat{\theta}_T \in \Theta_T$, we have

$$\|\widehat{\theta}_T - \theta^\star\|_F \leq \operatorname{diam}(\Theta_T) \leq \widetilde{O}\left(\frac{n_x^{1.5}(n_x + n_u)^2}{T}\right).$$

---

[4]A detailed formula is provided in Appendix 4.E.

Corollary 12.1 indicates that the estimation error of any point in the uncertainty set Θ_T can be bounded by $\widetilde{O}\left(\frac{n_x^{1.5}(n_x+n_u)^2}{T}\right)$ when $q_w(\epsilon) \geq \Omega(\epsilon)$.

Dynamical systems without control inputs. SME also applies to dynamical systems with no control inputs, i.e., $x_{t+1} = A^\star x_t + w_t$, where the uncertainty set of $A^\star$ can be computed by $\mathbb{A}_T = \bigcap_{t=0}^{T-1}\{\widehat{A} : \|x_{t+1} - \widehat{A}x_t\|_\infty \leq w_{\max}\}$. Its convergence rate can be similarly derived via the proof of Theorem 12.

Corollary 12.2 (Convergence rate with $B^\star = 0$ (informal)). *For stable $A^\star$, for any $m > 0, \delta > 0, T > m$, we have*

$$\mathbb{P}(\operatorname{diam}(\mathbb{A}_T) > \delta) \leq \frac{T}{m}\widetilde{O}(n_x^{2.5})a_2^{n_x}\exp(-a_3 m)$$
$$+ \widetilde{O}(n_x^5)a_4^{n_x^2}(1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{\lceil T/m\rceil}$$

Consequently, when $q_w(\epsilon) = \Omega(\epsilon)$, e.g. uniform or truncated Gaussian, we have $\operatorname{diam}(\mathbb{A}_T) \leq \widetilde{O}(n_x^{3.5}/T)$.

Note that [96] have shown a lower bound $\Omega(\sqrt{n_x}/\sqrt{T})$ for the estimation of linear systems with no control inputs when w_t follows an (unbounded) Gaussian distribution. Interestingly, Corollary 12.2 reveals that, for some bounded-support distributions of w_t, e.g. Uniform and truncated Gaussian, SME is able to converge at a faster rate $\widetilde{O}(1/T)$ in terms of the sample size T. This does not conflict with the lower bound in [96] because SME's rate only holds for bounded disturbances. In fact, from (4.2), it is straightforward to see that SME does not even converge under Gaussian disturbances. Therefore, SME is mostly useful in applications with bounded disturbances, e.g. robust constrained control, safety-critical systems, etc., while LSE's confidence regions are preferred for unbounded disturbances.

Lastly, Corollary 12.2 shows that SME's convergence rate has a poor dependence with respect to n_x: $\widetilde{O}(n_x^{3.5})$. This is likely a proof artifact because we do not observe such poor dimension scaling in simulation (see Figure 4.3). It is left as future work to refine the dimension dependence.

SME with Unknown $w_{\max}$

Next, we discuss the convergence rates of SME without knowing a tight bound $w_{\max}$ in three steps: 1) only knowing a conservative upper bound of $w_{\max}$, 2) learning $w_{\max}$ from data, and 3) a variant of SME that converges without prior knowledge of w_{max}.

1) SME with a conservative upper bound for $w_{\max}$**.** In many practical scenarios, it is easier to obtain an over-estimation of the range of the disturbances instead of a tight upper bound, i.e., $\widehat{w}_{\max} \geq w_{\max}$. In this case, we can show that the uncertainty set converges to a small neighborhood around $\theta^\star$ of size $O(\sqrt{n_x}(\widehat{w}_{\max} - w_{\max}))$ at the same convergence rate as Theorem 12.

Theorem 13 (Convservative bound on $w_{\max}$). *When $w_{\max}$ in Assumption 5 is unknown but an upper bound $\widehat{w}_{\max} \geq w_{\max}$ is known, consider the following SME algorithm:*

$$\widehat{\Theta}_T(\widehat{w}_{\max}) = \bigcap_{t=0}^{T-1} \{\widehat{\theta} : \|x_{t+1} - \widehat{\theta} z_t\|_\infty \leq \widehat{w}_{\max}\}.$$

For any $m > 0$, $\delta > 0$, $T > m$, we have

$$\mathbb{P}(\operatorname{diam}(\widehat{\Theta}_T) > \delta + a_5 \sqrt{n_x}(\widehat{w}_{\max} - w_{\max})) \leq \mathbb{T}_1 + \mathbb{T}_2(\delta),$$

where $a_5 = \frac{4}{a_1}$, $\mathbb{T}_1, \mathbb{T}_2(\delta)$ are defined in Theorem 12.

2) Learning $w_{\max}$**.** When $w_{\max}$ is not accurately known, we can try to learn it from the data. Let's first consider the learning algorithm studied in [112].

$$\bar{w}_{\max}^{(T)} = \min_\theta \max_{0 \leq t \leq T-1} \|x_{t+1} - \theta z_t\|_\infty. \tag{4.4}$$

Though algorithm (4.4) cannot provide an upper bound on $w_{\max}$ under finite samples because $\bar{w}_{\max}^{(T)} \leq w_{\max}$ for finite T,[5] it can be shown that $\bar{w}_{\max}^{(T)}$ converges to $w_{\max}$ as $T \to +\infty$. The convergence for linear regression has been established in [112]. The following theorem establishes the convergence and convergence rate of algorithm (4.4) for linear dynamical systems. Based on this convergence rate, we will design an online learning algorithm (4.5) that generates converging upper bounds of $w_{\max}$.

Theorem 14. *The estimation $\bar{w}_{\max}^{(T)}$ of $w_{\max}$ satisfies:*

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \underbrace{b_z \operatorname{diam}(\Theta_T)}_{\mathbb{T}_3} + \underbrace{w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty}_{\mathbb{T}_4}$$

Therefore, for any $\delta > 0$,

$$\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) \leq \mathbb{T}_1 + \mathbb{T}_2\left(\frac{\delta}{2b_z}\right) + \mathbb{T}_5\left(\frac{\delta}{2}\right),$$

where $\mathbb{T}_5(\delta) = (1 - q_w(\delta))^T$.

---

[5]If SME does not use an upper bound on $w_{\max}$, the generated uncertainty set may not contain the true parameter $\theta^\star$.

Notice that $\mathbb{T}_4$ is the smallest possible learning error of $w_{\max}$ from history w_t, which can be achieved if one can directly measure w_t. However, with unknown $\theta^\star$, it is challenging to measure/compute w_t exactly, then Theorem 14 shows that the learning error of $w_{\max}$ has an additional term $\mathbb{T}_3$ that depends on the uncertainty around $\theta^\star$. Therefore, the convergence rate of $\bar{w}_{\max}^{(T)}$ can be obtained by our non-asymptotic analysis of SME in Theorem 12.

Further, when $q_w(\epsilon) = \Omega(\epsilon)$, the convergence rate of $\bar{w}_{\max}^{(T)}$ can be explicitly bounded by $\widetilde{O}(n_x^{1.5} n_z^2 / T)$, which is of the same order as the convergence rate of the diameter of Θ_T.

Corollary 14.1. *For any $0 < \epsilon < 1/3$ and any $T \geq 1$, there exists $\delta_T > 0$ satisfying* $\lim_{T \to \infty} \delta_T = 0$ *such that*

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \delta_T$$

with probability at least $1 - 3\epsilon$.

In particular, when $q_w(\delta) = O(\delta)$, with probability $1 - 3\epsilon$,

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \delta_T = \widetilde{O}(n_x^{1.5} n_z^2 / T)$$

3) SME with unknown $w_{\max}$. Unfortunately, $\bar{w}_{\max}^{(T)}$ cannot be directly applied to SME because $\bar{w}_{\max}^{(T)} \leq w_{\max}$, which may cause $\theta^\star \notin \widehat{\Theta}_T(\bar{w}_{\max}^{(T)})$. However, by leveraging our convergence rate bound in Theorem 14, we can construct an upper confidence bound (UCB) of $w_{\max}$ and a corresponding UCB-SME algorithm:

$$\widehat{w}_{\max}^{(T)} = \bar{w}_{\max}^{(T)} + \delta_T, \quad \widehat{\Theta}_T^{\text{ucb}} = \widehat{\Theta}_T(\widehat{w}_{\max}^{(T)}), \tag{4.5}$$

where δ_T is defined in Corollary 14.1.

Then, by combining Theorem 13 and Corollary 14.1, we can verify the well-definedness of UCB-SME and obtain its convergence rate.

Theorem 15. *For any $0 < \epsilon < 1/3$, any $T \geq 1$, with probability at least $1 - 3\epsilon$, we have*

$$\theta^\star \in \widehat{\Theta}_T^{\text{ucb}}, \quad \text{diam}(\widehat{\Theta}_T^{\text{ucb}}) \leq O(\sqrt{n_x} \delta_T).$$

In particular, if $q_w(\epsilon) = \Omega(\epsilon)$, then $\text{diam}(\widehat{\Theta}_T^{\text{ucb}}) \leq O(n_x^2 n_z^2 / T)$ with probability at least $1 - 3\epsilon$.

Notice that UCB-SME converges at the same rate in terms of T but $\sqrt{n_x}$-worse in terms of dimensionality when compared with SME knowing a tight bound $w_{\max}$.

Remark 2 (Computation complexity). *SME can be computed by linear programming since all constraints are linear in (4.2). Further, UCB-SME can also be computed by linear programming because (4.4) can be reformulated as a linear program. However, the number of constraints for SME and UCB-SME increases linearly with T. To address the computation issue of SME, many computationally efficient algorithms have been proposed based on approximations of (4.2), e.g. [55], [115], [118]. The convergence rates of these approximate algorithms are unknown and how to design computationally efficient UCB-SME remains open.*

4.4 Proof Sketch

The major technical novelty of this chapter is the proof of Theorem 12, thus we describe the key ideas here. The complete proof is provided in Appendix 4.D. For ease of notation and without loss of generality, we assume T/m is an integer in the following.

Specifically, we first define a set Γ_T on the model estimation error $\gamma = \widehat{\theta} - \theta^\star$ by leveraging the observation that $x_{s+1} - \widehat{\theta}z_s = w_s - (\widehat{\theta} - \theta^\star)z_s$,

$$\Gamma_t = \bigcap_{s=0}^{t-1}\{\gamma : \|w_s - \gamma z_s\|_\infty \leq w_{\max}\}, \quad \forall\, t \geq 0. \tag{4.6}$$

Notice that $\Theta_t = \theta_\star + \Gamma_t$, so $\mathrm{diam}(\Theta_t) = \mathrm{diam}(\Gamma_t)$, and

$$\mathrm{diam}(\Gamma_t) = \sup_{\gamma,\gamma'\in\Gamma_t} \|\gamma - \gamma'\|_F \leq 2\sup_{\gamma\in\Gamma_t} \|\gamma\|_F.$$

Thus, we can define $\mathcal{E}_1 := \{\exists\, \gamma \in \Gamma_T, \text{ s.t. } \|\gamma\|_F \geq \frac{\delta}{2}\}$ such that $\mathbb{P}(\mathrm{diam}(\Theta_T) > \delta) \leq \mathbb{P}(\mathcal{E}_1)$.

Next, we define an event $\mathcal{E}_2$ below, which is essentially PE on every time segments $km + 1 \leq t \leq km + m$ for $k \geq 0$, where the choice of m will be specified later.

$$\mathcal{E}_2 = \left\{\frac{1}{m}\sum_{s=1}^{m} z_{km+s}z_{km+s}^\top \geq a_1^2 I_{n_z}, \forall\, 0 \leq k \leq \left\lceil\frac{T}{m}\right\rceil - 1\right\},$$

where $a_1 = \frac{\sigma_z p_z}{4}$. Now, by dividing the event $\mathcal{E}_1$ based on $\mathcal{E}_2$, we obtain

$$\mathbb{P}(\mathrm{diam}(\Theta_T) > \delta) \leq \mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2^c) + \mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2).$$

The proof can be completed by establishing the following bounds on $\mathbb{P}(\mathcal{E}_2^c)$ and $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2)$.

(a) Truncated Gaussian   (b) Uniform Distribution   (c) $\widehat{w}_{\max}^{(T)}$ and $\bar{w}_{\max}^{(T)}$ in UCB-SME
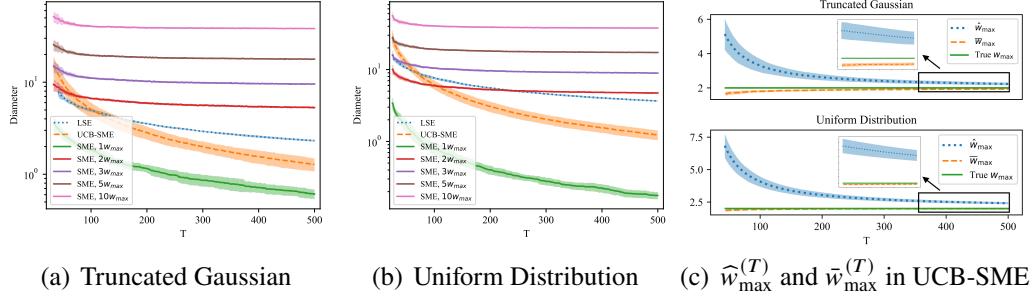
Figure 4.2: **Figures (a)-(b)** compares the diameters of SME, UCB-SME, and SME with loose disturbance upper bounds that are 2, 3, 5, and 10 times larger than the true disturbance bound $w_{\max}$, as well as the baseline uncertainty set from the 90% confidence region of LSE. **Figure (c)** shows the convergence to the true bound $w_{\max}$ of the lower estimation $\bar{w}_{\max}$ in (4.4) and the UCB $\widehat{w}_{\max}$ generated by the UCB-SME algorithm in **Figures (a)-(b)**.

Lemma 4 (Bound on $\mathbb{P}(\mathcal{E}_2^c)$). $\mathbb{P}(\mathcal{E}_2^c) \leq \mathbb{T}_1$, where $a_2 = \frac{64b_z^2}{\sigma_z^2 p_z^2}$ and $a_3 = \frac{p_z^2}{8}$.

Lemma 5 (Bound on $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2)$). $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{T}_2(\delta)$, where $a_4 = \max(1, 4b_z\sqrt{n_x}/a_1)$.

Roughly, Lemma 4 indicates that PE holds with high probability, which is proved by leveraging the BMSB assumption and set discretization. The proof of Lemma 5 is more involved and is our major technical contribution. On a high level, the proof relies on two technical lemmas below.

Lemma 6 (Discretization of $\mathcal{E}_1 \cap \mathcal{E}_2$ (informal)). *Let $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\}$ denote an ϵ_γ-net of $\{\gamma : \|\gamma\|_F = 1\}$. Under a proper choice of ϵ_γ, we have $v_\gamma = \widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}$.[6] We can construct $\widetilde{\Gamma}_T$ such that*

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\{\exists\, 1 \leq i \leq v_\gamma, d \geq 0, \text{ s.t. } d\gamma_i \in \widetilde{\Gamma}_T\} \cap \mathcal{E}_2)$$

$$\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2)$$

where $\mathcal{E}_{1,i} = \{\exists\, d \geq 0, \text{ s.t. } d\gamma_i \in \widetilde{\Gamma}_T\}$.

Lemma 6 leverages finite set discretization to bound the existence of a feasible element in an infinite continuous set. The formal version of Lemma 6 is provided as Lemma 15 in the appendix.

---

[6]The exact formulas of v_γ and ϵ_γ are in Lemma 11.

Lemma 7 (Construction of event $G_{i,k}$ via stopping times (informal)). *Consider $\mathcal{F}_t$ as defined in Assumption 4. Under the conditions in Lemma 6, we construct $G_{i,k}$ for all i and all $0 \leq k \leq T/m - 1$ by*

$$G_{i,k} = \left\{ b_{i,km+L_{i,k}} w^{j_{i,km+L_{i,k}}}_{km+L_{i,k}} \geq \frac{a_1 \delta}{4\sqrt{n_x}} - w_{\max}, \text{ and} \right.$$

$$\left. \frac{1}{m} \sum_{s=1}^{m} z_{km+s} z_{km+s}^{\top} \geq a_1^2 I_{n_z} \right\},$$

where $b_{i,t}, j_{i,t}$ are measurable in $\mathcal{F}_t$, and $L_{i,k}$ is constructed as a stopping time with respect to $\{\mathcal{F}_{km+l}\}_{l \geq 0}$. The formal definitions of $b_{i,t}, j_{i,t}, L_{i,k}$ are provided in Appendix 4.D.

Then, we have

$$\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}\left(\bigcap_{k=0}^{T/m-1} G_{i,k} \right) \leq \left(1 - q_w \left(\frac{a_1 \delta}{4\sqrt{n_x}} \right) \right)^{\frac{T}{m}}$$

The constructions of $G_{i,k}$ and $L_{i,k}$ in Lemma 7 are our major technical contribution. With the constructions above, the proof can be completed by leveraging the conditional independence property of stopping times, which is briefly discussed below. Notice that by conditioning on the event $\{L_{i,k} = l\}$, we have $w_{km+L_{i,k}} = w_{km+l}$ and w_{km+l} is independent of $\mathcal{F}_{km+l}$. Consequently, w_{km+l} is also independent of $b_{i,km+L_{i,k}}, j_{i,km+L_{i,k}}$ conditioning on $\{L_{i,k} = l\}$ since $b_{i,km+l}, j_{i,km+l}$ are measurable in $\mathcal{F}_{km+l}$. Therefore, the probability of $G_{i,k}$ conditioning on $\{L_{i,k} = l\}$ can be bounded by the probability distribution of w_t, which enjoys good properties such as Assumption 5. More details of the proof are in Appendix 4.D.

In conclusion, Lemma 5 follows directly from Lemma 6 and Lemma 7. Combining Lemma 5 and Lemma 4 completes the proof of Theorem 12.

Remark 3 (Convergence rate of SME for general time series). *Similar to Theorem 2.4 in [96], our results for linear dynamical systems can also be generalized to general time series with linear responses:*

$$y_t = \theta^{\star} z_t + w_t, \quad t \geq 0,$$

where $\mathcal{F}_t^y = \mathcal{F}(w_0, \ldots, w_t, z_0, \ldots, z_t)$, $y_t \in \mathbb{R}^{n_y}$ is measurable in $\mathcal{F}_t^y$ but not in $\mathcal{F}_{t-1}^y$. The SME algorithm is

$$\Theta_T^y = \bigcap_{t=0}^{T-1} \{\widehat{\theta} : y_t - \widehat{\theta} z_t \in \mathcal{W}\}.$$

Under Assumptions 3, 4, and 5, we have

$$\mathbb{P}(\mathrm{diam}(\Theta_T^y) > \delta) \leq \frac{T}{m}\widetilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m)$$

$$+ \widetilde{O}((n_y n_z)^{2.5})a_4^{n_y n_z}\left(1 - q_w\left(\frac{a_1\delta}{4\sqrt{n_y}}\right)\right)^{\lceil T/m \rceil},$$

where a_1, a_2, a_3 are defined in Theorem 12 and $a_4 = \frac{4b_z\sqrt{n_y}}{a_1}$.

4.5   Applications to Robust Adaptive Control

Robust adaptive control usually involves two steps: updating the uncertainty set estimation, and designing robust controllers based on the updated uncertainty set. SME can be naturally applied to robust adaptive control as the updating rule of the uncertainty set estimation. To illustrate this, we discuss the applications of SME to two popular controllers, robust adaptive MPC and robust SLS. We focus on the implications of our convergence rates.

Application of SME to robust adaptive MPC. SME has long been adopted in the robust adaptive MPC design (see e.g., [49], [107], [117]). Despite the regret analysis for unconstrained MPC and its variants (e.g. [128], [129]), the non-asymptotic analysis for robust adaptive MPC remains unsolved. Applying Theorem 12 straightforwardly, we can obtain a non-asymptotic estimation error bound for robust adaptive MPC below, which lays a foundation for future regret analysis. For simplicity, we consider a tight bound $\mathcal{W}$ is known below, but our results for unknown $\mathcal{W}$ can also be applied similarly.

Corollary 15.1. *Consider the robust adaptive MPC controller introduced in Example 5, where Θ_t is updated by SME and $\mathcal{W}$ is known.[7] Under the conditions of Corollary 12.1, the estimation error for any $\widehat{\theta}_T \in \Theta_T$ can be bounded by $\|\widehat{\theta}_T - \theta^\star\|_F \leq \widetilde{O}(\frac{n_x^{1.5}n_z^2}{T})$ with high probability.*

Application of SME to robust SLS. Robust SLS has been proposed in [24] for robust constrained control under system uncertainties [24]. Since [24] assumes bounded disturbances, one can apply SME for the uncertainty set estimation in place of the LSE's confidence regions in [24]. Then, by leveraging Theorems 3.1, 4.1 in [24] and our Theorem 12, we can directly obtain a non-asymptotic suboptimality gap for learning-based robust SLS with SME as the uncertainty set estimation. For

---

[7]When $\mathcal{W}$ is unknown, Theorems 13-15 all apply.

simplicity, we consider a known tight bound $\mathcal{W}$, but our results for unknown $\mathcal{W}$ can also be similarly applied here.

Corollary 15.2. *Under the conditions in Theorem 3.1 in [24] and Corollary 12.1, for large enough T, we have $\frac{J(A^\star, B^\star, \widehat{\mathbf{K}}) - J^\star}{J^\star} \leq \widetilde{O}(n_x^{1.5} n_z^2 / T)$, where $\widehat{\mathbf{K}}$ denotes the robust SLS controller in [24] under the uncertainty set Θ_T constructed by SME, $J(A^\star, B^\star, \widehat{\mathbf{K}}) = \lim_{T \to +\infty} \mathbb{E} \frac{1}{T} \sum_{t=0}^{T-1} (x_t^\top Q x_t + u_t^\top R u_t)$ denotes the infinite-horizon averaged total cost by implementing the robust SLS controller $\widehat{\mathbf{K}}$, and $J^\star$ denotes the optimal infinite-horizon averaged total cost.*

4.6 Simulation

We evaluate the empirical performance of SME on various systems and applications. For all experiments, we use the 90% confidence regions of LSE computed by Lemma E.3 in [59] and Theorem 1 in [60] as the baseline. The details of the simulation settings are provided in Appendix 4.I.[8]

Comparison of SME, SME with loose bound, UCB-SME, and LSE. This experiment is based on the linearized longitudinal flight control dynamics of Boeing 747 as studied in recent literature on learning-based control of linear systems [17], [130].

In Figure 4.2, we show the diameters of SME, SME with loose disturbance bounds, and UCB-SME on the identification problem of the Boeing 747 dynamics with i.i.d. truncated Gaussian (Figure 4.2(a)) and uniform (Figure 4.2(b)) disturbances. We use control actions sampled from a uniform distribution in both cases. In Figure 4.2(c), we show that both the upper bound $\widehat{w}_{\max}$ used for UCB-SME and the lower bound $\bar{w}_{\max}$ in (4.4) converge to the true bound $w_{\max}$ as T increases. The quantitative behaviors of SME and its variants are consistent with those predicted by our theoretical results. In particular, in Figure 4.2(a) and Figure 4.2(b), SME and UCB-SME outperform the 90% confidence regions of LSE in both the magnitude and the convergence rate. In Figure 4.2(c), we verify that the UCB estimation $\widehat{w}_{\max}^{(T)}$ converges to the true disturbance bound $w_{\max}$ from above, while the estimation $\bar{w}_{\max}^{(T)}$ converges from below. It is worth noting that $\bar{w}_{\max}^{(T)}$ converges to $w_{\max}$ very quickly in the simulations, allowing $\bar{w}_{\max}^{(T)}$ to be another potential approximation of $w_{\max}$ for SME when T is very large.

---

[8]The code to reproduce all the experimental results can be found at `https://github.com/jy-cds/non-asymptotic-set-membership`.
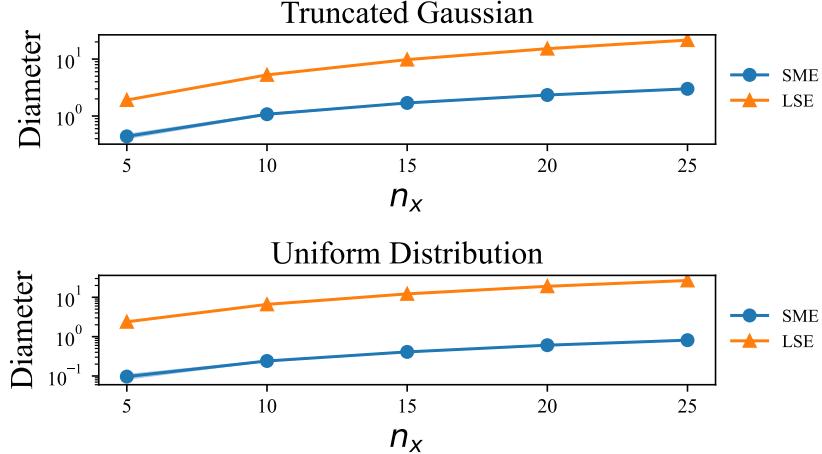
Figure 4.3: Diameters of the uncertainty sets constructed by SME, UCB-SME, and LSE for systems with different dimensions.

Scaling with dimension. We compare the scaling of SME, SME-UCB, and LSE with respect to the system dimensions in Figure 4.3. We use an autonomous system $x_{t+1} = A^\star x_t + w_t$, where $A \in \mathbb{R}^{n_x \times n_x}$ has varying n_x. Disturbances w_t are sampled from a truncated Gaussian distribution and uniform distribution with $w_{\max} = 2$. Surprisingly, the scaling of SME with respect to the dimension of the system is not significantly worse than that of LSE in the simulation. This suggests that the convergence rate in Corollary 12.1 can potentially be improved in terms of the dimension dependence, which is left for future investigation.

Figure 4.4: Linear quadratic tracking of robust adaptive MPC based on SME, LSE's confidence regions, and the accurate model (OPT).

Application to robust adaptive MPC. We provide an example of the quantitative

impact of using SME for adaptive robust MPC in Figure 4.4. We consider the task of constrained linear quadratic tracking problem as in [125]. The model uncertainty set is estimated online with SME and LSE's 90 % confidence region. Control actions are computed usin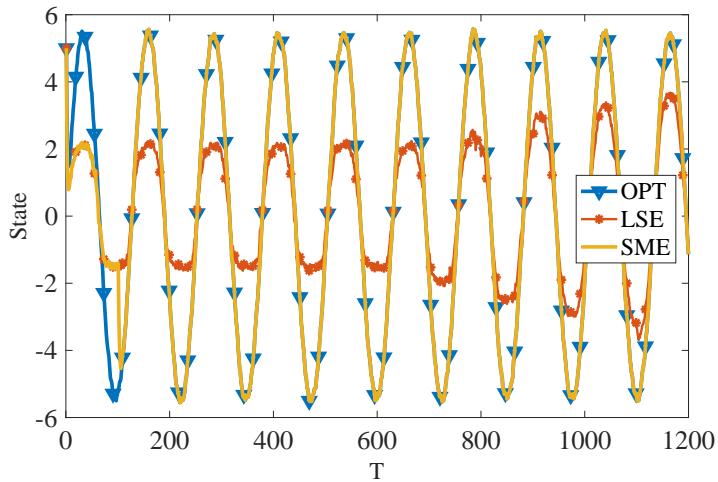g the tube-based robust MPC [131], [132] with the uncertainty sets. We also plot the optimal MPC controller with accurate model information. Thanks to the fast convergence of SME, the tracking performance of the tube-based robust MPC with SME estimation quickly coincides with OPT, while the same controller based on LSE's confidence region estimation converges more slowly.

4.7   Conclusion

This work provides the first convergence rates for SME in linear dynamical systems with bounded disturbances and discusses variants of SME with unknown bound on w_t. Numerical experiments demonstrate SME's promising performance under bounded disturbances.

Regarding future directions, this work only considers box constraints on w_t, so it is worth extending the analysis to more general constraints. In this work, we only measure the size of the uncertainty sets by their diameters. We leave for future work to consider other metrics, such as volume. Further, our bounds suffer poor dependence on the system dimension, which is not reflected in simulations. Hence, it is important to refine the bounds and discuss the fundamental limits. Another exciting direction is to speed up the computation of SME since the current computation complexity increases linearly with the sample size. The convergence rate of such algorithms is an important *open question*. Other interesting directions include the extensions of the SME analysis to nonlinear systems, where recent nonlinear system identification literature [133], [134] may provide insights; and analyzing SME in the presence of other uncertainties, e.g. measurement noises [135].

SME is a valid estimation for bounded non-stochastic disturbances [48], [100], [101], [136]. Thus, a fruitful direction is to study SME's convergence rates under non-stochastic w_t. Another method for uncertainty set estimation is the credible regions of Bayesian approaches, e.g. Thompson sampling for linear systems [28], [137] and Gaussian processes for nonlinear systems [138]. A future direction is to study the convergence rates of credible regions.

Roadmap for the Appendix

- Section 4.A introduces additional notation used throughout the Appendix.

- Section 4.B provides more literature review on LSE and SM, and a more detailed discussion on the technical contributions of this chapter.

- Section 4.C provides more discussions on examples that satisfy Assumptions 4 and 5.

- Section 4.D presents the proof of Theorem 12. In particular, we provide helper lemmas in Section 4.D and prove Lemma 4, Lemma 5 in Section 4.D and Section 4.D respectively. A more precise upper bound for Theorem 12 (without the $\widetilde{O}(\cdot)$ notation) is provided in Appendix 4.D.

- Section 4.E presents a proof of Corollary 12.1

- Section 4.F provides a proof of Corollary 12.2.

- Section 4.G presents a proof of Theorem 13.

- Appendix 4.H provides proofs of Theorem 14, Corollary 14.1, and Theorem 15.

- Section 4.I provides details of the simulation.

4.A Additional Notations

Let $\mathbb{S}_n(0, 1)$ denote the unit sphere in $\mathbb{R}^n$ in l_2 norm, i.e., $\mathbb{S}_n(0, 1) = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. Let $\mathbb{S}_{n \times m}(0, 1)$ denote the unit sphere in $\mathbb{R}^{n \times m}$ with respect to the Frobenius norm, i.e., $\mathbb{S}_{n \times m}(0, 1) = \{M \in \mathbb{R}^{n \times m} : \|M\|_F = 1\}$. Let $\bar{B}_n(0, 1)$ denote the closed unit ball in $\mathbb{R}^n$ in l_2 norm, i.e., $\bar{B}_n(0, 1) = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$. Let $\bar{B}_{n \times m}(0, 1)$ denote the closed unit ball in $\mathbb{R}^{n \times m}$ in Frobenius norm, i.e., $\bar{B}_{n \times m}(0, 1) = \{M \in \mathbb{R}^{n \times m} : \|M\|_F \leq 1\}$. For a matrix $M \in \mathbb{R}^{n \times m}$, $\text{vec}(M)$ is the vectorization of M. Moreover, we define the inverse mapping of $\text{vec}(\cdot)$ as $\text{mat}(\cdot)$, i.e., for a vector $d \in \mathbb{R}^{nm}$, $\text{mat}(d) \in \mathbb{R}^{n \times m}$. Consider a σ-algebra $\mathcal{F}$ and a random variable X, we write $X \in \mathcal{F}$ if X is measurable with respect to $\mathcal{F}$, i.e., for all Borel measurable sets $B \subseteq \mathbb{R}$, we have $X^{-1}(B) \in \mathcal{F}$. We can similarly define $\mathcal{F}$-measurable random matrices and random vectors. Further, consider a polyhedral $\mathbb{D} = \{x : Ax \leq b\}$, we write $\mathbb{D} \in \mathcal{F}$ if matrix A and vector b are measurable with respect to $\mathcal{F}$. Consider two symmetric matrices $A, B \in \mathbb{R}^{n \times n}$, we write $A \succeq B$ if $A - B$ is a positive definite matrix. We define $\min \emptyset = +\infty$. For a set $\mathcal{E}$, let $\mathbb{1}_{\mathcal{E}}$ denote the indicator function on $\mathcal{E}$. For a vector $x \in \mathbb{R}^n$, we use x^j to denote the jth coordinate of x. Throughout the chapter, we use $\text{TrunGauss}(0, \sigma_w, [-w_{\max}, w_{\max}])$ to refer to the truncated Gaussian distribution generated by Gaussian distribution with zero mean and σ_w^2 variance

with truncated range $[-w_{\max}, w_{\max}]$. The same applies to multi-variate truncated Gaussian distributions.

4.B   Dicussion on Least Squares and SME

System identification studies the problem of estimating the parameters of an unknown dynamical systems from trajectory data. There are two main classes of estimation methods: point estimator such as least squares estimation (LSE), and set estimator such as set membership estimation (SME). In the following, we provide more discussions and literature review on LSE and SME. We will also discuss the major technical novelties of this work.

Least Squares Estimation

For linear dynamical systems $x_{t+1} = A^\star x_t + B^\star u_t + w_t = \theta^\star z_t + w_t$, given a trajectory of data $\{x_t, u_t\}_{t\geq 0}$, least squares estimation generates a point estimator that minimizes the following quadratic error [139], [140]:

$$\widehat{\theta}_{\text{LSE}} = \min_{\widehat{\theta}} \sum_{t=0}^{T-1} \|x_{t+1} - \widehat{\theta} z_t\|_2^2.$$

Least-square estimation is widely used and its convergence (rate) guarantees have been investigated for a long time. In particular, non-asymptotic convergence rate guarantees of LSE has become increasingly important as these guarantees are the foundations for non-asymptotic performance analysis of learning-based/adaptive control algorithms. Earlier non-asymptotic analysis of LSE focused on the simpler regression model $y_t = \theta^\star x_t + w_t$, where x_t and y_t are independent [141]–[143].

Recently, there is one major breakthrough in [96] that provides LSE's convergence rate analysis for linear dynamical system $x_{t+1} = \theta^\star z_t + w_t$, where x_{t+1} and $z_t = [x_t^\top, u_t^\top]^\top$ are correlated. More specifically, [96] establishes a fundamental property, *block-martingale small-ball (BMSB)*, to analyze LSE under correlated data. BMSB enables a long list of subsequent literature on LSE's non-asymptotic analysis for different types of dynamical systems, e.g., [19], [95], [109], [144]–[149].

Though LSE is a point estimator, one can establish confidence region of LSE based on proper statistical assumptions on w_t. The pioneer works on the confidence region of LSE for linear dynamical systems are [27], [60], which construct ellipsoid confidence regions for LSE. Moreover, the non-asymptotic bounds on estimation errors established in [24], [96] can also be viewed as confidence bounds. Further, the estimation error $\widetilde{O}(\frac{\sqrt{n_x+n_z}}{\sqrt{T}})$ has been shown to match the fundamental lower

bound for any estimation methods for unbounded disturbances in [96]. However, these confidence bounds all rely on statistical inequalities, which may result in loose constant factors despite an optimal convergence rate. When applying these confidence bounds to robust control, where the controller is required to satisfy certain stability and constraint satisfaction properties for every possible system in the confidence region, a loose constant factor will result in a larger confidence region and a more conservative control design. Finally, in robust control and many practical applications, the disturbances are usually bounded, and it will be interesting to see how the knowledge of the boundedness will improve the uncertainty set estimation.

On a side note, this chapter is also related with the ambiguity set estimation for the transition probabilities in robust Markov decision processes [31]. There are attempts on improving the ambiguity set estimation based on LSE for less conservative robust MDP [31].

Set Membership Estimation

Set membership is commonly used in robust control for uncertainty set estimation [50]–[54], [150]–[152]. There is a long history of research on SME for both deterministic disturbances, such as [48], [100], [101], [115], [136], [153], and stochastic disturbances, such as [112]–[115], [118]. For the stochastic disturbances, both convergence and convergence rate analysis have been investigated under the persistent excitation (PE) condition. However, the existing convergence rates are only established for simpler regression problems, $y_t = \theta^\star x_t + w_t$, where y_t and x_t are independent [112]–[115].

Recently, [118] provided an initial attempt to establish the convergence guarantee of SME for linear dynamical systems $x_{t+1} = \theta^\star z_t + w_t$ for correlated data x_{t+1} and z_t. However, [118] assumes that PE holds deterministically, and designs a special control design based on constrained optimization to satisfy PE deterministically. Therefore, the convergence for general control design and the convergence rate analysis remain open questions for correlated data arising from dynamical systems.

In this work, we establish the convergence rate guarantees of SME on linear dynamical systems under the BMSB conditions in [96]. Compared with [118], BMSB condition can be satisfied by adding an i.i.d. random noise to a general class of control designs [109].

Technically, one major challenge of SME analysis compared with the LSE analysis is that the diameter of the membership set does not have an explicit formula,

which is in stark contrast with LSE, where the point estimator is the solution to a quadratic program and has explicit form. A common trick to address this issue in the analysis of SME is to connect the diameter bound with the values of disturbances subsequences $\{w_{s_k}\}_{k\geq 0}$: it can be generally shown that a large diameter indicates that a long subsequence of disturbances are far away from the boundary of $\mathcal{W}$. However, existing construction methods of $\{w_{s_k}\}_{k\geq 0}$ will cause the time indices $\{s_k\}_{k\geq 0}$ to *correlate* with the realization of the sequences $\{x_t, u_t, w_t\}_{t\geq 0}$ [113], [115], [118].[9] Consequently, in the correlated-data scenario and when PE does not hold deterministically, under the existing construction methods in [113], [115], [118], the probability of $\{w_{s_k}\}_{k\geq 0}$ with correlated time indices *cannot* be bounded by the probability of the *independent* sequence $\{w_t\}_{t\geq 0}$. One major **technical contribution** of this chapter is to provide a novel construction of $\{w_{s_k}\}_{k\geq 0}$ based on a sequence of stopping times and establish conditional independence properties despite correlated data and stochastic PE condition (BMSB). More details can be found in Lemma 7 and the proof or Lemma 5.

Though we only consider box constraints for w_t, it is worth mentioning that SME can be applied to much more general forms of disturbances. For example, a common alternative is the ellipsoidal-bounded disturbance where $\mathcal{W} := \{w \in \mathcal{R}^{n_x} : w^\top P w \leq 1\}$ with positive definite $P \in \mathcal{R}^{n_x \times n_x}$ [115], [154]–[156] and polytopic-bounded disturbance $\mathcal{W} := \{w \in \mathcal{R}^{n_x} : Gw \leq h\}$ for positive definite $G \in \mathcal{R}^{n_x \times n_x}$ and $h \in \mathcal{R}^{n_x}$ [48], [107], [118]. There are also SME literature assuming bounded energy of the disturbance sequences [115]. It is an interesting future direction to extend the analysis in this chapter to more general disturbance constraints.

Further, exact SME involves the intersection of an increasing number of sets, thus causing the computation complexity increases with time t, which can become prohibitive when t is large. There are many methods trying to reduce the computation complexity by approximating the membership sets (see e.g., [101], [118], etc.). It is an exciting future direction to study the diameter bounds of the approximated SME methods.

Lastly, it is worthing mentioning that SME can also be applied to the uncertainty set estimation in perception-based control [157], [158].

---

[9]In [118], the correlation between $\{s_k\}_{k\geq 0}$ and $\{x_t, u_t, w_t\}_{t\geq 0}$ is via the PE condition, but [118] assume deterministic PE to avoid this correlation issue.

4.C   Discussion on Assumptions

Assumption 4

The BMSB condition has been widely used in learning-based control. It has been shown that BMSB can be satisfied in many scenarios. For example, [96], [120] showed that linear systems with i.i.d. perturbed linear control policies, i.e., $x_{t+1} = Ax_t + B(Kx_t + \eta_t) + w_t$,[10] satisfy BMSB if the disturbances w_t and η_t are i.i.d. and follow Gaussian distributions with positive definite covariance matrices. Later, [24] showed that $x_{t+1} = Ax_t + B(Kx_t + \eta_t) + w_t$ can still satisfy BMSB even for non-Gaussian distributions of w_t, η_t, as long as w_t and η_t have independent coordinates and finite fourth moments. Recently, [109] extended the results to linear systems with nonlinear policies, i.e., $x_{t+1} = Ax_t + B(\pi_t(x_t) + \eta_t) + w_t$, and showed that BMSB still holds as long as the nonlinear policies π_t generate bounded trajectories of states and control inputs, and w_t, η_t are bounded and follow distributions with certain anti-concentrated properties (a special case is positive definite covariance matrix).

Assumption 5

In this subsection, we provide two example distributions, truncated Gaussian and uniform distributions, and discuss their corresponding $q_w(\epsilon)$ functions. It will be shown that for both distributions below, $q_w(\epsilon) = O(\epsilon)$.

Lemma 8 (Example of uniform distribution). *Consider w_t that follows a uniform distribution on $[-w_{\max}, w_{\max}]^{n_x}$. Then, $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}}$.*

Proof. Since $\mathrm{Unif}(\mathcal{W})$ is symmetric, we only need to consider one direction $j = 1$.

$$\mathbb{P}(w^j + w_{\max} \leq \epsilon) = \int_{w^1 + w_{\max} \leq \epsilon} \int_{w^2,\ldots,w^{n_x} \in [-w_{\max}, w_{\max}]} \frac{1}{(2w_{\max})^{n_x}} \mathbb{1}_{(w \in \mathcal{W})} \, \mathrm{d}w$$

$$= \int_{w^1 \leq \epsilon - w_{\max}} \frac{1}{2w_{\max}} \mathbb{1}_{(w \in \mathcal{W})} \mathrm{d}w^1 = \frac{\epsilon}{2w_{\max}}$$

Similarly, $\mathbb{P}(w_{\max} - w^1 \leq \epsilon) = \int_{w^1 \geq w_{\max} - \epsilon} \frac{1}{2w_{\max}} \mathbb{1}_{(w \in \mathcal{W})} \mathrm{d}w^1 = \frac{\epsilon}{2w_{\max}}$. $\square$

Lemma 9 (Example of truncated Gaussian distribution). *Consider w_t follows a truncated Gaussian distribution on $[-w_{\max}, w_{\max}]^{n_x}$ generated by a Gaussian distribution with zero mean and $\sigma_w I_{n_x}$ covariance matrix. Then, $q_w(\epsilon) = \frac{1}{\min(\sqrt{2\pi}\sigma_w, 2w_{\max})} \exp(\frac{-w_{\max}^2}{2\sigma_w^2})\epsilon$.*

---

[10]Though we only describe a static linear policy $u_t = Kx_t$ here, the results in [24], [96], [120] hold for dynamic linear policies.

Proof. Since this distribution is symmetric and each coordinate is independent, we only need to consider one direction j. Let X denote a Gaussian distribution with zero mean and σ_w^2 variance. By the definition of truncated Gaussian distributions, we have

$$\mathbb{P}(w^j + w_{\max} \le \epsilon) = \frac{\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon)}{\mathbb{P}(-w_{\max} \le X \le w_{\max})} \,.$$

Notice that X/σ_w follows the standard Gaussian distribution, so we can obtain the following bounds,

$$\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon) = \int_{-w_{\max}/\sigma_w}^{(-w_{\max}+\epsilon)/\sigma_w} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \, dz$$

$$\ge \frac{1}{\sqrt{2\pi}} \exp(-w_{\max}^2/(2\sigma_w^2)) \frac{\epsilon}{\sigma_w}$$

and

$$\mathbb{P}(-w_{\max} \le X \le w_{\max}) = \int_{-w_{\max}/\sigma_w}^{w_{\max}/\sigma_w} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \, dz$$

$$\le \min(1, \frac{1}{\sqrt{2\pi}} \frac{2w_{\max}}{\sigma_w}) \,.$$

Therefore, we obtain

$$\mathbb{P}(w^j + w_{\max} \le \epsilon) = \frac{\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon)}{\mathbb{P}(-w_{\max} \le X \le w_{\max})}$$

$$\ge \max(\frac{1}{\sqrt{2\pi}} \exp(-w_{\max}^2/\sigma_w^2) \frac{\epsilon}{\sigma_w}, \frac{\epsilon}{2w_{\max}} \exp(\frac{-w_{\max}^2}{2\sigma_w^2}))$$

$$= \frac{1}{\min(\sqrt{2\pi}\sigma_w, 2w_{\max})} \exp(\frac{-w_{\max}^2}{2\sigma_w^2}) \epsilon \,.$$

Finally, $\mathbb{P}(w_{\max} - w^1 \le \epsilon)$ can be bounded similarly. □

Lemma 10 (Example of uniform distribution on the boundary of $\mathcal{W}$ (a generalization of Rademacher distribution))**.** *Consider w_t follows a uniform distribution on $\{w : \|w\|_\infty = w_{\max}\}$. Then $q_w(\epsilon) = \frac{1}{2n_x}$.*

Proof. Since the hyper-cube $\{w : \|w\|_\infty = w_{\max}\}$ has $2n_x$ facets, the probability on each facet is $\frac{1}{2n_x}$. Therefore, $\mathbb{P}(w^j \le \epsilon - w_{\max}) \ge \mathbb{P}(w^j = -w_{\max}) = \frac{1}{2n_x}$ for all j. The same applies to $\mathbb{P}(w^j \ge -\epsilon + w_{\max})$. □

4.D Proof of Theorem 12

The section provides more details for the proof of Theorem 12. In particular, we first provide technical lemmas for set discretization, then prove Lemma 4 and Lemma 5 respectively. The proof of Theorem 12 follows naturally by combining Lemma 4 and Lemma 5.

Technical Lemmas: Set Discretization

This subsection provide useful technical lemmas for the proofs of Lemma 4 and Lemma 5. The results are based on a finite-ball covering result that is classical in the literature [159] [160].

Theorem 16 (Theorem 1.1 and 1.2 in [160] and Theorem 2 in [159] (revised to match the setting of this chapter)). *Consider a closed ball $\bar{\mathbb{B}}_n(0, 1) = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ in l_2 norm. Considering covering this ball $\bar{\mathbb{B}}_n(0, 1)$ with smaller closed balls $\bar{\mathbb{B}}_n(z, \epsilon)$ for $z \in \mathbb{R}^n$. Let $v_{\epsilon,n}$ denote the minimal number of smaller balls needed to cover $\bar{\mathbb{B}}_n(0, 1)$. For $n \geq 1$ and $0 < \epsilon < 1/2$, we have*

$$v_{\epsilon,n} \leq 544 n^{2.5} \log(n/\epsilon)(\frac{1}{\epsilon})^n$$

Proof. Theorem 1.1 and 1.2 in [160] and Theorem 2 in [159] discuss the upper bounds of $v_{\epsilon,n}$ in several different cases. These upper bounds in these different cases are unified by the upper bound in the theorem above by algebraic manipulations. □

We apply Theorem 16 to obtain the number of covering balls in the two settings below. These two settings will be considered in the proofs of Lemma 4 and 5 respectively.

Corollary 16.1. *There exists a finite set $\mathcal{M}' = \{\lambda_1, \ldots, \lambda_{v_\lambda}\} \subseteq \mathbb{S}_{n_z}(0, 1)$ such that for any $\lambda \in \mathbb{R}^{n_z}$ with $\|\lambda\|_2 = 1$, there exists $\lambda_i \in \mathcal{M}'$ such that $\|\lambda - \lambda_i\|_2 \leq 2\epsilon_\lambda$.*

In the following, we consider $\epsilon_\lambda = \sigma_z^2 p_z^2/(64 b_z^2) = 1/a_2$. Notice that $\epsilon_\lambda < 1/2$. Accordingly,

$$v_\lambda \leq 544 n_z^{2.5} \log(a_2 n_z) a_2^{n_z}. \tag{4.7}$$

Proof. $\epsilon_\lambda \leq 1/64 < 1/2$ because $p_z \leq 1$ and $\sigma_z \leq b_z$ by the definitions of BMSB and b_z. Then, the bound on v_λ follows from Theorem 16. □

Lemma 11. *There exists a finite set $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\} \subseteq \mathbb{S}_{n_x \times n_z}(0, 1)$ such that for any $\gamma \in \mathbb{R}^{n_x \times n_z}$ and $\|\gamma\|_F = 1$, there exists $\gamma_i \in \mathcal{M}$ such that $\|\gamma - \gamma_i\|_F \leq 2\epsilon_\gamma$. Consider $\epsilon_\gamma = \frac{a_1}{4 b_z \sqrt{n_x}} = 1/a_4$. Notice that $\epsilon_\gamma < 1/2$. Accordingly,*

$$v_\gamma \leq 544 n_x^{2.5} n_z^{2.5} \log(a_4 n_x n_z) a_4^{n_z n_x}.$$

Proof. The proof is basically by mapping the matrices to vectors based on matrix vectorization, then mapping the vectors back to matrices. These two mappings are isomorphism.

Specifically, consider a closed unit ball in $\mathbb{R}^{n_x n_z}$. There exist $v_{\epsilon, n_x n_z}$ smaller closed balls to cover it, denoted by $\mathbb{B}_1, \ldots, \mathbb{B}_{v_{\epsilon, n_x n_z}}$. Consider the non-empty sets from $\mathbb{B}_1 \cap \mathbb{S}_{n_x n_z}(0, 1), \ldots, \mathbb{B}_{v_{\epsilon, n_x n_z}} \cap \mathbb{S}_{n_x n_z}(0, 1)$. For any $1 \leq i \leq v_{\epsilon, n_x n_z}$, if $\mathbb{B}_i \cap \mathbb{S}_{n_x n_z}(0, 1) \neq \emptyset$, select a point $\text{vec}(\gamma) \in \mathbb{B}_i \cap \mathbb{S}_{n_x n_z}(0, 1)$. Notice that $\|\text{vec}(\gamma)\|_2 = 1$. In this way, we construct a finite sequence $\{\text{vec}(\gamma_1), \ldots, \text{vec}(\gamma_{v_\gamma})\}$ where $v_\gamma \leq v_{\epsilon_\gamma, n_x n_z}$.[11]

For any $\gamma \in \mathbb{R}^{n_x \times n_z}$, we have $\text{vec}(\gamma) \in \mathbb{R}^{n_x n_z}$ and $\|\text{vec}(\gamma)\|_2 = 1$. Hence, there exists $1 \leq i \leq v_\gamma$ such that $\text{vec}(\gamma) \in \mathbb{B}_i \cap \mathbb{S}_{n_x n_z}(0, 1)$. Hence, $\|\text{vec}(\gamma) - \text{vec}(\gamma_i)\|_2 \leq 2\epsilon_\gamma$. Moreover, $\|\gamma_i\|_F = \|\text{vec}(\gamma_i)\|_2 = 1$. Therefore, $\|\gamma_i - \gamma\|_F \leq 2\epsilon_\gamma$. So the set $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\}$ satisfies our requirement. □

Proof of Lemma 4

Essentially, Lemma 4 shows that PE holds with high probability under the BMSB condition. This result has been established in Proposition 2.5 in [96], though in a different form. The rest of this subsection will prove the PE condition needed in this chapter based on Proposition 2.5 in [96].

Firstly, we review Proposition 2.5 in [96] for the convenience of the reader.

Theorem 17 (Proposition 2.5 in [96] when $k = 1$). *Let $\{Z_t\}_{t \geq 1}$ be an $\{\mathcal{F}_t^Z\}_{t \geq 1}$-adapted random process taking values in $\mathbb{R}$. Z_0 is given. If $\{Z_t\}_{t \geq 0}$ is $(1, v, p)$-BMSB, then*

$$\mathbb{P}(\sum_{t=1}^{T} Z_t^2 \leq v^2 p^2 T/8) \leq \exp(-T p^2/8)$$

Next, we prove the PE in one segment of data sequence.

Lemma 12 (Probability of PE in one segment). *For any $m \geq 1$, for any $k \geq 0$, we have*

$$\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top > (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) \geq 1 - v_\lambda \exp(-m p_z^2/8))$$

---

[11]Here, without loss of generality, we consider $\mathbb{B}_1 \cap \mathbb{S}_{n_x n_z}(0, 1), \ldots, \mathbb{B}_{v_\gamma} \cap \mathbb{S}_{n_x n_z}(0, 1)$ are not empty.

Proof. Consider $\mathcal{M}' = \{\lambda_1, \ldots, \lambda_{\nu_\lambda}\}$ defined in Corollary 16.1. For any $\lambda_i \in \mathcal{M}'$, $\lambda_i^\top z_t$ satisfies the $(1, \sigma_z, p_z)$-BMSB condition. Therefore, by Theorem 17, we have

$$\mathbb{P}(\sum_{t=1}^{T} \lambda_i^\top z_t z_t^\top \lambda_i \leq \sigma_z^2 p_z^2 T/8) \leq \exp(-Tp_z^2/8).$$

Notice that the horizon length T is arbitrary and the starting stage $t = 1$ can also be different because we consider a time-invariant dynamical system in this chapter. Therefore, for any $m \geq 1, k \geq 0$, for any $\lambda_i \in \mathcal{M}'$, we have

$$\mathbb{P}(\sum_{i=1}^{m} \lambda_i^\top z_{km+i} z_{km+i}^\top \lambda_i \leq \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km}) \leq \exp(-mp_z^2/8),$$

where we condition on $\mathcal{F}_{km}$ to make sure z_{km} is known under $\mathcal{F}_{km}$, which is required by Theorem 17.

For arbitrary λ such that $\|\lambda\|_2 = 1$, there exists $\lambda_i \in \mathcal{M}'$ such that $\|\lambda - \lambda'\|_2 \leq 2\epsilon_\lambda$. Therefore, we can bound $\sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda$ by $\sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i$.

$$
\begin{aligned}
\sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda &= \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i + \sum_{t=km+1}^{km+m} (\lambda + \lambda_i)^\top z_t z_t^\top (\lambda - \lambda_i) \\
&\geq \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sum_{t=km+1}^{km+m} \|\lambda + \lambda_i\|_2 \|z_t\|_2^2 \|\lambda_i - \lambda\|_2 \\
&\overset{(a)}{\geq} \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sum_{t=km+1}^{km+m} 4b_z^2 \epsilon_\lambda \\
&= \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - 4b_z^2 \epsilon_\lambda m \overset{(b)}{\geq} \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sigma_z^2 p_z^2 m/16,
\end{aligned}
$$

where (a) is by Assumption 4, $\|\lambda - \lambda_i\|_2 \leq 2\epsilon_\lambda$, and $\|\lambda\|_2 = \|\lambda_i\|_2 = 1$; and (b) is by choosing $\epsilon_\lambda \leq \sigma_z^2 p_z^2/(64b_z^2)$.

Therefore, by the definition of positive definiteness and the inequalities above, we can complete the proof by the following:

$$
\begin{aligned}
\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top > (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) &= \mathbb{P}(\forall \|\lambda\|_2 = 1, \sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda > \sigma_z^2 p_z^2 m/16 \mid \mathcal{F}_{km}) \\
&\geq \mathbb{P}(\forall 1 \leq i \leq \nu_\lambda, \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i > \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km}) \\
&\geq 1 - \sum_{i=1}^{\nu_\lambda} \mathbb{P}(\sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i \leq \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km})
\end{aligned}
$$

$$\geq 1 - v_\lambda \exp(-mp_z^2/8)),$$

which completes the proof. □

Now, we are ready for the proof of Lemma 4.

Proof of Lemma 4. Recall that $\mathcal{E}_2 = \{\frac{1}{m}\sum_{s=1}^{m} z_{km+s}z_{km+s}^\top \geq a_1^2 I_{n_z}, \forall 0 \leq k \leq \lceil T/m \rceil - 1\}$, where $a_1 = \sigma_z p_z/4$. Hence

$$\mathcal{E}_2 = \bigcap_{k=0}^{T/m-1} \{ \sum_{t=km+1}^{km+m} z_t z_t^\top > (\sigma_z^2 p_z^2 m/16) I_{n_z} \}.$$

Therefore,

$$\mathbb{P}(\mathcal{E}_2) \geq 1 - \sum_{k=0}^{T/m-1} \mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \leq (\sigma_z^2 p_z^2 m/16) I_{n_z})$$

$$\geq 1 - \frac{T}{m} v_\lambda \exp(-mp_z^2/8)$$

$$= 1 - \frac{T}{m}(544 n_z^{2.5} \log(a_2 n_z) a_2^{n_z}) \exp(-mp_z^2/8),$$

where we use Lemma 12 and the fact that if $\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \leq (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) \leq v_\lambda \exp(-mp_z^2/8)$, then $\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \leq (\sigma_z^2 p_z^2 m/16) I_{n_z}) \leq v_\lambda \exp(-mp_z^2/8)$. □

Proof of Lemma 5

This proof takes four major steps:

 (i) Define $b_{i,t}, j_{i,t}, L_{i,k}$.

 (ii) Provide a formal definition of $\mathcal{E}_{1,k}$ based on $b_{i,t}, j_{i,t}, L_{i,k}$ and prove a formal version of Lemma 6.

(iii) Prove Lemma 7.

(iv) Prove Lemma 5 by the formal version of Lemma 6 and Lemma 7.

It is worth mentioning that the formal definition of $\mathcal{E}_{1,k}$ is slightly different from the definition in Lemma 6, but we still have $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,k} \cap \mathcal{E}_2)$, which is the key property that will be used in the proof of Lemma 5.

Step (i): definitions of $b_{i,t}, j_{i,t}, L_{i,k}$.

Recall the discretization of $\mathbb{S}_{n_x \times n_z}(0, 1)$ in Lemma 11, which generates the set $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{\nu_\gamma}\}$. We are going to define $b_{i,t}, j_{i,t}, L_{i,k}$ for $\gamma_i \in \mathcal{M}$ for each $1 \le i \le \nu_\gamma$. Notice that $\mathcal{M}$ is a deterministic set of matrices.

Lemma 13 (Definition of $b_{i,t}, j_{i,t}$). *For any $\gamma_i \in \mathcal{M}$, any $0 \le t \le T$, there exist $b_{i,t} \in \{-1, 1\}$ and $1 \le j_{i,t} \le n_x$ such that $b_{i,t}, j_{i,t} \in \mathcal{F}(z_t) \subseteq \mathcal{F}_t$ and*

$$\|\gamma_i z_t\|_\infty = b_{i,t}(\gamma_i z_t)^{j_{i,t}}.$$

Note that one way to determine $b_{i,t}, j_{i,t}$ from z_t is by the following: first pick the smallest j such that $|(\gamma_i z_t)^j| = \|\gamma_i z_t\|_\infty$, then let $b_{i,t} = \text{sgn}((\gamma_i z_t)^j)$, where $\text{sgn}(\cdot)$ denotes the sign of a scalar argument.

Proof. For any $\gamma_i \in \mathcal{M}$, any $0 \le t \le T$, we have

$$\|\gamma_i z_t\|_\infty = \max_{1 \le j \le n_x} \max_{b \in \{-1,1\}} b(\gamma_i z_t)^j .$$

Hence, there exist $b_{i,t}, j_{i,t}$ such that $\|\gamma_i z_t\|_\infty = b_{i,t}(\gamma_i z_t)^{j_{i,t}}$. Further, $b_{i,t}, j_{i,t}$ only depend on γ_i and z_t, so they are $\mathcal{F}(z_t)$-measurable, and $\mathcal{F}(z_t) \subseteq \mathcal{F}_t$. $\square$

Lemma 14 (Definition of stopping times $L_{i,k}$). *Let $\eta = \frac{a_1}{\sqrt{n_x}}$. For any $\gamma_i \in \mathcal{M}$, any $0 \le k \le T/m - 1$, we can define a random time index $1 \le L_{i,k} \le m + 1$ by*

$$L_{i,k} = \min(m + 1, \min\{l \ge 1 : \|\gamma_i z_{km+l}\|_\infty \ge \eta\}).$$

Then, we have $1 \le L_{i,k} \le m + 1$. Further, for any $1 \le l \le m$, $\{L_{i,k} = l\} \in \mathcal{F}_{km+l}$, and $\{L_{i,k} = m + 1\} \in \mathcal{F}_{km+m} \subseteq \mathcal{F}_{km+m+1}$. In other words, $L_{i,k}$ is a stopping time with respect to filtration $\{F_{km+l}\}_{l \ge 1}$.

Proof. For any i and any k, it is straightforward to see that $L_{i,k}$ is well-defined and $1 \le L_{i,k} \le m + 1$.

When $L_{i,k} = l \le m$, this is equivalent with $\|\gamma_i z_{km+l}\|_\infty \ge \eta$ but $\|\gamma_i z_{km+s}\| < \eta$ for $1 \le s < l$. Notice that this event is only determined by $z_{km+l}, \ldots, z_{km+1}$, so $\{L_{i,k} = l\} \in \mathcal{F}_{km+l}$.

When $L_{i,k} = m + 1$, this is equivalent with $\|\gamma_i z_{km+s}\| < \eta$ for $1 \le s \le m$. Notice that this event is only determined by $z_{km+m}, \ldots, z_{km+1}$, so $\{L_{i,k} = m + 1\} \in \mathcal{F}_{km+m}$.

Therefore, by definition, $L_{i,k}$ is a stopping time with respect to filtration $\{\mathcal{F}_{km+l}\}_{l \ge 1}$.

$\square$

Step (ii): a formal version of Lemma 6 and its proof

Lemma 15 (Discretization of $\mathcal{E}_1 \cap \mathcal{E}_2$ (Formal version of Lemma 6)). *Let* $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\}$ *be an* ϵ_γ*-net of* $\{\gamma : \|\gamma\|_F = 1\}$ *as defined in Lemma 11, where* $\epsilon_\gamma = \min(\frac{a_1}{4b_z\sqrt{n_x}}, 1)$, $v_\gamma = \widetilde{O}(n_x^{2.5}n_z^{2.5})a_4^{n_x n_z}$, *and* $a_4 = \frac{4b_z\sqrt{n_x}}{a_1}$. *Define*

$$\mathcal{E}_{1,i} = \{\exists \gamma \in \Gamma_T, \text{ s.t. } b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1\delta}{4\sqrt{n_x}}, \forall k \geq 0\}.$$

Then, we have

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2).$$

The rest of this subsubsection is dedicated to the proof of Lemma 15. As an overview: firstly, we will discuss the implications of $\mathcal{E}_2$ on $\gamma_i \in \mathcal{M}$. Then, we discuss the implications of $\mathcal{E}_2$ on any γ. Lastly, we prove Lemma 15 by combining the implications of $\mathcal{E}_2$ on any γ and $\|\gamma\|_F \geq \delta/2$.

Lemma 16 (The implication of $\mathcal{E}_2$ on γ_i). *If* $\mathcal{E}_2$ *happens, then for any* $\gamma_i \in \mathcal{M}$, *any* $0 \leq k \leq T/m - 1$, *we have*

$$\max_{1 \leq s \leq m} \|\gamma_i z_{km+s}\|_\infty \geq \frac{a_1}{\sqrt{n_x}}.$$

Therefore, almost surely, we have $1 \leq L_{i,k} \leq m$ *and*

$$b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}.$$

Proof. If $\mathcal{E}_2$ happens, then by definition, we have

$$\frac{1}{m}\sum_{s=1}^{m} z_{km+s}z_{km+s}^\top \geq a_1^2 I_{n_z},$$

for all $0 \leq k \leq T/m - 1$.

Now, for any $\gamma_i \in \mathcal{M}$, we have that

$$\frac{1}{m}\sum_{s=1}^{m} \gamma_i z_{km+s}z_{km+s}^\top\gamma_i^\top \geq a_1^2\gamma_i\gamma_i^\top. \tag{4.8}$$

Therefore, by taking trace at each side of (4.8), we obtain

$$\frac{1}{m}\sum_{s=1}^{m} \text{tr}(\gamma_i z_{km+s}z_{km+s}^\top\gamma_i^\top) \geq a_1^2\text{tr}(\gamma_i\gamma_i^\top). \tag{4.9}$$

Since $\gamma_i \in \mathbb{S}_{n_x \times n_z}(0, 1)$, we have $\|\gamma_i\|_F = 1$, so $\text{tr}(\gamma_i \gamma_i^\top) = \text{tr}(\gamma_i^\top \gamma_i) = \|\gamma_i\|_F^2 = 1$. Further, we have

$$\text{tr}(\gamma_i z_{km+s} z_{km+s}^\top \gamma_i^\top) = \text{tr}(z_{km+s}^\top \gamma_i^\top \gamma_i z_{km+s}) = z_{km+s}^\top \gamma_i^\top \gamma_i z_{km+s} = \|\gamma_i z_{km+s}\|_2^2.$$

Consequently, we have

$$\frac{1}{m} \sum_{s=1}^m \|\gamma_i z_{km+s}\|_2^2 \geq a_1^2$$

for all k.

By the pigeonhole principle, we have that

$$\max_{1 \leq s \leq m} \|\gamma_i z_{km+s}\|_2^2 \geq a_1^2.$$

This is equivalent with $\max_{1 \leq s \leq m} \|\gamma_i z_{km+s}\|_2 \geq a_1$.

Notice that $\|\gamma_i z_{km+s}\|_2 \leq \sqrt{n_x} \|\gamma_i z_{km+s}\|_\infty$, so $\max_{1 \leq s \leq m} \sqrt{n_x} \|\gamma_i z_{km+s}\|_\infty \geq a_1$, which completes the proof of the first inequality in the lemma statement.

Next, we prove the second inequality in the lemma statement. Notice that by the definition of $L_{i,k}$ in Lemma 14 and by $\eta = \frac{a_1}{\sqrt{n_x}}$, we have $1 \leq L_{i,k} \leq m$ and $\|\gamma_i z_{km+L_{i,k}}\|_\infty \geq \frac{a_1}{\sqrt{n_x}}$ for all k. Further, by Lemma 13, we have $\|\gamma_i z_{km+L_{i,k}}\|_\infty = b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}}$ almost surely. Hence, we have $b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}$, which completes the proof.

$\square$

Lemma 17 (The implication of $\mathcal{E}_2$ on γz_t)**.** *If $\mathcal{E}_2$ happens, then for any $\gamma \in \mathbb{R}^{n_x \times n_z}$, there exists $1 \leq i \leq v_\gamma$, such that*

$$b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}} \|\gamma\|_F,$$

for all $0 \leq k \leq T/m - 1$.

Proof. Firstly, when $\gamma = 0$, the inequality holds because both sides are 0.

Next, when $\gamma \neq 0$, it suffices to prove $b_{i,km+L_{i,k}}(\frac{\gamma}{\|\gamma\|_F} z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}}$. Therefore, we will only consider $\gamma \in \mathbb{S}_{n_x \times n_z}(0, 1)$. By Lemma 11, there exists $\gamma_i \in \mathcal{M}$ such that $\|\gamma - \gamma_i\|_F \leq 2\epsilon_\gamma = \min(\frac{a_1}{2b_z \sqrt{n_x}}, 2)$. Notice that by Lemma 16, if $\mathcal{E}_2$ happens, for all k, we have

$$b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}.$$

Therefore,

$$
\begin{aligned}
b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} &= b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \\
&\quad - b_{i,km+L_{i,k}}((\gamma_i - \gamma) z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \\
&\geq \frac{a_1}{\sqrt{n_x}} - |b_{i,km+L_{i,k}}((\gamma_i - \gamma) z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}}| \\
&\geq \frac{a_1}{\sqrt{n_x}} - \|(\gamma_i - \gamma) z_{km+L_{i,k}}\|_2 \\
&\geq \frac{a_1}{\sqrt{n_x}} - \|\gamma_i - \gamma\|_2 \|z_{km+L_{i,k}}\|_2 \\
&\geq \frac{a_1}{\sqrt{n_x}} - 2\epsilon_\gamma b_z \geq \frac{a_1}{2\sqrt{n_x}}.
\end{aligned}
$$

$\square$

Proof of Lemma 15. By Lemma 17, under $\mathcal{E}_2$, for any $\gamma \in \mathbb{R}^{n_x \times n_z}$, there exists $1 \leq i \leq v_\gamma$, such that

$$
b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}} \|\gamma\|_F,
$$

for all $0 \leq k \leq T/m - 1$. Therefore, if $\mathcal{E}_1 \cap \mathcal{E}_2$ happens, there exists $\gamma \in \Gamma_T$ and a corresponding i, such that

$$
b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}} \|\gamma\|_F \geq \frac{a_1 \delta}{4\sqrt{n_x}}.
$$

Therefore,

$$
\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\bigcup_{i=1}^{v_\gamma} \mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2),
$$

which completes the proof. $\square$

Proof of Lemma 7

Notice that Lemma 7 states two inequalities: in the following, we will first prove the first inequality $\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}(\cap_{k=0}^{T/m-1} G_{i,k})$, then prove the second inequality on $\mathbb{P}(G_{i,k} \mid \cap_{k'=0}^{k-1} G_{i,k'})$.

Lemma 18 (Bound $\mathcal{E}_{1,i} \cap \mathcal{E}_2$ by $G_{i,k}$). *Under the conditions in Lemma 7, for any* $1 \leq i \leq v_\gamma$, *we have*

$$
\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}\left(\bigcap_{k=0}^{T/m-1} G_{i,k} \right).
$$

Proof. Firstly, for any $\gamma \in \Gamma_T$, we have $\|w_t - \gamma z_t\|_\infty \le w_{\max}$ for all $t \ge 0$. This suggests that, for any $1 \le j \le n_x$, we have

$$-w_{\max} \le w_t^j - (\gamma z_t)^j \le w_{\max}.$$

Hence, we have $b(\gamma z_t)^j \le bw_t^j + w_{\max}$ for any $b \in \{-1, 1\}$, $1 \le j \le n_x$, and $t \ge 0$.

Next, by $\mathcal{E}_{1,i}$, there exists $\gamma \in \Gamma_T$ such that $b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \ge \frac{a_1 \delta}{4\sqrt{n_x}}$ for all $k \ge 0$. Therefore, $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}}$ for all k.

Finally, $\mathcal{E}_{1,i} \cap \mathcal{E}_2$ implies that $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}}$ and $\frac{1}{m}\sum_{s=1}^m z_{km+s} z_{km+s}^\top \ge a_1^2 I_{n_z}$ for all k, which is $\bigcap_k G_{i,k}$ by the definition of $G_{i,k}$.

$\square$

Lemma 19 (Bound on $\mathbb{P}(G_{i,k} \mid \cap_{k'=0}^{k-1} G_{i,k'})$). *Under the conditions in Lemma 7, for any $1 \le i \le v_\gamma$ and any $k \ge 0$, we have*

$$\mathbb{P}(G_{i,k} \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \le 1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}).$$

Proof. Firstly, notice that when $\frac{1}{m}\sum_{s=1}^m z_{km+s} z_{km+s}^\top \ge a_1^2 I_{n_z}$, we have $1 \le L_{i,k} \le m$ by the proof of Lemma 16. Therefore, we have

$$\mathbb{P}(G_{i,k} \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \le \mathbb{P}(b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}}, 1 \le L_{i,k} \le m \mid \bigcap_{k'=0}^{k-1} G_{i,k'})$$

$$\le \sum_{l=1}^m \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}}, L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'})$$

$$\le \sum_{l=1}^m \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \mathbb{P}(L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'})$$

$$\overset{(a)}{\le} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}})) \sum_{l=1}^m \mathbb{P}(L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'})$$

$$\le 1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}).$$

The inequality (c) is proved in the following:

$$\mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'})$$

$$= \int_{v_{0:km+l}} \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \ge \frac{a_1 \delta}{4\sqrt{n_x}}, w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) dv_{0:km+l}$$

$$= \int_{v_{0:km+l} \in S_{km+l}} \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1 \delta}{4\sqrt{n_x}} \mid w_{0:km+l} = v_{0:km+l})$$

$$\times \mathbb{P}(w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) dv_{0:km+l}$$

$$\overset{(b)}{\leq} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}})) \int_{v_{0:km+l} \in S_{km+l}} \mathbb{P}(w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) dv_{0:km+l}$$

$$= 1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}),$$

where we define a shorthand notation $w_{0:km+l} = (w_0, \ldots, w_{km+l-1})$, and we use $v_{0:km+l}$ to denote a realization of $w_{0:km+l}$, then we define the set of values of $w_{0:km+l}$ as S_{km+l} such that $L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}$ holds. Notice that $L_{i,k} = l$ can be determined by a set of values of $w_{0:km+l}$ because $L_{i,k}$ is a stopping time of $\{F_{km+l}\}_{l \geq 1}$ and thus $\{L_{i,k} = l\} \in \mathcal{F}_{km+l}$. The inequality (b) above is because of the following: firstly, notice that $b_{i,km+l}, j_{i,km+l} \in \mathcal{F}_{km+l}$, so $b_{i,km+l}, j_{i,km+l}$ are deterministic values when $w_{0:km+l} = v_{0:km+l}$. Further, since w_{km+l} is independent of $w_{0:km+l}$, we have $\mathbb{P}(w_{max} + bw_{km+l}^j \geq \epsilon \mid w_{0:km+l} = v_{0:km+l}) \leq 1 - q_w(\epsilon)$ for any deterministic b, j and any $\epsilon > 0$ by Assumption 5. Hence, we have $\mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1 \delta}{4\sqrt{n_x}} \mid w_{0:km+l} = v_{0:km+l}) \leq 1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}})$. □

Proof of Lemma 5

The proof is by leveraging Lemma 15 and Lemma 7.

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2)$$

$$\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k})$$

$$= \sum_{i=1}^{v_\gamma} \mathbb{P}(G_{i,0})\mathbb{P}(G_{i,1} \mid G_{i,0}) \cdots \mathbb{P}(G_{i,T/m-1} \mid \bigcap_{k=0}^{T/m-2} G_{i,k})$$

$$\leq \sum_{i=1}^{v_\gamma} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}))^{T/m}$$

$$\leq 544 n_x^{2.5} n_z^{2.5} \log(a_4 n_x n_z) a_4^{n_z n_x} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}))^{T/m}.$$

A More Precise Upper Bound for Theorem 12

By the proof of Lemma 4 and Lemma 5 above, we have

$$\mathbb{P}(\text{diam}(\Theta_T) > \delta) \leq 544 \frac{T}{m} n_z^{2.5} \log(a_2 n_z) a_2^{n_z} \exp(-a_3 m)$$
$$+ 544 n_x^{2.5} n_z^{2.5} \log(a_4 n_x n_z) a_4^{n_z n_x} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}))^{T/m} \quad (4.10)$$

4.E   Proof of Corollary 12.1

The proof involves two parts. Firstly, we will show that Term $1 \leq \epsilon$ under our choice of m. Secondly, we will let Term $2 = \epsilon$, then we will show $\delta \leq \widetilde{O}(n_x^{1.5} n_z^2 / T)$, which completes the proof.

Step 1: show Term 1 $\leq \epsilon$. Notice that when $m \geq \frac{1}{a_3}(\log(\frac{T}{\epsilon}) + n_z \log(a_2) + 2.5 \log(n_z) + \log \log(a_2 n_z) + 7) = O(n_z + \log T + \log(1/\epsilon))$, we have $T\widetilde{O}(n_z^{2.5}) a_2^{n_z} \exp(-a_3 m) \leq \epsilon$. Since $m \geq 1$, we obtain Term $1 \leq \epsilon$.

Step 2: let Term 2 $= \epsilon$ and show $\delta \leq \widetilde{O}(n_x^{1.5} n_z^2 / T)$. Let Term $2 = \epsilon$, then we have $(1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}))^{T/m} = \frac{\epsilon}{\widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}$. Then, we obtain $(1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}})) = \left(\frac{\epsilon}{\widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}$, which is equivalent with

$$q_w(\frac{a_1 \delta}{4\sqrt{n_x}}) = 1 - \left(\frac{\epsilon}{\widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}.$$

When $q_w(\frac{a_1 \delta}{4\sqrt{n_x}}) = O(\frac{a_1 \delta}{4\sqrt{n_x}})$, we obtain

$$\delta = O(\frac{4\sqrt{n_x}}{a_1})\left(1 - \left(\frac{\epsilon}{\widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}\right)$$

$$\leq O(\frac{-4\sqrt{n_x}}{a_1}) \log\left(\left(\frac{\epsilon}{\widetilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}\right)$$

$$= O(\frac{4\sqrt{n_x} m}{a_1 T})(\log(1/\epsilon) + n_x n_z + \log(n_x n_z))$$

$$= \widetilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right).$$

Step 3: prove Corollary 12.1. By leveraging the bounds above and Theorem 12, we have $\mathbb{P}(\text{diam}(\Theta_T) \leq \widetilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right)) \geq \mathbb{P}(\text{diam}(\Theta_T) \leq \delta) \geq 1 - 2\epsilon$.

Since $\theta^\star \in \Theta_T$ by definition, for any $\widehat{\theta}_T \in \Theta_T$, we have $\|\widehat{\theta}_T - \theta^\star\|_F \leq \text{diam}(\Theta_T) \leq \widetilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right)$ with probability at least $1 - 2\epsilon$.

4.F Proof of Corollary 12.2

We provide a formal version of Corollary 12.2 and its proof below.

Corollary 17.1 (Convergence rate when $B^\star = 0$ (formal version))**.** *When $A^\star$ is (κ, ρ)-stable, i.e., $\|(A^\star)^t\|_2 \le \kappa(1 - \rho)^t$ for all t with $\rho < 1$, for any $m > 0$ and any $\delta > 0$, when $T > m$, we have*

$$\mathbb{P}(\text{diam}(\mathbb{A}_T) > \delta) \le \frac{T}{m}\widetilde{O}(n_x^{2.5})a_2^{n_x}\exp(-a_3 m) + \widetilde{O}(n_x^5)a_4^{n_x^2}(1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{\lceil T/m \rceil}$$

where $b_x = \kappa\|x_0\|_2 + \kappa\sqrt{n_x}/\rho$, $p_x = 1/192$, $\sigma_x = \sqrt{\lambda_{\min}(\Sigma_w)/2}$, $a_1 = \frac{\sigma_x p_x}{4}$, $a_2 = \frac{64w_{\max}}{\sigma_x^2 p_x^2}$, $a_3 = \frac{p_x^2}{8}$, $a_4 = \frac{4b_x\sqrt{n_x}}{a_1}$.
Consequently, when the distribution of w_t satisfies $q_w(\epsilon) = O(\epsilon)$, e.g. uniform or truncated Gaussian, we have $\|\widehat{\theta} - \theta_\star\| \le \widetilde{O}(n_x^{3.5}/T)$.

The proof of Corollary 12.2 is exactly the same as the proofs of Theorem 12 and Corollary 12.1. When $A^\star$ is stable, we can show that $\|x_t\|_2 \le b_x$ for all t. Further, by [24], the sequence $\{x_t\}_{t \ge 0}$ satisfies the $(1, \sigma_x, p_x)$-BMSB condition. Therefore, we complete the proof.

4.G Proof of Theorem 13

Specifically, we define $\epsilon_0 = \frac{4\sqrt{n_x}}{a_1}(\widehat{w}_{\max} - w_{\max})$.

The proof is similar to the proof of Theorem 12. Firstly, we define $\widehat{\Gamma}_T$ as a translation of the set $\widehat{\Theta}_T$:

$$\widehat{\Gamma}_t = \bigcap_{s=0}^{t-1}\{\gamma : \|w_s - \gamma z_s\|_\infty \le \widehat{w}_{\max}\}, \quad \forall t \ge 0. \tag{4.11}$$

Notice that

$$\widehat{\Theta}_T = \theta^\star + \widehat{\Gamma}_T$$

by considering $\gamma = \widehat{\theta} - \theta^\star$. Therefore, we can upper bound our goal event $\{\text{diam}(\widehat{\Theta}_T) > \delta + \epsilon_0\}$ by the event $\mathcal{E}_3$ defined below.

$$\mathbb{P}(\text{diam}(\widehat{\Theta}_T) > \delta + \epsilon_0) \le \mathbb{P}(\mathcal{E}_3), \text{ where } \mathcal{E}_3 := \{\exists\, \gamma \in \widehat{\Gamma}_T, \text{ s.t. } \|\gamma\|_F \ge \frac{\delta + \epsilon_0}{2}\}. \tag{4.12}$$

Next, notice that

$$\mathbb{P}(\text{diam}(\widehat{\Theta}_T) > \delta + \epsilon_0) \le \mathbb{P}(\mathcal{E}_3) \le \mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) + \mathbb{P}(\mathcal{E}_2^c).$$

By Lemma 4, we have already shown $\mathbb{P}(\mathcal{E}_2^c) \le$ Term 1. So we only need to discuss $\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2)$.

Lemma 20.

$$\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) \leq \text{Term 2}$$

Proof. Firstly, define

$$\mathcal{E}_{3,i} = \{\exists \gamma \in \widehat{\Gamma}_T, \text{ s.t. } b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1(\delta + \epsilon_0)}{4\sqrt{n_x}}, \ \forall k \geq 0\}.$$

We have $\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{3,i} \cap \mathcal{E}_2)$ based on the same proof ideas of Lemma 15.

Next, we will show that

$$\Pr(\mathcal{E}_{3,k} \cap \mathcal{E}_2) \leq \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}). \tag{4.13}$$

This is because for any $\gamma \in \widehat{\Gamma}_T$, we have $b(\gamma z_t)^j \leq bw_t^j + \widehat{w}_{\max}$ for any $b \in \{-1, 1\}$, $1 \leq j \leq n_x$, and $t \geq 0$. By $\mathcal{E}_{3,i}$, there exists $\gamma \in \widehat{\Gamma}_T$ such that $b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1(\delta+\epsilon_0)}{4\sqrt{n_x}}$ for all $k \geq 0$. Thus, $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + \widehat{w}_{\max} \geq \frac{a_1(\delta+\epsilon_0)}{4\sqrt{n_x}}$ for all k. Notice that this is equivalent with $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}$ for all k because $\epsilon_0 = \frac{4\sqrt{n_x}}{a_1}(\widehat{w}_{\max} - w_{\max})$. In this way, we can prove (4.13).

Finally, we can complete the proof by the following:

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) &\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{3,i} \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}) \\
&= \sum_{i=1}^{v_\gamma} \mathbb{P}(G_{i,0})\mathbb{P}(G_{i,1} \mid G_{i,0}) \cdots \mathbb{P}(G_{i,T/m-1} \mid \bigcap_{k=0}^{T/m-2} G_{i,k}) \\
&\leq \sum_{i=1}^{v_\gamma} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{T/m} \leq \text{Term 1},
\end{aligned}$$

where the second last inequality is by Lemma 19 and the last inequality uses the definition of v_γ in Lemma 11. $\qquad\square$

4.H   Proofs of Theorem 14, Corollary 14.1, and Theorem 15

This section provides proofs of the main results related to the SME with unknown $w_{\max}$ as discussed in section 4.3. Namely, Theorem 14 and Corollary 14.1 provide the rate of convergence of the estimator $\bar{w}_{\max}^{(T)}$ defined in (4.4) to $w_{\max}$, and Theorem 15 states the rate of convergence of UCB-SME algorithm introduced in (4.5).

For ease of notation, we introduce the following function indexed by the time horizon $T > 0$,

$$W_T : \theta \mapsto \max_{0 \leq t \leq T-1} \|x_{t+1} - \theta z_t\|_\infty. \qquad (4.14)$$

The estimator $\bar{w}_{\max}^{(T)}$ is simply the infimum of this function, i.e., $\bar{w}_{\max}^{(T)} = \inf_\theta W_T(\theta)$.

Proof of Theorem 14

The proof of Theorem 14 involves two steps:

- *Step 1:* We demonstrate that the learning error of $w_{\max}$ incurred by the estimator $\bar{w}_{\max}^{(T)}$ is governed by the diameter of the uncertainty set Θ_T and the minimum learning error achievable if θ^* were known.

- *Step 2:* We then provide an upper bound the probability of learning error exceeding a fixed threshold.

Before we proceed with the proof of Theorem 14, we present the the following technical lemma.

Lemma 21. *Consider the sequence of functions $\{W_T\}_{T>0}$ defined in (4.14). The following holds:*

- i. W_T *is convex in* $\mathbb{R}^{n_x \times n_z}$,

- ii. *The sequence* $\{\inf_\theta W_T(\theta)\}_{T>0}$ *is bounded and monotonically non-decreasing, i.e.,*

$$0 \leq \inf_\theta W_T(\theta) \leq \inf_\theta W_{T+1}(\theta) \leq w_{\max},$$

  for all $T > 0$,

- iii. W_T *attains its minimum in* Θ_T, *i.e.,* $\arg\min_\theta W_T(\theta) \subset \Theta_T$.

Proof. (*i.*) For $0 \leq t \leq T - 1$, the function $\theta \mapsto \|x_{t+1} - \theta z_t\|_\infty$ is convex due to convexity of norms. Since the maximum of convex functions is convex [161], convexity of W_T follows.

(*ii.*) Notice that W_{T+1} can be defined in terms of W_T recursively as $W_{T+1}(\theta) = \max(W_T(\theta), \|x_{T+1} - \theta z_T\|_\infty)$. Thus, $W_T(\theta) \leq W_{T+1}(\theta)$ for all $\theta \in \mathbb{R}^{n_x \times n_z}$, implying monotonicity of $\{\inf_\theta W_T(\theta)\}_{T>0}$. To see boundedness, first notice that

$$W_T(\theta^\star) = \max_{0 \leq t \leq T-1} \|x_{t+1} - \theta^\star z_t\|_\infty = \max_{0 \leq t \leq T-1} \|w_t\|_\infty \leq w_{\max},$$

since $x_{t+1} = \theta^\star z_t + w_t$. Therefore, for any $T > 0$, we have that

$$\inf_\theta W_T(\theta) = \inf_\theta \max_{0 \le t \le T-1} \|x_{t+1} - \theta z_t\|_\infty \le \max_{0 \le t \le T-1} \|x_{t+1} - \theta^\star z_t\|_\infty \le w_{\max}$$

.

(*iii*.) First, we show that W_T attains its minimum on $\mathbb{R}^{n_x \times n_z}$. If $z_t = 0$ for $t \in [T]$, then W_T is a constant function and any $\theta \in \mathbb{R}^{n_x \times n_z}$ is a minimum of W_T. Now, suppose $z_t \ne 0$ for some $t \in [T]$. Then, W_T diverges at the infinity, i.e., $\lim_{k \to \infty} W_T(\theta_k) = \infty$ for any sequence $\{\theta_k\}_{k \in \mathbb{N}}$ such that $\|\theta_k\| \to \infty$ as $k \to \infty$. Since W_T is convex and bounded below with finite infimum, there exists a global minimizer $\bar{\theta}_T \in \mathbb{R}^{n_x \times n_z}$ such that $W_T(\bar{\theta}_T) = \inf_\theta W_T(\theta) = \bar{w}_{\max}^{(T)}$. Furthermore, by (*ii*), we have that $\|x_{t+1} - \bar{\theta}_T z_t\|_\infty \le w_{\max}$ for all $t \in [T]$ and any global minimizer $\bar{\theta}_T \in \arg\min_\theta W_T(\theta)$, hence $\bar{\theta}_T \in \Theta_T$ by definition. $\qquad\square$

Step 1 of the proof of Theorem 14: We first show that the error margin of the estimate $\bar{w}_{\max}^{(T)}$ from $w_{\max}$ is governed by the sum of two factors: (i) the diameter of Θ_T, which arises due to the lack of knowledge of θ^*, and (ii) the minimum learning error achievable if θ^* were known, namely

$$0 \le w_{\max} - \bar{w}_{\max}^{(T)} \le b_z \operatorname{diam}(\Theta_T) + w_{\max} - \max_{0 \le t \le T-1} \|w_t\|_\infty. \tag{4.15}$$

First, $0 \le w_{\max} - \bar{w}_{\max}^{(T)}$ is simply due to Lemma 21. Next, we prove the second inequality $w_{\max} - \bar{w}_{\max}^{(T)} \le b_z \operatorname{diam}(\Theta_T) + w_{\max} - \max_{0 \le t \le T-1} \|w_t\|_\infty$. By Lemma 21, there exists $\bar{\theta}_T \in \Theta_T$ such that $W_T(\bar{\theta}_T) = w_{\max}$ and

$$
\begin{aligned}
w_{\max} &= \max_{0 \le t \le T-1} \|x_{t+1} - \bar{\theta}_T z_t\|_\infty, \\
&= \max_{0 \le t \le T-1} \|x_{t+1} - \theta^\star z_t + (\theta^\star - \bar{\theta}_T) z_t\|_\infty, \\
&\ge \max_{0 \le t \le T-1} \left(\|x_{t+1} - \theta^\star z_t\|_\infty - \|(\theta^\star - \theta_k) z_t\|_\infty \right),
\end{aligned}
$$

where the inequality is due to reverse triangle inequality. Furthermore, by using the equivalence of ℓ_2 and ℓ_∞ norms, i.e., $\|x\|_2 \le \|x\|_\infty$ for $x \in \mathbb{R}^{n_x}$, we bound $w_{\max}$ further below by

$$
\begin{aligned}
w_{\max} &\ge \max_{0 \le t \le T-1} \left(\|x_{t+1} - \theta^\star z_t\|_\infty - \|(\theta^\star - \bar{\theta}_T) z_t\|_2 \right), \\
&\ge \max_{0 \le t \le T-1} \left(\|x_{t+1} - \theta^\star z_t\|_\infty - \|\theta^\star - \bar{\theta}_T\|_2 \|z_t\|_2 \right), \\
&\ge \max_{0 \le t \le T-1} \|w_t\|_\infty - b_z \operatorname{diam}(\Theta_T),
\end{aligned}
$$

where the second inequality is due to $\|\theta^\star - \bar{\theta}_T\|_2 := \sup_{z \neq 0} \frac{\|(\theta^\star - \bar{\theta}_T)z\|_2}{\|z\|_2} \leq \frac{\|(\theta^\star - \bar{\theta}_T)z_t\|_2}{\|z_t\|_2}$ and the third inequality follows from the assumption $\|z_t\|_2 \leq b_z$, the equivalence of Frobenius and spectral norms $\|\theta^\star - \bar{\theta}_T\|_2 \leq \|\theta^\star - \bar{\theta}_T\|_F$, and $\theta^\star, \bar{\theta}_T \in \Theta_T$. Consequently,

$$w_{\max} - \bar{w}_{\max}^{(T)} \leq w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty + b_z \operatorname{diam}(\Theta_T).$$

This completes the proof of the first step. □

Step 2 of the proof of Theorem 14: Using the learning error bound in (4.15), we obtain an upper bound on the probability of learning error exceeding a fixed $\delta > 0$ as shown below

$$\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) \leq \mathbb{T}_1 + \mathbb{T}_2\left(\frac{\delta}{2b_z}\right) + \mathbb{T}_5\left(\frac{\delta}{2}\right), \tag{4.16}$$

where $\mathbb{T}_5(\delta) := (1 - q_w(\delta))^T$.

First, using the the fact that $\{w_t\}_{t=0}^{T-1}$ are iid, we show that

$$\begin{aligned}
\mathbb{P}\left(w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta\right) &= \mathbb{P}(w_{\max} - \delta > \|w_t\|_\infty, \ \forall \, 0 \leq t \leq T-1), \\
&= \prod_{t=0}^{T-1} \mathbb{P}(w_{\max} - \delta > \|w_t\|_\infty), \\
&\leq \prod_{t=0}^{T-1} \mathbb{P}(w_{\max} - \delta > w_t^1), \\
&\leq (1 - q_w(\delta))^T,
\end{aligned}$$

where the first inequality is due to $w_t^1 \leq \|w_t\|_\infty$ and the second inequality is from Assumption 5. Finally, we obtain the desired convergence rate using the error bound in (4.15) as follows:

$$\begin{aligned}
\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) &\leq \mathbb{P}\left(b_z \operatorname{diam}(\Theta_T) + w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta\right) \\
&\leq \mathbb{P}\left(b_z \operatorname{diam}(\Theta_T) > \delta/2 \text{ or } w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta/2\right) \\
&\leq \mathbb{P}\left(\operatorname{diam}(\Theta_T) > \frac{\delta}{2b_z}\right) + \mathbb{P}\left(w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta/2\right) \\
&\leq \mathbb{T}_1 + \mathbb{T}_2\left(\frac{\delta}{2b_z}\right) + \mathbb{T}_5(\delta/2).
\end{aligned}$$

where the last inequality is by Theorem 12.

This completes the second and the last step of the proof. □

Proof of Corollary 14.1

First, by the proof of Corollary 12.1 in section 4.E, we have that $\mathbb{T}_1 = \frac{T}{m}\widetilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m) \leq \epsilon$ whenever $m \geq O(n_z + \log T + \log\frac{1}{\epsilon})$.

Next, we show $\mathbb{T}_5(\delta_T/2) \leq \mathbb{T}_2(\frac{\delta_T}{2b_z})$. Since $b_z \geq \sigma_z$ by the definition of BMSB, we have $\frac{a_1\delta_T}{8\sqrt{n_x}b_z} \leq \frac{\delta_T}{2}$. Since $q_w(\cdot)$ is a non-decreasing function, we have $1 - q_w(\frac{a_1\delta_T}{8\sqrt{n_x}b_z}) \geq 1 - q_w(\frac{\delta_T}{2})$. Notice that $m \geq 1$, and the constant factors in front of the $(1-q_w(\cdot))^{\lceil(T/m)\rceil}$ in $\mathbb{T}_2$ is also larger than 1. Consequently, $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \geq \mathbb{T}_5(\delta/2)$. Therefore, the choice of δ_T for the second term $\mathbb{T}_2$ also guarantees $\mathbb{T}_5(\delta_T/2) \leq \epsilon$.

Therefore, it suffices to ensure $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$. Notice that, when $\frac{\delta_T}{2b_z} = 2w_{\max}$, then $\mathbb{T}_2(\frac{\delta_T}{2b_z}) = 0 \leq \epsilon$, so there exists δ_T such that $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$.

Next, we will show that there exists such δ_T that diminishes to zero as T goes to infinity. Notice that we need

$$1 - q_w\left(\frac{a_1\delta_T}{8b_z\sqrt{n_x}}\right) \leq \left(\frac{\epsilon}{\widetilde{O}((n_x n_z)^{2.5}a_4^{n_x n_z})}\right)^{1/\lceil T/m\rceil},$$

so that

$$q_w\left(\frac{a_1\delta_T}{8b_z\sqrt{n_x}}\right) \geq 1 - \left(\frac{\epsilon}{\widetilde{O}((n_x n_z)^{2.5}a_4^{n_x n_z})}\right)^{1/\lceil T/m\rceil},$$

where the right hand side converges to zero as $T \to \infty$.

Now, consider $\delta(k) = 1/k$. Since $q_w\left(\frac{a_1\delta(k)}{8b_z\sqrt{n_x}}\right) > 0$, there exists a large enough T_k for any $k > 0$ such that for any $T \geq T_k$, we have that

$$q_w\left(\frac{a_1\delta(k)}{8b_z\sqrt{n_x}}\right) \geq 1 - \left(\frac{\epsilon}{\widetilde{O}((n_x n_z)^{2.5}a_4^{n_x n_z})}\right)^{1/\lceil T_k/m\rceil}.$$

Furthermore, for any $T > 0$, we can define

$$\delta_T = \begin{cases} \delta(k), & \text{if } T_k \leq T < T_{k+1}, \text{ for } k > 0, \\ 2w_{\max}, & \text{if } T < T_1. \end{cases}$$

In this way, δ_T satisfies $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$ and $\delta_T \to 0$ as $T \to +\infty$.

Finally, using the proof of Corollary 12.1, we can show that there exists $\frac{\delta_T}{2b_z} = \widetilde{O}(n_x^{1.5}n_z^2/T)$ such that $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$ whenever $q_w(\delta) = O(\delta)$. This implies $\delta_T = \widetilde{O}(n_x^{1.5}n_z^2/T)$ and completes the proof. $\qquad\square$

Proof of Theorem 15

We first show that the unknown $\theta^\star$ is a member of USC-SME uncertainty set $\widehat{\Theta}_T^{\text{ucb}}$ with high probability. By Theorem 14, Corollary 14.1, and the definition in (4.5), we have

$$\mathbb{P}(w_{\max} > \widehat{w}_{\max}^{(T)}) = \mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta_T) \le 3\epsilon,$$

which implies $1 - 3\epsilon \le \mathbb{P}(w_{\max} \le \widehat{w}_{\max}^{(T)}) \le \mathbb{P}(\theta^\star \in \widehat{\Theta}_T^{\text{ucb}})$.

Next, we show that the diameter of the UCB-SME uncertainty set is controlled by δ_T with high probability. Notice that $\widehat{\Theta}_T^{\text{ucb}} \subseteq \Theta_T(w_{\max} + \delta_T)$ because $\bar{w}_{\max}^{(T)} \le w_{\max}$. Therefore, by Theorem 13, the following holds for any constant $r > 0$:

$$\mathbb{P}(\text{diam}(\widehat{\Theta}_T^{\text{ucb}}) > r + a_5\sqrt{n_x}\delta_T) \le \mathbb{P}(\text{diam}(\Theta_T(w_{\max} + \delta_T)) > r + a_5\sqrt{n_x}\delta_T),$$
$$\le \mathbb{T}_1 + \mathbb{T}_2(r).$$

Let $r = \delta_T$, then, using the inequality $\mathbb{T}_2(\delta_T) \le \mathbb{T}_2(\delta_T/2b_z)$, we have that

$$\mathbb{P}(\text{diam}(\widehat{\Theta}_T^{\text{ucb}}) > \delta_T + a_5\sqrt{n_x}\delta_T) \le \mathbb{P}(\text{diam}(\Theta_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T),$$
$$\le 2\epsilon.$$

Therefore, with probability $1 - 2\epsilon$, the diameter of $\widehat{\Theta}_T^{\text{ucb}}$ is bounded above by

$$\text{diam}(\widehat{\Theta}_T^{\text{ucb}}) \le \delta_T + a_5\sqrt{n_x}\delta_T = O(\sqrt{n_x}\delta_T).$$

Finally, we can verify that the event $\{\text{diam}(\widehat{\Theta}_T^{\text{ucb}}) \le \delta_T + a_5\sqrt{n_x}\delta_T = O(\sqrt{n_x}\delta_T)\}$ and the event $\{\theta^\star \in \widehat{\Theta}_T^{\text{ucb}}\}$ simultaneously happen with probability at least $1 - 3\epsilon$ as follows:

$$\mathbb{P}\left(\theta^\star \notin \widehat{\Theta}_T(\widehat{w}_{\max}^{(T)}), \text{ or } \text{diam}(\widehat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T\right)$$

$$\le \mathbb{P}\left(w_{\max} - \max_{0 \le t \le T-1} \|w_t\|_\infty \ge \delta_T/2, \text{ or } \text{diam}(\Theta_T) > \delta_T/2b_z, \text{ or } \text{diam}(\widehat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T\right)$$

$$\le \epsilon + \mathbb{P}(\mathcal{E}_2) + \sum_{i=1}^{v_\gamma} \mathbb{P}\left(\bigcap_k G_{i,k}(\min(\delta_T/2b_z, \delta_T))\right)$$

$$\le 3\epsilon.$$

The third inequality follows from

- the proof of Theorem 14 in section 4.H,

- Theorem 13,

- the fact that the probabilties $\mathbb{P}(\text{diam}(\Theta_T) > \delta_T/2b_z)$ and $\mathbb{P}(\text{diam}(\widehat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T)$ are bounded by the same events $\mathcal{E}_2$,

- and $G_{i,k}(\delta_T), G_{i,k}(\delta_T/2b_z) \subseteq G_{i,k}(\min(\delta_T/2b_z, \delta_T))$, where $G_{i,k}(\delta)$ is defined in Lemma 7 as a function of δ.

This completes the proof. □

4.I Simulation Details

This section provides the details on the simulation experiments, along with some additional results. The code for replicating the presented results can be found in the github repository: https://github.com/jy-cds/non-asymptotic-set-membership.

Baseline: LSE Confidence Regions

In all our experiments, we use the 90% confidence region of the LSE as the baseline uncertainty set. The diameters of LSE's confidence regions are computed by taking minimum of the formulas provided in the following two papers: Lemma E.3 in [59] and Theorem 1 in [60]. To apply Theorem 1 in [60], we used regularization parameter $\lambda = 0.1$, $\delta = 0.1$ for 90% confidence, $S = \sqrt{\text{tr}(\theta^{\star,\top}\theta^\star)}$, variance proxy $L = 1$ for truncated Gaussian distribution and $L = 4/3$ for uniform distribution.

To determine the parameters in Lemma E.3 of [59], we approximately optimize the projection matrix P in Lemma E.3 as follows. First, we consider an orthogonal transformation of the empirical covariance matrix $\Lambda = \sum_{t=1}^{T} z_t z_t^\top$ with $\Lambda = GMG^\top$ where G is unitary. This transforms the event $\mathcal{E}$ in Lemma E.3 to $M \geq \lambda_1 P_0 + \lambda_2(I - P_0)$, where $GP_0G^\top = P$. We select P_0 as a block matrix $[[I_p, 0], [0, 0]]$, then optimize over the block size p in search of the tightest LSE confidence bound.

Figure 4.1: SME and LSE Uncertainty Set Visualization

In this experiment, we consider $x_{t+1} = A^\star x_t + B^\star u_t + w_t$, where $A^\star = 0.8$ and $B^\star = 1$ are unknown. $w_t \sim \text{TrunGauss}(0, \sigma_w, [-w_{\max}, w_{\max}])$ is i.i.d. and $u_t \sim \text{TrunGauss}(0, \sigma_u, [-u_{\max}, u_{\max}])$ are also i.i.d generated, where $\sigma_w = \sigma_u = 0.5$, and $w_{\max} = u_{\max} = 1$. We compare SME that knows $w_{\max} = 1$ and LSE's 90% confidence region computed based on Section 4.I.

Figure 4.2

In this experiment, we consider the the linearized longitudinal flight control dynamics of Boeing 747 [17], [130] with i.i.d. bounded inputs and disturbances sampled from

truncated Gaussian and uniform distribution. The dynamics is $x_{t+1} = A^\star x_t + B^\star u_t + w_t$ with

$$A = \begin{bmatrix} 0.99 & 0.03 & -0.02 & -0.32 \\ 0.01 & 0.47 & 4.7 & 0 \\ 0.02 & -0.06 & 0.4 & 0 \\ 0.01 & -0.04 & 0.72 & 0.99 \end{bmatrix} \qquad B = \begin{bmatrix} 0.01 & 0.99 \\ -3.44 & 1.66 \\ -0.83 & 0.44 \\ -0.47 & 0.25 \end{bmatrix}.$$

Disturbances are sampled from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^4)$ and $\text{Unif}([-w_{\max}, w_{\max}]^4)$, while control inputs are samples from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^2)$ in both disturbance settings, with $w_{\max} = 2$. To compute the UCB for SME using (4.5), we heuristically define $\delta_T = \beta \frac{n_x^{1.5} \cdot n_z^2 \cdot (\max_t \|x_t\|)}{T}$, where $n_x = 4$ and $n_z = 6$ are the system dimension, while β is a tunable parameter. This definition matches the dimension and time order of the theoretical analysis in Corollary 14.1. In both experiments of Figure 4.2, we fix $\beta = 0.01$.

In Figure 2(a)-(b), we plot SME with accurate and conservative bounds of $w_{\max}$, UCB-SME, and LSE's 90% confidence regions computed by Section 4.I. We use 10 different seeds to generate the disturbance sequences for each plot, and use the shaded region to denote 1 standard deviation from the mean (colored lines).

Figure 4.3

In this experiment, we consider autonomous systems of the form $x_{t+1} = A^\star x_t + w_t$, where $A^\star \in \mathbb{R}^{n_x}$ is randomly sampled and its spectral radius is normalized to be 0.9. We simulate SME and LSE for $n_x = 5, 10, 15, 20, 25$. The disturbances are sampled from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^{n_x})$ as well as $\text{Unif}([-w_{\max}, w_{\max}]^{n_x})$ with $w_{\max} = 2$. This simulation is run on 10 random seeds and the total length of the simulation is set to be $T = 1000$ across all n_x experiments. The mean is plotted as solid lines and the shaded regions denote 1 standard deviation from the mean.

Though SME's theoretical bound with respect to the dimension is $\widetilde{O}(n_x^{1.5} n_z^2)$ from Corollary 12.2, which is much worse than LSE's bound, it is not reflected in Figure 4.3. Therefore, it is promising that the dimension scaling in the analysis in Section 4.3 can be further tightened. We leave this for future work.

Figure 4.4

To illustrate the quantitative impact of using SME for adaptive tube-based robust MPC, we study tube-based robust MPC for a system $x_{t+1} = A^\star x_t + B^\star u_t + w_t$ with nominal system $A^\star = 1.2$, $B^\star = 0.9$ with an initial model uncertainty set

$\Theta_0 := [1, 1.2] \times [0.9, 1.1]$. We use the basic tube-based robust MPC method [131], [132] and parameterize the control policy as $u_k = Kx_k + v_k + \eta_k$, where $K = -1$, v_k is determined by the tube-based robust MPC algorithm, and η_k is a bounded exploration injection with $\eta_k \sim \text{Unif}([-0.01, 0.01])$. The disturbance w_k has a known bound of $w_{\max} = 0.1$ and is generated to be i.i.d. $\text{Unif}([-0.1, 0.1])$. The horizon of the tube-based robust MPC is set to be 5. The state and input constraints are such that $x_k \in [-10, 10]$ and $u_k \in [-10, 10]$ for all $k \geq 0$. We consider the task of constrained LQ tracking problem with a time-varying cost function $c_t := (x_t - g_t)^\top Q(x_t - g_t) + u_t^\top R u_t$ where the target trajectory is generated as $g_t = 8\sin(t/20)$.

We compare the performance of an adaptive tube-based robust MPC controller that uses the SME for uncertainty set estimation against one that uses the LSE 90% confidence region (LSE). For better visualization of the trajectory difference as a result of different estimation methods, we used the minimum of the the dominant factors in Dean, Mania, Matni, *et al.* [58, equation C.12] and the LSE 90% confidence region for the LSE uncertainty set. We also plot the offline optimal RMPC controller, i.e., the controller that has knowledge of the true underlying system parameters (OPT).

Since the controller has to robustly satisfy constraints against the worst-case model in the uncertainty set, smaller uncertainty set for the tube-based robust MPC means more optimal trajectories can be computed. This observation is consistent with the extensive empirical results in the control literature [49], [117], [118].

C h a p t e r   5

ONLINE ADVERSARIAL STABILIZATION OF UNKNOWN TIME-VARYING SYSTEMS

In the previous chapter, we introduced SME, an uncertainty set estimation method. We demonstrated that for linear control dynamical systems, the uncertainty sets generated by SME converge linearly (with respect to the number of samples) to the true model under stochastic perturbations. Importantly, even when the stochastic assumptions in Chapter 4 fail to hold, SME still guarantees that the uncertainty sets contain the true model, provided we have an upper bound on the disturbances. This makes SME an ideal method for estimating uncertainty sets in both stochastic and non-stochastic settings when perturbations are bounded.

Recall that in Chapter 1 (Figure 1.1), we introduced a learning-based control framework that leverages SME uncertainty sets to integrate learning algorithms with model-based control design. Thanks to the robustness of SME as discussed above, the framework can be used to guarantee safety of the closed loop system despite learning under potentially adversarial disturbances.

In this chapter, we will study the canonical problem of online stabilization of *unknown linear time-varying* (LTV) system under bounded non-stochastic (potentially adversarial) disturbances. The study of LTV systems are crucial for CPS such as sustainable energy systems. For instance, as power systems continue to adopt more renewable energy supply, the system dynamics for frequency regulation will become time-varying due to the intermittency of the renewables. Moreover, LTV system can be used to approximate practical applications where system dynamics are often nonlinear.

To address this, we instantiate the framework introduced in Chapter 1 and propose a novel algorithm based on convex body chasing (CBC), an online learning technique, and classical linear qudratic regulator (LQR) control. Assuming infrequently changing or slowly drifting dynamics, our algorithm guarantees bounded-input-bounded-output stability for the unknown LTV system. Our approach avoids system identification and requires minimal disturbance assumptions. This chapter is based on the following paper:

[1] J. Yu, V. Gupta, and A. Wierman, "Online adversarial stabilization of unknown linear time-varying systems," *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 8320–8327, 2023. DOI: 10.1109/CDC49753.2023.1038384.

5.1 Introduction

Learning-based control of linear-time invariant (LTI) systems in the context of linear quadratic regulators (LQR) has seen considerable progress. However, many real-world systems are time-varying in nature. For example, the grid topology in power systems can change over time due to manual operations or unpredictable line failures [66]. Therefore, there is increasing recent interest in extending learning-based control of LTI systems to the linear time-varying (LTV) setting [162]–[166].

LTV systems are widely used to approximate and model real-world dynamical systems such as robotics [167] and autonomous vehicles [168]. In this chapter, we consider LTV systems with dynamics of the following form:

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \tag{5.1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ and w_t denotes the state, the control input, and the bounded and potentially adversarial disturbance, respectively. We use $\theta_t = [A_t \ B_t]$ to succinctly denote the system matrices at time step t.

On the one hand, **offline control** design for LTV systems is well-established in the setting where the underlying LTV model is *known* [169]–[173]. Additionally, recent work has started focusing on regret analysis and non-stochastic disturbances for known LTV systems [162], [174].

On the other hand, **online control** design for LTV systems where the model is *unknown* is more challenging. Historically, there is a rich body of work on adaptive control design for LTV systems [175]–[177]. Also related is the system identification literature for LTV systems [178]–[180], which estimates the (generally assumed to be stable) system to allow the application of the offline techniques.

In recent years, the potential to leverage modern data-driven techniques for controller design of unknown linear systems has led to a resurgence of work in both the LTI and LTV settings. There is a growing literature on "learning to control" unknown LTI systems under stochastic or no noise [17], [24], [181]. Learning under bounded and potentially adversarial noises poses additional challenges, but online stabilization [99] and regret [26] results have been obtained.

In comparison, there is much less work on learning-based control design for unknown LTV systems. One typical approach, exemplified by [163], [182], [183], derives stabilizing controllers under the assumption that *offline* data representing the input-output behavior of (5.1) is available and therefore an *offline* stabilizing controller can be pre-computed. Similar *finite-horizon* settings where the algorithm has access to offline data [184], or can *iteratively* collect data [185] were also considered. In the context of *online* stabilization, i.e., when offline data is not available, work has derived stabilizing controllers for LTV systems through the use of predictions of θ_t, e.g., [18]. Finally, another line of work focuses on designing regret-optimal controllers for LTV systems [164]–[166], [186], [187]. However, with the exception of [18], existing work on *online* control of unknown LTV systems share the common assumption of either of open-loop stability or knowledge of an offline stabilizing controller. Moreover, the disturbances are generally assumed to be zero or stochastic noise independent of the states and inputs.

In this chapter, we propose an online algorithm for stabilizing unknown LTV systems under bounded, potentially adversarial disturbances. Our approach uses convex body chasing (CBC), which is an online learning problem where one must choose a sequence of points within sequentially presented convex sets with the aim of minimizing the sum of distances between the chosen points [188], [189]. CBC has emerged as a promising tool in online control, with most work making connections to a special case called *nested* convex body chasing (NCBC), where the convex sets are sequentially nested within the previous set [190], [191]. In particular, [13] first explored the use of NCBC for learning-based control of time-invariant nonlinear systems. NCBC was also used in combination with System Level Synthesis to design a distributed controller for networked systems [99] and in combination with model predictive control [55] for LTI system control as a promising alternative to system identification based methods. However, this line of work depends fundamentally on the time invariance of the system, which results in nested convex sets. LTV systems do not yield nested sets and therefore represent a significant challenge.

This work addresses this challenge and presents a novel online control scheme (Algorithm 2) based on CBC (non-nested) techniques that guarantees bounded-input-bounded-output (BIBO) stability as a function of the total model variation $\sum_{t=1}^{\infty} \|\theta_t - \theta_{t-1}\|$, without predictions or offline data under bounded and potentially adversarial disturbances for unknown LTV systems (Theorem 18). This result implies that when the total model variation is finite or growing sublinearly, BIBO

stability of the closed loop is guaranteed (Corollaries 18.1 and 18.2). In particular, our result depends on a refined analysis of the CBC technique (Lemma 22) and is based on the perturbation analysis of the Lyapunov equation. This contrasts with previous NCBC-based works for time-invariant systems, where the competitive ratio guarantee of NCBC directly applies and the main technical tool is the robustness of the model-based controller, which is a proven using a Lipschitz bound of a quadratic program in [99] and is directly assumed to exist in [13].

We illustrate the proposed algorithm via numerical examples in Section 5.4 to corroborate the stability guarantees. We demonstrate how the proposed algorithm can be used for data collection and complement data-driven methods like [163], [183], [184]. Further, the numerics highlight that the proposed algorithm can be efficiently implemented by leveraging the linearity of (5.1) despite the computational complexity of CBC algorithms in general (see Line 5 for details).

Notation. We use $\mathbb{S}^{n-1}$ to denote the unit sphere in $\mathbb{R}^n$ and $\mathbb{N}_+$ for positive integers. For $t, s \in \mathbb{N}_+$, we use $[t:s]$ as shorthand for the set of integers $\{t, t+1, \ldots, s\}$ and $[t]$ for $\{1, 2, \ldots, t\}$. Unless otherwise specified, $\|\cdot\|$ is the operator norm. We use $\rho(\cdot)$ for the spectral radius of a matrix.

5.2 Preliminaries

In this section, we state the model assumptions underlying our work and review key results for convex body chasing, which we leverage in our algorithm design and analysis.

Stability and model assumptions

We study the dynamics in (5.1) and make the following standard assumptions about the dynamics.

Assumption 6. *The disturbances are bounded:* $\|w_t\|_\infty \leq W$ *for all* $t \geq 0$.

Assumption 7. *The unknown time-varying system matrices* $\{\theta_t\}_{t=0}^\infty$ *belong to a known (potentially large) polytope* Θ *such that* $\theta_t \in \Theta$ *for all* t. *Moreover, there exists* $\kappa > 0$ *such that* $\|\theta\| \leq \kappa$ *and* θ *is stabilizable for all* $\theta \in \Theta$.

Bounded and non-stochastic (potentially adversarial) disturbances is a common model both in the online learning and control problems [115], [192]. Since we make no assumptions on how large the bound W is, Assumption 6 models a variety of scenarios, such as bounded and/or correlated stochastic noise, state-dependent

disturbances, e.g., the linearization and discretization error for nonlinear continuous-time dynamics, and potentially adversarial disturbances. Assumption 7 is standard in learning-based control, e.g. [23], [193].

We additionally assume there is a quadratic known cost function of the state and control input at every time step t to be minimized, e.g. $x_t^\top Q x_t + u_t^\top R u_t$, with $Q, R > 0$. For a given LTI system model $\theta = [A \; B]$ and cost matrices Q, R, we denote $K = \text{LQR}(\theta; Q, R)$ as the optimal feedback gain for the corresponding infinite-horizon LQR problem.

Remark 4. *Representing model uncertainty as convex compact parameter sets where every model is stabilizable is not always possible. In particular, if a parameter set Θ has a few singular points where (A, B) loses stabilizability such as when $B = 0$, a simple heuristic is to ignore these points in the algorithm since we assume the underlying true system matrices θ_t must be stabilizable.*

Convex body chasing

Convex Body Chasing (CBC) is a well-studied online learning problem [190], [191]. At every round $t \in \mathbb{N}_+$, the player is presented a convex body/set $\mathcal{K}_t \subset \mathbb{R}^n$. The player selects a point $q_t \in \mathcal{K}_t$ with the objective of minimizing the cost defined as the total path length of the selection for T rounds, e.g., $\sum_{t=1}^{T} \|q_t - q_{t-1}\|$ for a given initial condition $q_0 \notin \mathcal{K}_1$. There are many known algorithms for the CBC problem with a *competitive ratio* guarantee such that the cost incurred by the algorithm is at most a constant factor from the total path length incurred by the offline optimal algorithm which has the knowledge of the entire sequence of the bodies. We will use CBC to select θ_t's that are consistent with observed data.

The nested case

A special case of CBC is the *nested* convex body chasing (NCBC) problem, where $\mathcal{K}_t \subseteq \mathcal{K}_{t-1}$. A known algorithm for NCBC is to select the *Steiner point* of $\mathcal{K}_t$ at t [191]. The Steiner point of a convex set $\mathcal{K}$ can be interpreted as the average of the extreme points of $\mathcal{K}$ and is defined as $\text{st}(\mathcal{K}) := \mathbb{E}_{v:\|v\| \leq 1}[g_{\mathcal{K}}(v)]$, where $g_{\mathcal{K}}(v) := \text{argmax}_{x \in \mathcal{K}} v^\top x$ and the expectation is taken with respect to the uniform distribution over the unit ball. The intuition is that Steiner point remains "deep" inside of the (nested) feasible region so that when this point becomes infeasible due to a new convex set, this convex set must shrink considerably, which indicates that the offline optimal must have moved a lot. Given the initial condition $q_0 \notin \mathcal{K}_1$, the

Steiner point selector achieves competitive ratio of $O(n)$ against the offline optimal such that for all $T \in \mathbb{N}_+$, $\sum_{t=1}^{T} \|\mathsf{st}(\mathcal{K}_t) - \mathsf{st}(\mathcal{K}_{t-1})\| \leq O(n) \cdot \text{OPT}$, where OPT is the offline optimal total path length. There are many works that combine the Steiner point algorithm for NCBC with existing control methods to perform learning-based online control for LTI systems, e.g., [13], [55], [99].

General CBC

For general CBC problems, we can no longer take advantage of the nested property of the convex bodies. One may consider naively applying NCBC algorithms when the convex bodies happen to be nested and restarting the NCBC algorithm when they are not. However, due to the myopic nature of NCBC algorithms, which try to remain deep inside of each convex set, they no longer guarantee a competitive ratio when used this way. Instead, [188] generalizes ideas from NCBC and proposes an algorithm that selects the *functional Steiner point* of the *work function*.

Definition 5.2.1 (Functional Steiner point). *For a convex function $f : \mathbb{R}^n \to \mathbb{R}$, the functional Steiner point of f is*

$$\mathsf{st}(f) = -n \cdot \fint_{v:\|v\|=1} f^*(v) \, v \, dv, \tag{5.2}$$

where $\fint_{x \in \mathcal{S}} f(x)dx$ denotes the normalized value $\frac{\int_{x \in \mathcal{S}} f(x)dx}{\int_{x \in \mathcal{S}} 1 dx}$ of $f(x)$ on the set $\mathcal{S}$, and

$$f^*(v) := \inf_{x \in \mathbb{R}^n} f(x) - \langle x, v \rangle \tag{5.3}$$

is the Fenchel conjugate of f.

The CBC algorithm selects the functional Steiner point of the *work function*, which records the smallest cost required to satisfy a sequence of requests while ending in a given state, thereby encapsulating information about the offline-optimal cost for the CBC problem.

Definition 5.2.2 (Work function). *Given an initial point $q_0 \in \mathbb{R}^n$, and convex sets $\mathcal{K}_1, \ldots, \mathcal{K}_t \subset \mathbb{R}^n$, the work function at time step t evaluated at a point $x \in \mathbb{R}^n$ is given by:*

$$\omega_t(x) = \min_{q_s \in \mathcal{K}_s} \|x - q_t\| + \sum_{s=1}^{t} \|q_s - q_{s-1}\|. \tag{5.4}$$

Importantly, it is shown that the functional Steiner points of the work functions are valid, i.e., $\mathsf{st}(\omega_t) \in \mathcal{K}_t$ for all t [188]. On a high level, selecting the functional Steiner

point of the work function helps the algorithm stay competitive against the currently estimated offline optimal cost via the work function, resulting in a competitive ratio of *n* against the offline optimal cost (OPT) for general CBC problems,

$$\sum_{t=1}^{T} \|\mathsf{st}(\omega_t) - \mathsf{st}(\omega_{t+1})\| \leq n \cdot \text{OPT}. \tag{5.5}$$

Given the non-convex nature of (6.3) and (5.4), we note that, in general, it is challenging to compute the functional Steiner point of the work function. However, in the proposed algorithm, we are able to leverage the linearity of the LTV systems and numerically approximate both objects with efficient computation in Line 5.

5.3 Main Results

We present our proposed online control algorithm to stabilize the unknown LTV system (5.1) under bounded and potentially adversarial disturbances in Algorithm 2. After observing the latest transition from x_t, u_t to x_{t+1} at $t + 1$ according to (5.1) (line 2), the algorithm constructs the set of all feasible models $\widehat{\theta}_t$'s (line 3) such that the model is *consistent* with the observation, i.e., there exists an admissible disturbance $\widehat{w}_t$ satisfying Assumption 6 such that the state transition from x_t, u_t to x_{t+1} can be explained by the tuple $(\widehat{\theta}_t, \widehat{w}_t)$. We call this set the *consistent model set* $\mathcal{P}_t$ and we note that the unknown true dynamics $\theta_t = [A_t \ B_t]$ belongs to $\mathcal{P}_t$. The algorithm then selects a *hypothesis* model out of the consistent model set $\mathcal{P}_t$ using the CBC algorithm by computing the functional Steiner point (6.3) of the work function (5.4) with respect to the history of the consistent parameter sets $\mathcal{P}_1, \ldots, \mathcal{P}_t$ (line 4). In particular, we present an efficient implementation of the functional Steiner point chasing algorithm in Line 5 by taking advantage of the fact that $\mathcal{P}_t$'s are polytopes that can be described by intersection of half-spaces. The implementation is summarized in Algorithm 3. Based on the selected hypothesis model $\widehat{\theta}_t$, a certainty-equivalent LQR controller is synthesized (line 5) and the state-feedback control action is computed (line 6).

Note that, by construction, at time step $t \in \mathbb{N}_+$ we perform certainty-equivalent control $\widehat{K}_{t-1}$ based on a hypothesis model $\widehat{\theta}_{t-1}$ computed using retrospective data, even though the control action ($u_t = \widehat{K}_{t-1} x_t$) is applied to the dynamics (θ_t) that we do not yet have any information about. In order to guarantee stability, we would like for $\widehat{K}_{t-1}$ to be stabilizing the "future" dynamics (θ_t). This is the main motivation behind our choice of the CBC technique instead of regression-based techniques for model selection. Thanks to the competitive ratio guarantee (5.5) of the functional Steiner

point selector, when the true model variation is "small," our previously selected hypothesis model will stay "consistent" in the sense that $\widehat{K}_{t-1}$ can be stabilizing for θ_t despite the potentially adversarial or state-dependent disturbances. On the other hand, when the true model variation is "large," $\widehat{K}_{t-1}$ does not stabilize θ_t, and we see growth in the state norm. Therefore, our final state bound is in terms of the total variation of the true model.

We show in the next section that, by drawing connections between the stability of the closed-loop system and the path length cost of the selected hypothesis model via CBC, we are able to stabilize the unknown LTV system without any identification requirements, e.g., the selected hypothesis models in Algorithm 2 need not be close to the true models. It is observed that even in the LTI setting, system identification can result in large-norm transient behaviors with numerical stability issues if the underlying unknown system is open-loop unstable or under non-stochastic disturbances, thus motivating the development of NCBC-based online control methods [13], [26], [99]. In the LTV setting, it is not sufficient to use NCBC ideas due to the time-variation of the model; however, the intuition for the use of CBC is similar. In fact, it can be additionally beneficial to bypass identification in settings where the true model is a moving target, thus making identification more challenging. We illustrate this numerically in Section 5.4.

Algorithm 2: UNKNOWN LTV STABILIZATION

Input: $W > 0, \Theta \subset \mathbb{R}^{n \times (n+m)}$
Initialize : $u_0 = 0, \widehat{\theta}_0 \in \Theta$
1 **for** $t + 1 = 1, 2, \ldots$ **do**
2 $\quad$ Observe x_{t+1}
3 $\quad$ Construct consistent set
$\quad\quad \mathcal{P}_t := \{\theta = [A, B] : \|x_{t+1} - Ax_t - Bu_t\|_\infty \le W\} \cap \Theta$
4 $\quad$ Select hypothesis model $\widehat{\theta}_t \leftarrow \text{CBC}(\{\mathcal{P}_s\}_{s=1}^t; \widehat{\theta}_0)$
5 $\quad$ Synthesize controller $\widehat{K}_t \leftarrow \text{LQR}\left(\widehat{\theta}_t; Q, R\right)$
6 $\quad$ Compute feedback control input $u_{t+1} = \widehat{K}_t x_{t+1}$
7 **end**

Stability Analysis

The main result of this paper is the BIBO stability guarantee for Algorithm 2 in terms of the true model variation and the disturbance bound. We sketch the proof in this section and refer Section 5.C for the formal proof. This result depends on a refined

---
Algorithm 3: CBC

Input: $\mathcal{P}_1, \ldots, \mathcal{P}_t, \widehat{\theta}_0, N$
Output: $\widehat{\theta}_t$
1 **for** $k = 0, 1, \ldots N$ **do**
2 $\quad$ Sample v_i uniformly from $\mathbb{S}^{n-1}$
3 $\quad h_i \leftarrow$ (5.12)
4 **end**
5 $\widehat{\theta}_t \leftarrow \text{proj}_{\Theta \cap \mathcal{P}_t} \left(-\frac{n}{N} \sum_{i=1}^{N} h_i v_i \right)$

---

analysis of the competitive ratio for the functional Steiner point chasing algorithm introduced in [188], which is stated as follows.

Lemma 22 (Partial-path competitive ratio). *For $t \in \mathbb{N}_+$, let s, $e \in [t]$ and $s < e$, and let $\Theta \subset \mathbb{R}^n$ be a convex compact set. Denote $\widehat{\Delta}_{[s,e]} := \sum_{\tau=s+1}^{e} \|\text{st}(\omega_\tau) - \text{st}(\omega_{\tau-1})\|_F$ as the partial-path cost of the functional Steiner point selector during interval $[s, e]$ and $\{OPT_\tau\}_{\tau=1}^{t}$ as the (overall) offline optimal selection for $\mathcal{K}_1, \ldots, \mathcal{K}_t \subset \Theta$. The functional Steiner point chasing algorithm has the following competitive ratio,*

$$\widehat{\Delta}_{[s,e]} \leq n \left(\text{dia}(\Theta) + 2\kappa + \sum_{\tau=s+1}^{e} \|OPT_\tau - OPT_{\tau-1}\|_F \right)$$

on interval $[s, e]$, where $\text{dia}(\Theta) := \max_{\theta_1, \theta_2 \in \Theta} \|\theta_1 - \theta_2\|_F$ denotes the diameter of Θ and $\kappa := \max_{\theta \in \Theta} \|\theta\|_F$.

Proof. See Section 5.A. $\qquad \square$

Theorem 18 (BIBO Stability). *Under Assumption 6 and 7, the closed loop of (5.1) under Algorithm 2 is BIBO stable such that for all $t \geq 0$,*

$$\|x_t\| \leq W \cdot c_1 \sum_{s=0}^{t-2} c_2^{\Delta_{[s,t-1]}} \rho_L^{t-s}$$

where $\Delta_{[s,t-1]} := \sum_{\tau=s+1}^{t-1} \|\theta_\tau - \theta_{\tau-1}\|_F$ is the true model variation, W is the disturbance bound, and c_1, $c_2 > 0$, $\rho_L \in (0, 1)$ are constants that depend on the system-theoretical quantities of the worst-case model in the parameter set Θ.

Proof Sketch: At a high level, the structure of our proof is as follows. We first use the fact that our time-varying feedback gain $\widehat{K}_t$ is computed according to a hypothesis model from the *consistent* model set. Therefore, we can characterize the closed-loop

dynamics in terms of the consistent models $\widehat{\theta}_t$ and $\widehat{K}_t$. Specifically, consider a time step t where we take the action $u_t = \widehat{K}_{t-1} x_t$ after observing x_t. Then, we observe $x_{t+1} = A_t x_t + B_t u_t + w_t$ and select a new hypothesis model $\widehat{\theta}_t = [\widehat{A}_t \ \widehat{B}_t]$ that is consistent with this new observation. Since we have selected a consistent hypothesis model, there is some admissible disturbance $\widehat{w}_t$ satisfying Assumption 6 such that

$$x_{t+1} = \left(A_t + B_t \widehat{K}_{t-1}\right) x_t + w_t = \left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}\right) x_t + \widehat{w}_t.$$

Without loss of generality, we assume initial condition $x_0 = 0$. We therefore have

$$x_t = \widehat{w}_{t-1} + \sum_{s=0}^{t-2} \prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right) \widehat{w}_s. \tag{5.6}$$

We have two main challenges in bounding $\|x_t\|$ in (5.6):

1. $\widehat{K}_t$ is computed using $\widehat{\theta}_t$ in Algorithm 2, but is applied to the next time step $\widehat{\theta}_{t+1}$. While we know $\rho(\widehat{A}_t + \widehat{B}_t \widehat{K}_t) < 1$, in (5.6) we have $\widehat{K}_{t-1}$ instead of $\widehat{K}_t$.

2. Naively applying submultiplicativity of the operator norm for (5.6) results in bounding $\left\|\left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right)\right\|$. However, even if $\widehat{K}_{t-1}$ satisfies $\rho(\widehat{A}_t + \widehat{B}_t \widehat{K}_t) < 1$, in general the operator norm can be greater than 1.

To address the first challenge, our key insight is that by selecting hypothesis models via CBC technique, in any interval where the true model variation is small, our selected hypothesis model also vary little. Specifically, by Lemma 22, we can bound the partial-path variation of the selected hypothesis models with the true model partial-path variation $\Delta_{[s,e]}$ as follows:

$$\widehat{\Delta}_{[s,e]} \le n \left(\text{dia}(\Theta) + 2\kappa + \sum_{\tau=s}^{e-1} \|\text{OPT}_{\tau+1} - \text{OPT}_\tau\|_F\right)$$
$$\le n \left(\text{dia}(\Theta) + 2\kappa + \Delta_{[s,e]}\right), \tag{5.7}$$

where Θ and κ are from Assumption 7. A consequence of (5.7) is that, during intervals where the true model variation is small, we have $\left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}\right) \approx \left(\widehat{A}_t + \widehat{B}_t \widehat{K}_t\right)$.

For the second challenge, we leverage the concept of sequential strong stability [21], which allows bounding $\left\|\prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right)\right\|$ approximately with $\prod_{\tau \in [t-1:s+1]} \rho\left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_\tau\right)$ times $O\left(\exp(\Delta_{[s,t-1]})\right)$.

We now sketch the proof. The helper lemmas are summarized in Section 5.B and the formal proof can be found in Section 5.C. Consider L_t, $H_t \in \mathbb{R}^{n \times n}$ with $H_t > 0$ such that

$$\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1} := H_t^{1/2} L_t H_t^{-1/2}.$$

We use I_s as shorthand for the interval $[t - 1 : s + 1]$. Then each summand in (5.6) can be bounded as

$$\left\| \prod_{\tau \in I_s} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1} \right) \right\|$$
$$\leq \underbrace{\left\| H_{t-1}^{1/2} \right\| \left\| H_{s+1}^{-1/2} \right\|}_{(a)} \underbrace{\prod_{k \in I_{s+1}} \left\| H_k^{-1/2} H_{k-1}^{1/2} \right\|}_{(b)} \underbrace{\prod_{\tau \in I_s} \| L_\tau \|}_{(c)} . \tag{5.8}$$

Therefore showing BIBO stability comes down to bounding individual terms in (5.8). In particular we will show that by selecting appropriate H_t and L_t, term (a) is bounded by a constant C_H that depends on system theoretical properties of the worst-case parameter in Θ. For (b) and (c), we isolate the instances when

$$\left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F \leq \epsilon \tag{5.9}$$

for some chosen $\epsilon > 0$. For instances where (5.9) holds, we use the perturbation analysis of the Lyapunov equation involving the matrix $\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}$ (Lemma 27 for (b) and Lemma 25 for (c)) to bound (b) and (c) in terms of the partial-path movement of the selected parameters $\widehat{\Delta}_{[s,e]} := \sum_{\tau=s+1}^{e} \| \mathsf{st}(\omega_{\tau+1}) - \mathsf{st}(\omega_\tau) \|_F$. Specifically, Lemma 27 implies

$$\left\| H_t^{-1/2} H_{t-1}^{1/2} \right\| \leq \begin{cases} e^{\frac{\beta \left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F}{2}}, & \text{if (5.9) holds} \\ \bar{H} & \text{otherwise,} \end{cases} \tag{5.10}$$

where β, $\bar{H} > 1$ are constants. We also show that from Lemma 25,

$$\| L_t \| \leq \begin{cases} \rho_L & \text{if (5.9) holds} \\ \bar{L} & \text{otherwise,} \end{cases} \tag{5.11}$$

for $\rho_L \in (0, 1)$ and $\bar{L} > 1$ a constant.

We now plug (5.10) and (5.11) into (5.8). Denote by $n_{[s,t]}$ the number of pairs $(\tau, \tau-1)$ with $s + 1 \leq \tau \leq t - 1$ where (5.9) fails to hold. Let $\Delta_{[s,e]} := \sum_{\tau=s+1}^{e} \| \theta_\tau - \theta_{\tau-1} \|_F$

be the true model partial-path variation. Then (5.8) can be bounded as

$$\left\| \prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1} \right) \right\|$$

$$\leq C_H \cdot \bar{H}^{n_{[s,t]}} \cdot e^{\frac{\beta \widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \bar{L}^{n_{[s,t]}} \cdot \rho_L^{t-s-\widehat{n}_{[s,t]}-1}$$

$$\leq C_H \left(\frac{\bar{L}\bar{H}}{\rho_L} \right)^{\frac{\widehat{\Delta}_{[s,t-1]}}{\epsilon_*}} e^{\frac{\beta \widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \rho_L^{t-s-1}$$

$$\leq C_H \left(\frac{\bar{L}\bar{H}}{\rho_L} \right)^{\frac{\bar{n}\left(\mathrm{dia}(\Theta) + 2\kappa + \Delta_{[s,t-1]} \right)}{\epsilon_*}} e^{\frac{\beta \bar{n} \left(\mathrm{dia}(\Theta) + 2\kappa + \Delta_{[s+1,t-1]} \right)}{2}} \cdot \rho_L^{t-s-1}$$

$$=: c \cdot c_2^{\Delta_{[s,t-1]}} \rho_L^{t-s},$$

for constants c, c_2 and $\bar{n} := n(n+m)$ for the dimension of the parameter space for A_t, B_t. In the second inequality, we used the observation that $n_{[s,t]} \leq \frac{\widehat{\Delta}_{[s,t-1]}}{\epsilon}$ and in the last inequality we used Lemma 22. Combined with (5.6) and Assumption 6, this proves the desired bound. ∎

An immediate consequence of Theorem 18 is that when the model variation in (5.1) is bounded or sublinear, Algorithm 2 guarantees BIBO stability. This is summarized below.

Corollary 18.1 (Bounded variation). *Suppose* (5.1) *has model variation* $\Delta_{[0,t]} \leq M$ *for a constant M. Then,*

$$\sup_t \|x_t\| \leq \frac{c_1 \cdot c_2^M}{1 - \rho_L}.$$

Corollary 18.2 (Unbounded but sublinear variation). *Let $\alpha \in (0,1)$ and $t \in \mathbb{N}_+$. Suppose* (5.1) *is such that for each $k \leq t$, $\Delta_{[k,k+1]} \leq \delta_t := 1/t^{(1-\alpha)}$, implying a total model variation $\Delta_{[0,t]} = O(t^\alpha)$. Then for large enough t, $\rho_L c_2^{\delta_t} \leq \frac{1+\rho_L}{2}$, and therefore*

$$\|x_k\| \leq c_1 \sum_{i=0}^{k} \left(\rho_L c_2^{\delta_t} \right)^i \leq \frac{2c_1}{1 - \rho_L}.$$

Corollary 18.1 can be useful for scenarios where the mode of operation of the system changes infrequently and for systems such that $\theta(t) \to \theta^\star$ as $t \to \infty$ [194]. As an example, consider power systems where a prescribed set of lines can potentially become disconnected from the grid and thus change the grid topology. Corollary 18.2 applies to slowly drifting systems [195].

Efficient implementation of CBC

In general, implementation of the functional Steiner point of the work function may be computationally inefficient. However, by taking advantage of the LTV structure, we are able to design an efficient implementation in our setting. The key observation here is that for each t, $\mathcal{P}_t$ (Algorithm 2, line 3) can be described by the intersection of half-spaces because the ambient parameter space Θ is assumed to be a polytope and the observed online transition data from x_t, u_t to x_{t+1} specifies two half-space constraints at each time step due to linearity of (5.1). Our approach to approximate the functional Steiner point for chasing the consistent model sets is inspired by [189] where second-order cone programs (SOCPs) are used to approximate the (nested set) Steiner point of the sublevel set of the work functions for chasing half-spaces.

Denote $\{(a_i, b_i)\}_{i=1}^{p_t}$ as the collection of p_t half-space constraints describing $\mathcal{P}_t$, i.e., $a_i^\top \theta \le b_i$. To approximate the integral for the functional Steiner point (6.3) of ω_t, we sample N number of random directions $v \in \mathbb{S}^{n-1}$, evaluate the Fenchel conjugate of the work function ω_t^* at each v with an SOCP, and take the empirical average. Finally we project the estimated functional Steiner point back to the set of consistent model $\mathcal{P}_t \cap \Theta$. Even though the analytical functional Steiner point (6.3) is guaranteed to be a member of the consistent model set, the projection step is necessary because we are integrating numerically, which may result in an approximation that ends up outside of the set. We summarize this procedure in Algorithm 3. Specifically, given a direction $v \in \mathbb{S}^{n-1}$, the Fenchel conjugate of the work function at time step t is

$$
\begin{aligned}
\omega_t^*(v) &= \inf_{x \in \mathbb{R}^n} \omega_t(x) - \langle x, v \rangle \\
&= \min_{\substack{x \in \mathbb{R}^n \\ q_s \in \mathcal{K}_s}} \sum_{s=1}^{t} \|q_s - q_{s-1}\| + \|x - q_t\| - \langle x, v \rangle .
\end{aligned}
$$

This can be equivalently expressed as the following SOCP with decision variables $x, q_1, \ldots, q_t, \lambda, \lambda_1, \ldots, \lambda_t$:

$$
\begin{aligned}
\min_{\substack{x, q_1, \ldots, q_t \\ \lambda, \lambda_1, \ldots, \lambda_t}} \quad & \lambda + \sum_{s=1}^{t} \lambda_s - \langle v, x \rangle \\
\text{s.t.} \quad & \|q_s - q_{s-1}\| \le \lambda_s, \quad \text{for } s \in [t] \\
& \|x - q_t\| \le \lambda \\
& a_i^\top q_s \le b_i, \quad \text{for } i \in [p_s], \ s \in [t] .
\end{aligned}
\tag{5.12}
$$

Another potential implementation challenge is that the number of constraints in the SOCP (5.12) grows linearly with time due to the construction of the work function

(5.4). This is a common drawback of online control methods based on CBC and NCBC techniques and can be overcome through truncation or over-approximation in of the work functions in practice. Additionally, if the LTV system is periodic with a known period, then we can leverage Algorithm 2 during the initial data collection phase. Once representative (persistently exciting) data is available, one could employ methods like [163] to generate a stabilizing controller for the unknown LTV system. In Section 5.4, we show that data collection via Algorithm 2 results in a significantly smaller state norm than random noise injection when the system is unstable.

5.4 Simulation

In this section, we demonstrate Algorithm 2 in two LTV systems. Both of the systems we consider are open-loop unstable, thus the algorithms must work to stabilize them. We use the same algorithm parameters for both, with $\Theta = [-2, 3]^2$, LQR cost matrices $Q = I$ and $R = 1$.

Example 1: Markov linear jump system

We consider the following Markov linear jump system (MLJS) model from [196], with

$$A_1 = \begin{bmatrix} 1.5 & 1 \\ 0 & 0.5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.6 & 0 \\ 0.1 & 1.2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \Pi = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix},$$

where Π is the transition probability matrix from θ_1 to θ_2 and vice versa. We inject uniformly random disturbances such that $w_t \in \{-10\mathbb{1}, -3\mathbb{1}, 3\mathbb{1}\}$ where $\mathbb{1}$ is the all-one vector. We set the disturbances to be zero for the last 10 time steps to make explicit the stability of the closed loop. We implement certainty-equivalent control based on online least squares (OLS) with different sliding window sizes $L = 5, 10, 20$ and a exponential forgetting factor of 0.95 [197] as the baselines.

We show two different MLJS models generated from 2 random seeds and show the results in Figure 5.1. For both systems, the open loop is unstable. In Figure 5.1(a) the OLS-based algorithms fail to stabilize the system for window size of $L = 20$, while stabilizing the system but incurring larger state norm than the proposed algorithm for $L = 5, 10$. On the other hand, in Figure 5.1(b), OLS with $L = 5$ results in unstable closed loop. This example highlights the challenge of OLS-based methods, where the choice of window size is crucial for the performance. Since the underlying LTV system is unknown and our goal is to control the system *online*, it is unclear how to

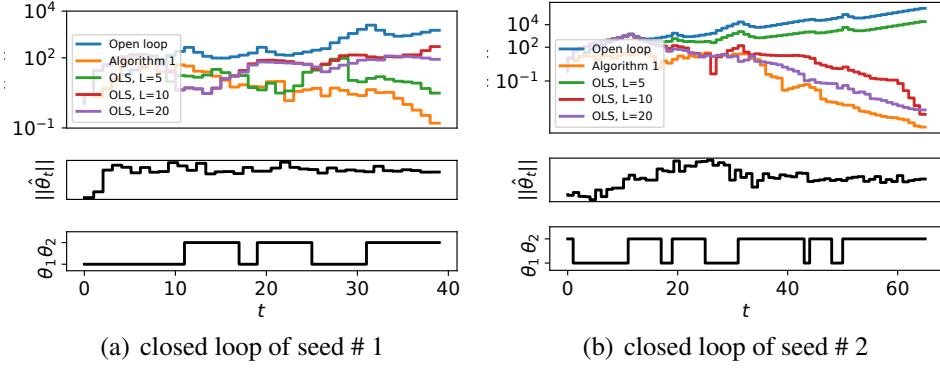(a) closed loop of seed # 1   (b) closed loop of seed # 2

Figure 5.1: Markov linear jump system for two different random seeds. For each seed: **Top** plot shows the state norm trajectories of the proposed algorithm, certainty-equivalent control based on online least squares (OLS) with different sliding window sizes, and the open loop. **Middle** plot shows the norm of the selected hypothesis model via Algorithm 3. **Bottom** plot shows the true model switches.

select the appropriate window size to guarantee stability for OLS-based methods a priori. In contrast, Algorithm 2 does not require any parameter tuning.

We note that while advanced least-squares based identification techniques that incorporate sliding window with *variable* length exist, e.g. [164], [197], due to the unknown system parameters, it is unclear how to choose the various algorithm parameters such as thresholds for system change detection. Therefore, we only compare Algorithm 1 against fixed-length sliding window OLS methods as baselines.

Example 2: LTV system

Our second example highlights that Algorithm 2 is a useful data-collection alternative to open-loop random noise injection. We consider the LTV system from [163], [184], with

$$A(k) = \begin{bmatrix} 1.5 & 0.0025k \\ -0.1\cos(0.3k) & 1 + 0.05^{3/2}\sin(0.5k)\sqrt{k} \end{bmatrix},$$

$$B(k) = 0.05 \begin{bmatrix} 1 \\ \frac{0.1k+2}{0.1k+3} \end{bmatrix}.$$

where we modified $A(1, 1)$ from 1 to 1.5 to increase the instability of the open loop in the beginning; thus making it more challenging to stabilize. We consider no disturbances here, which is a common setting in direct data-driven control, e.g., [163], [182], [183]. In particular, we compare the proposed algorithm against randomly generated bounded inputs from $\mathsf{UNIF}[-1, 1]$. We also modify the control inputs from Algorithm 2 to be $u_t = \widehat{K}_{t-1}x_t + \eta_t \cdot \mathbb{1}$ with $\eta_t \sim \mathsf{UNIF}[-1, 1]$ so that we can collect
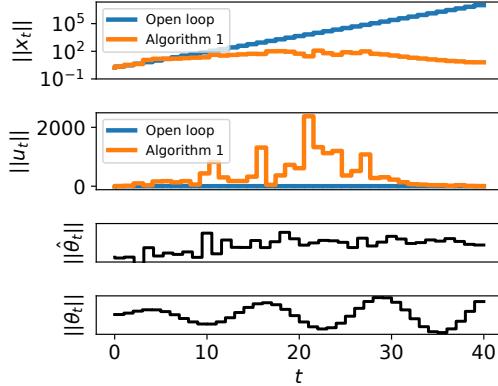
Figure 5.2: Simulation result for the LTV system in example 2. Here we plot the the state and control norm, as well as the selected hypothesis model via CBC $\widehat{\theta}_t$ and true models θ_t.

rich data in the closed loop. This is motivated by the growing body of data-driven control methods such as [163], [183], [184] that leverage sufficiently rich offline data to perform control design for unknown LTV systems. However, most of these works directly inject random inputs for data collection. It is evident in Figure 5.2 that when the open-loop system is unstable it may be undesirable to run the system without any feedback control. Therefore, Algorithm 2 complements existing data-driven methods by allowing safe data collection with significantly better transient behavior.

5.5 Conclusion

In this chapter, we propose a model-based approach for stabilizing an unknown LTV system under arbitrary non-stochastic disturbances in the sense of bounded input bounded output under the assumption of infrequently changing or slowly drifting dynamics. Our approach uses ideas from convex body chasing (CBC), which is an online problem where an agent must choose a sequence of points within sequentially presented convex sets with the aim of minimizing the sum of distances between the chosen points. The algorithm requires minimal tuning and achieves significantly better performance than the naive online least squares based control. Future work includes sharpening the stability analysis to go beyond the BIBO guarantee in this work, which will require controlling the difference between the estimated disturbances and true disturbances. Another direction is to extend the current results to the networked case, similar to [99].

5.A   Proof of Lemma 22

We have

$$\sum_{\tau=s+1}^{e} \left\| \widehat{\theta}_\tau - \widehat{\theta}_{\tau-1} \right\|_F = \sum_{\tau=s+1}^{e} \| \mathsf{st}(\omega_\tau) - \mathsf{st}(\omega_{\tau-1}) \|_F$$

$$\stackrel{(a)}{\leq} n \fint_v \left(\sum_{\tau=s+1}^{e} \left| \omega_\tau^*(v) - \omega_{\tau-1}^*(v) \right| \right) v \, dv$$

$$\stackrel{(b)}{=} n \fint_v \left(\sum_{\tau=s+1}^{e} \omega_\tau^*(v) - \omega_{\tau-1}^*(v) \right) v \, dv$$

$$= n \fint_v \left(\omega_e^*(v) - \omega_s^*(v) \right) v \, dv$$

$$\stackrel{(c)}{\leq} n \cdot \left(\min_x \omega_e(x) - \min_y \omega_s(y) + 2\kappa \right), \qquad (5.13)$$

where (a) is due to the definition (6.3). For (b), we used the observation that $\omega_t^*(v)$ is non-decreasing in time. For (c), by definition of the Fenchel conjugate (5.3), we have that $\omega_e^*(v) = \inf_x \omega_e(x) - \langle x, v \rangle$. Denote $(x^\star, q_1^\star, \ldots, q_e^\star)$ as the optimal solution to the problem $\min_x \omega_e(x)$. It is clear that $\omega_e^*(v) \leq \omega_e(x^\star) - \langle x^\star, v \rangle \leq \min_x \omega_e(x) + \kappa$ where in the last inequality we used Cauchy-Shwarz and $\kappa := \max_{\theta \in \Theta} \|\theta\|_F$. Similarly, we also have $\omega_s^*(v) \geq \inf_y \omega_s(y) - \kappa$.

Denote $\mathrm{OPT}_{[0,e]}$ as the minimizing trajectory $(\mathrm{OPT}_0, \ldots, \mathrm{OPT}_e)$ to $\min_x \omega_e(x)$ where $\mathrm{argmin}_x \omega_e(x) = \mathrm{OPT}_e$. This last equality is by the observation that if $x^\star := \mathrm{argmin}_x \omega_e(x) \neq \mathrm{OPT}_e$, then $\omega_e(\mathrm{OPT}_e) \leq \omega_e(x^\star)$ by definition (5.4), thus contradicting that $x^\star$ is defined to be the minimizer of ω_e. We also denote $\mathrm{INT}_{[0,s]}$ as the minimizing trajectory to $\min_y \omega_s(y)$. To reduce notation, we denote $\Delta_{[s,e]}^{\mathrm{OPT}} := \sum_{\tau=s+1}^{e} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F$ and $\Delta_{[s,e]}^{\mathrm{INT}} := \sum_{\tau=s+1}^{e} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F$. Then we have

$$(5.13) = n \cdot \left(\Delta_{[0,e]}^{\mathrm{OPT}} - \Delta_{[0,s]}^{\mathrm{INT}} + 2\kappa \right)$$

$$\stackrel{(c)}{\leq} n \cdot \left(\Delta_{[0,e]}^{\mathrm{OPT}} - \Delta_{[0,s]}^{\mathrm{OPT}} + \mathsf{dia}(\Theta) + 2\kappa \right)$$

$$= n \cdot \left(\Delta_{[s,e]}^{\mathrm{OPT}} + \mathsf{dia}(\Theta) + 2\kappa \right),$$

where (c) holds because if $\sum_{\tau=1}^{s} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F > \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \mathsf{dia}(\Theta)$ and $\mathrm{OPT}_{[0,s]} \neq \mathrm{INT}_{[0,s]}$, then we can replace the $[0, s]$ portion of the optimal trajectory $\mathrm{OPT}_{[0,e]}$ with $\mathrm{INT}_{[0,s]}$ and achieve a lower cost for $\omega_e(\mathrm{OPT}_e)$, thus contradicting the optimality of $\mathrm{OPT}_{[0,e]}$. To see why the fictitious trajectory $\left(\mathrm{INT}_{[0,s]}, \mathrm{OPT}_{[s+1,e]} \right)$ achieves lower cost than $\mathrm{OPT}_{[0,e]}$, we compare the total

movement cost during the interval $[0, s+1]$,

$$\sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{INT}_s\|_F$$

$$\leq \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F$$

$$+ \|\mathrm{OPT}_s - \mathrm{INT}_s\|_F$$

$$\leq \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F + \mathrm{dia}(\Theta)$$

$$< \sum_{\tau=1}^{s} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F \,,$$

which means the fictitious trajectory achieves lower overall cost. Therefore (c) must hold. $\blacksquare$

5.B Auxiliary Results

Here we summarize the helper lemmas used in the proof sketch of Theorem 18. First, we define some useful notation.

Lyapunov equation. Let $X, Y \in \mathbb{R}^{n \times n}$ with $Y = Y^\top > 0$ and $\rho(X) < 1$. Define $\mathrm{dlyap}(X, Y)$ to be the unique positive definite solution Z to the Lyapunov equation $X^\top Z X - Z = Y$. For a stabilizable system (A, B) with optimal infinite-horizon LQR feedback $K := K^*([A \ B])$ with cost matrices $Q, R = I$, we define

$$P(A, B) = \mathrm{dlyap}(A + BK^*([A \ B]), I_n + K^*([A \ B])^\top K^*([A \ B]))$$

and

$$H(A, B) = \mathrm{dlyap}(A + BK^*([A \ B]), I_n).$$

We also define the shorthand for the following:

$$P_t := P(\widehat{A}_t, \widehat{B}_t), \quad H_t := H(\widehat{A}_t, \widehat{B}_t). \tag{5.14}$$

Constants. Throughout the proof, we will reference the following system-theoretical constants for the parameter set Θ defined in Assumption 7:

$$\|K_*\| := \sup_{[A \ B] \in \Theta} \|K^*([A \ B])\|, \gamma_* := \max_{[A \ B] \in \Theta} \|A + BK^*([A \ B])\|.$$

We also quantify the stability of every model in Θ under its corresponding optimal LQR gain. Let

$$C_* > 0, \quad r_* \in (0, 1)$$

be such that for all $\theta := [A \ B] \in \Theta$, $K := K^*(\theta)$, and $i \in \mathbb{N}_+$, $\left\| \left((A + BK)^T \right)^i \right\| \cdot \left\| (A + BK)^i \right\| \leq C_* r_*^{2i}$. By Lemma 23 which is stated below and Assumption 7, such C_* and r_* always exist. Further, we define

$$\|P_*\| := \sup_{[A \ B] \in \Theta} \|P(A, B)\|, \quad \|H_*\| := \sup_{[A \ B] \in \Theta} \|H(A, B)\|,$$

$$\epsilon_* := 1 / \left(54 \|P_*\|^5 \right), \quad c_* := \max_{[A \ B] \in \Theta} \frac{\lambda_{\max} H(A, B)}{\lambda_{\min} H(A, B)},$$

$$h_* := \sup_{[A_1 \ B_1], [A_2 \ B_2] \in \Theta} \left\| H(A_1, B_1)^{1/2} \right\| \left\| H(A_2, B_2)^{-1/2} \right\|.$$

To justify the existence of these constants, note that discrete-time optimal LQR controller has guaranteed stability margin [198] and that by Lemma 23 and the fact that the solution to Lyapunov equation has the following closed form:

$$P(A, B) = \sum_{i=0}^{\infty} \left((A + BK)^\top \right)^i (I + K^\top K)(A + BK)^i, \tag{5.15}$$

we have that for all $[A, B] \in \Theta$,

$$\|P(A, B)\| \leq \left(1 + \|K\|^2 \right) \left(1 + \sum_{i=1}^{\infty} \left\| \left((A + BK)^\top \right)^i \right\| \left\| (A + BK)^i \right\| \right)$$

$$\leq \frac{\left(1 + \|K_*\|^2 \right) \left(1 - r_*^2 + C_* \right)}{1 - r_*^2} =: \|P_*\|.$$

We can similarly derive $\|H_*\|$. By definition of the Lyapunov solution (5.15), $\|P_*\| \geq \|H_*\| \geq 1$.

Lemma 23 ([199, page 183]). *For a matrix $A \in \mathbb{R}^{n \times n}$, with $\rho := \rho(A)$, there exist constants κ_1, κ_2 such that for any positive integer i*

$$\kappa_1 \rho^i i^{n_1 - 1} \leq \left\| A^i \right\| \leq \kappa_2 \rho^i i^{n_1 - 1}$$

where n_1 is the size of the largest Jordan block corresponding to eigenvalue of ρ in Jordan block form representation of A.

Lemma 24 ([59, Proposition 6]). *Let $\Theta = [A \ B]$ be a stabilizable system, with optimal controller $K := K^*(\theta)$ and $P := P(A, B)$. Let $\widehat{\theta} = [\widehat{A} \ \widehat{B}]$ be an estimate of θ, $\widehat{K} := K^*(\widehat{\theta})$ the optimal controller for the estimate, and $\epsilon := \max \left\{ \left\| A - \widehat{A} \right\|, \left\| B - \widehat{B} \right\| \right\}$. Then if $\alpha := 8 \|P\|^2 \epsilon < 1$:*

$$\left\| B \left(\widehat{K} - K \right) \right\| \leq 8(1 - \alpha)^{-7/4} \|P\|^{7/2} \epsilon.$$

Lemma 25 ([59, Theorem 8]). *Let $\theta = [A\ B]$ be a stabilizable system, with $P := P(A, B)$, and $H = H(A, B)$. Let $\widehat{\theta} = [\widehat{A}\ \widehat{B}]$ be an estimate of θ satisfying $\max\left\{\left\|A - \widehat{A}\right\|, \left\|B - \widehat{B}\right\|\right\} \leq \epsilon$. Consider certainty equivalent controller $\widehat{K} = K^*(\widehat{\theta})$. Then if ϵ is such that $54\,\|P\|^5\,\epsilon \leq 1$, we have*

$$(A + B\widehat{K})^\top H(A + B\widehat{K}) \preceq \left(1 - \frac{1}{2}\,\|H\|^{-1}\right)H \preceq \left(1 - \frac{1}{2}\,\|P\|^{-1}\right)H.$$

Lemma 26 ([200]). *Let X be the solution to the Lyapunov equation $X - F^\top X F = M$, and let $X + \Delta X$ be the solution to the perturbed problem*

$$Z - (F + \Delta F)^\top Z(F + \Delta F) = M.$$

The following inequality holds for the spectral norm:

$$\frac{\|\Delta X\|}{\|X + \Delta X\|} \leq 2\left\|\sum_{k=0}^{+\infty} (F^\top)^k\,F^k\right\| \cdot (2\|F\| + \|\Delta F\|) \cdot \|\Delta F\|.$$

Lemma 27. *Suppose $\epsilon_{t+1} := \max\left\{\left\|\widehat{A}_{t+1} - \widehat{A}_t\right\|, \left\|\widehat{B}_{t+1} - \widehat{B}_t\right\|\right\}$ and $\alpha := 8\,\|P_*\|^2\,\epsilon_{t+1} \leq 1/2$, . Then H_t defined in (5.14) satisfies*

$$H_t \preceq H_{t+1}(1 + \eta_{t+1})$$

for $\eta_{t+1} := c_*\beta_*\epsilon_{t+1}$, and

$$\beta_* := \frac{2C_*}{1 - r_*^2}\,(2\gamma_* + 3 + \|K_*\|)\left(1 + 32\,\|P_*\|^2 + \|K_*\|\right).$$

Proof of Lemma 27. For notational brevity, we drop the time index for ϵ and η in the proof. Applying Lemma 26 with $X = H_t$, $X + \Delta X = H_{t+1}$ and $F = \widehat{A}_t + \widehat{B}_t\widehat{K}_t$ and $\Delta F = (\widehat{A}_{t+1} - \widehat{A}_t) + (\widehat{B}_{t+1}\widehat{K}_{t+1} - \widehat{B}_t\widehat{K}_t)$, and $M = I_n$ we have

$$\frac{\|H_{t+1} - H_t\|}{\|H_{t+1}\|} \leq 2\left\|\sum_{k=0}^{+\infty}\left((\widehat{A}_t + \widehat{B}_t\widehat{K}_t)^\top\right)^k (\widehat{A}_t + \widehat{B}_t\widehat{K}_t)^k\right\|$$

$$\cdot\left(2\left\|\widehat{A}_t + \widehat{B}_t\widehat{K}_t\right\| + \left\|\widehat{A}_{t+1} - \widehat{A}_t)\right\| + \right.$$

$$\left.\left\|\widehat{B}_{t+1}(\widehat{K}_{t+1} - \widehat{K}_t)\right\| + \left\|(\widehat{B}_{t+1} - \widehat{B}_t)\widehat{K}_t\right\|\right)$$

$$\cdot\left(\left\|\widehat{A}_{t+1} - \widehat{A}_t)\right\| + \left\|\widehat{B}_{t+1}(\widehat{K}_{t+1} - \widehat{K}_t)\right\| + \left\|(\widehat{B}_{t+1} - \widehat{B}_t)\widehat{K}_t\right\|\right)$$

$$\leq \epsilon\frac{2C_*}{1 - r_*^2}\left(2\gamma_* + \epsilon\left(1 + 32\,\|P_*\|^2 + \|K_*\|\right)\right)\cdot$$

$$\left(1 + 32\,\|P_*\|^2 + \|K_*\|\right)$$

$$\leq \epsilon \frac{2C_*}{1 - r_*^2} (2\gamma_* + 3 + \|K_*\|) \left(1 + 32 \|P_*\|^2 + \|K_*\|\right)$$

$$=: \epsilon\beta,$$

where in the second inequality we used Lemma 24 to bound $\left\|\widehat{B}_{t+1}(\widehat{K}_{t+1} - \widehat{K}_t)\right\| \leq$ $32 \|P_{t+1}\|^{7/2} \epsilon$ and in the last inequality we use the assumption $8\epsilon \|P_*\|^2 \leq 1/2$.

To show $H_t \preceq H_{t+1}(1 + \eta)$ for some η, it suffices to show that for all vectors $v \in RR^n$, $v^\top (H_t - H_{t+1})v \leq \eta v^\top H_{t+1}v$. With the preceding calculation, we have

$$\begin{aligned} v^\top (H_t - H_{t+1})v &\leq \|v\|^2 \|H_t - H_{t+1}\| \\ &\leq \epsilon\beta_* \|v\|^2 \|H_{t+1}\| \\ &\leq \epsilon\beta_* c_* \lambda_{\min}(H_{t+1}) \|v\|^2 \\ &\leq \epsilon\beta_* c_* v^\top H_{t+1}v \,. \end{aligned}$$

This proves the desired bound, with $\eta = c_*\beta_*\epsilon$ and

$$\beta_* = \frac{2C_*}{1 - r_*^2} (2\gamma_* + 3 + \|K_*\|) \left(1 + 32 \|P_*\|^2 + \|K_*\|\right).$$

$\blacksquare$

5.C   Proof of Theorem 18

Recall that the closed loop dynamics can be characterized as (5.6). Therefore,

$$\|x_t\| \leq W + W \sum_{s=0}^{t-2} \left\| \prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right) \right\|. \tag{5.16}$$

Define

$$L_t := H_t^{-1/2}(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1})H_t^{1/2},$$

where H_t is defined in (5.14). This gives,

$$\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1} := H_t^{1/2} L_t H_t^{-1/2}.$$

Therefore, each summand in (5.16) can be bounded as

$$\left\| \prod_{\tau \in I_s} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right) \right\|$$

$$\leq \underbrace{\left\|H_{t-1}^{1/2}\right\| \left\|H_{s+1}^{-1/2}\right\|}_{(a)} \underbrace{\prod_{k \in I_{s+1}} \left\|H_k^{-1/2} H_{k-1}^{1/2}\right\|}_{(b)} \underbrace{\prod_{\tau \in I_s} \|L_\tau\|}_{(c)} \tag{5.17}$$

where we used I_s as shorthand for the interval $[t - 1 : s + 1]$.

Bounding (a). We directly use the system-theoretical constant introduced in Section 5.B so that (a) $\leq h_*$.

Bounding (b). Lemma 27 directly implies that for all $t \in \mathbb{N}_+$, $H_{t-1}H_t^{-1} \leq (1 + \eta_t)I$. Therefore, we have

$$\left\| H_{t-1}^{1/2} H_t^{-1/2} \right\| \leq (1 + \eta_t)^{1/2} \leq 1 + \eta_t/2 \leq e^{\eta_t/2}.$$

Hence with the fact that H_t's are symmetric,

$$\left\| H_t^{-1/2} H_{t-1}^{1/2} \right\| \leq \begin{cases} e^{\frac{c_* \beta_* \left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F}{2}}, & \left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F \leq \epsilon_* \\ h_* & \text{otherwise.} \end{cases} \tag{5.18}$$

Bounding (c). Lemma 25 implies that if $\left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F \leq \epsilon_*$ then

$$\left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1} \right)^\top H_t \left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1} \right) \leq \left(1 - \frac{1}{2} \| P_t \|^{-1} \right) H_t.$$

This in turn implies that

$$L_t^\top L_t = H_t^{-1/2} (\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1})^\top H_t (\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}) H_t^{-1/2}$$

$$\leq H_t^{-1/2} \left(1 - \frac{1}{2} \| P_t \|^{-1} \right) H_t H_t^{-1/2}$$

$$\leq \left(1 - \frac{1}{2} \| P_t \|^{-1} \right) I_n.$$

This in turn implies that $\| L_t \| \leq \left(1 - \frac{1}{2\|P_*\|} \right)^{1/2}$. To summarize,

$$\| L_t \| \leq \begin{cases} \rho_L := \left(1 - \frac{1}{2\|P_*\|} \right)^{1/2} < 1, & \left\| \widehat{\theta}_t - \widehat{\theta}_{t-1} \right\|_F \leq \epsilon_* \\ \ell_* & \text{otherwise,} \end{cases} \tag{5.19}$$

for some constant ℓ_* such that for all $t \in \mathbb{N}_+$,

$$\left\| H_t^{1/2} (\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}) H_t^{-1/2} \right\| \leq \ell_*$$

Combining (a,b,c). We now plug in the bounds (5.18) and (5.19) into (5.17). Let $\widehat{\Delta}_{[s,e]} := \sum_{\tau=s+1}^{e} \left\| \widehat{\theta}_\tau - \widehat{\theta}_{\tau-1} \right\|_F$ be the partial-path movement of the selected hypothesis models and $\Delta_{[s,e]} := \sum_{\tau=s+1}^{e} \| \theta_\tau - \theta_{\tau-1} \|_F$ be the true model partial-path

variation. We also denote by $n_{s,t}$ the number of pairs $(\tau, \tau - 1)$ with $s + 1 \le \tau \le t - 1$ where $\left\|\widehat{\theta}_\tau - \widehat{\theta}_{\tau-1}\right\|_F > \epsilon_*$. Note that $n_{s,t} \le \widehat{\Delta}_{[s,t-1]}/\epsilon_*$. Therefore,

$$
\left\| \prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1} \right) \right\|
$$

$$
\le h_* \cdot h_*^{n_{s,t}} \cdot e^{\frac{c_* \beta_* \widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \ell_*^{n_{s,t}} \cdot \rho_L^{t-s-1-n_{s,t}}
$$

$$
\le h_* \left(\frac{\ell_* h_*}{\rho_L} \right)^{\frac{\widehat{\Delta}_{[s,t-1]}}{\epsilon_*}} e^{\frac{c_* \beta_* \widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \rho_L^{t-s-1}
$$

$$
\le h_* \left(\frac{\ell_* h_*}{\rho_L} \right)^{\frac{\bar{n}\left(\mathrm{dia}(\Theta) + 2\kappa + \Delta_{[s,t-1]}\right)}{\epsilon_*}} \cdot e^{\frac{c_* \beta_* \bar{n}\left(\mathrm{dia}(\Theta) + 2\kappa + \Delta_{[s+1,t-1]}\right)}{2}} \cdot \rho_L^{t-s-1}
$$

$$
=: c_0 \cdot c_1^{\Delta_{[s,t-1]}} \rho_L^{t-s-1},
$$

where $\bar{n} := n(n + m)$ is the dimension of the parameter space for $[A_t \; B_t]$. Finally plugging the above in (5.16) gives

$$
\|x_t\| \le W \left(1 + c_0 \sum_{s=0}^{t-2} c_1^{\Delta_{[s,t-1]}} \rho_L^{t-s-1} \right).
$$

∎

Chapter 6

ONLINE ADVERSARIAL STABILIZATION OF UNKNOWN NETWORKED SYSTEMS

In this chapter, we continue to explore the SME uncertainty set-based learning and control framework. In particular, we will instantiate the framework for another important class of problems where the goal is to stabilize *unknown networked* systems under communication constraints and non-stochastic (potentially adversarial) perturbations. Networked systems model a variety of CPS, such as power systems, connected vehicles, and building control systems. As sensing, actuation, and communication technologies continue to expand for CPS, these network models will be too large-scale and complex to model accurately. Moreover, due to the sheer scale of such systems, communication constraints, such as delay and locality as introduced in Part I, must be considered in the algorithm design. To add to the challenge, CPS are increasingly susceptible to attacks. These considerations motivate us to devise an algorithm that can guarantee network stability for unknown systems under communication constraints against adversarial attacks.

We will propose the first provably stabilizing algorithm for this setting. The algorithm uses a distributed version of nested convex body chasing to maintain a consistent estimate of the network dynamics and applies SLS to determine a distributed controller based on the selected model. Our approach accommodates a broad class of communication delays while enabling fully distributed execution and scaling favorably with the number of subsystems in the network.

[1] J. Yu, D. Ho, and A. Wierman, "Online adversarial stabilization of unknown networked systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems (SIGMETRICS)*, vol. 7, no. 1, pp. 1–43, 2023. DOI: 10.1145/3579452.

6.1 Introduction

Large-scale networked dynamical systems play a crucial role in many emerging engineering systems such as the power grid [69], autonomous vehicles [70], and swarm robots [201]. Motivated by the success of learning-based control methods for single-agent (centralized) linear systems, there has been growing interest in learning

distributed controllers for unknown networked systems composed of interconnected and spatially distributed linear time-invariant (LTI) subsystems [121], [202]–[205].

However, since most existing literature ports centralized learning-based control techniques over to the distributed setting, almost all previous work assumes that the underlying dynamics are stable, or that a stabilizing and distributed controller is known. For a large-scale networked system, such assumptions are often unrealistic, because designing stabilizing distributed controllers itself is a significant task even if the dynamics model is available [39], [44], [71], [73], [206], [207].

Recent work has begun to lift the assumption of the knowledge of a stabilizing controller in the centralized case, e.g. [26], [96], [208]. This line of work follows the approach of system identification, either by letting the unstable system run open-loop or by exciting the system via control inputs. However, such approaches induce explosive transient behaviors due to the instability of the underlying system. Without proper generalization to the networked setting, such explosive behavior can cause catastrophic system degradation before a proper stabilizing controller can be learned.

Further, until now, scalability and information constraints have only been considered separately in learning-based distributed controller design; no general approach exists. On the other hand, information constraints and scalability have been the central topics in distributed control for the past decade due to their theoretical challenge and practical importance [40]–[43], [78]. Therefore, it is crucial to simultaneously consider such constraints when designing learning-based distributed control algorithms for networked systems.

Contribution

We overcome the aforementioned challenges by leveraging recent advances in online learning and distributed control. In particular, we propose an approach that combines a distributed version of nested convex body chasing (NCBC), in order to maintain a consistent estimate of the network dynamics, with system level synthesis (SLS), in order to determine a distributed controller based on the selected consistent model. This combination yields the first online algorithm that provably stabilizes a networked LTI system with information constraints under adversarial disturbances (Theorem 22). The proposed algorithm (Algorithm 5) is distributed and scales favorably to the number of subsystems in the network.

The approach in this chapter is fundamentally different than traditional system identification based methods, which may incur prohibitively large state norm under

Table 6.1: Maximum and top 90% infinity norm of the state ($\|x(t)\|_\infty$) for different disturbance profiles averaged over 10 runs. The goal of the controller is to minimize the size of the state. Simulation details are provided in Section 6.5.

Algorithm	Correlated Gaussian (Top 90%)	Uniform (Top 90%)	State-dependent (Top 90%)
This work	$\mathbf{1.21 \times 10^1}$ (0.31×10^1)	$\mathbf{2.30 \times 10^1}$ (0.36×10^1)	$\mathbf{7.14 \times 10^1}$ (0.54×10^1)
SysID	5.12×10^{11} (1.71×10^{11})	5.12×10^{11} (1.71×10^{11})	5.12×10^{11} (1.71×10^{11})

adversarial disturbances, even in the simplest setting (see Table 6.1). The reason is that system identification-based approaches seek to learn the full system dynamics, which requires full excitation of the system against worst-case disturbances. On the other hand, our approach does not require precise knowledge of the system. Instead, we maintain model estimates that are consistent with the observations generated by the unknown system at all times. A consequence of focusing on consistency is a natural endogenous exploration-exploitation scheme where our algorithm performs well (small state norm) while the selected model stays consistent, and gains information about the system whenever it observes a large state norm that renders the selected model inconsistent.

The main result of this chapter is an input-to-state stability guarantee (Theorem 22), where we draw novel connections between the path length property of NCBC techniques and system stability analysis. This follows from a set of novel technical results for SLS in the learning-based control context. In particular, we generalize a previous result [73] on the characterization of the closed loop under SLS controllers that are synthesized from an arbitrary and potentially incorrect system model (Lemma 28). This result enables the analysis of our algorithm when each subsystem uses local, asynchronous, and wrong model information for local controller synthesis. Further, we derive a novel perturbation result with explicit constants for finite-horizon SLS synthesis (Theorem 21) that globally bounds the sensitivity of the optimal solution to the SLS problem (a quadratic program with equality and sparsity constraints) with respect to the model. This result is also applicable in other contexts such as a class of MPC problems under sparsity constraints [209]–[211].

Related Work

This work contributes to a large and growing body of work on the topics related to learning-based control design, online control, and distributed control. We briefly review the literature most related to this work below.

Stabilization of unknown systems. Stabilizing unknown linear systems has long

been a fundamental problem studied in adaptive control theory [212]. It recently reemerged as a learning problem and received considerable attention from the machine learning community [208], [213]–[215]. Most works have been developed under single-agent setting, with a no-noise assumption [181], [216] or Gausssian noise models [17], [147]. Under the adversarial noise setting, which is the focus of this chapter, the only work that guarantees stabilization for LTI systems is [26], with a system identification-based approach that achieves order-optimal regret. In contrast, we propose a novel framework for stabilization under adversarial noise that does not rely on accurate identification of the true dynamics. In particular, our method is the first algorithm to stabilize a networked LTI system under adversarial disturbances with information constraints while simultaneously achieving magnitudes of improvement in empirical performance over the state-of-the-art identification-based approach [26] in the single-agent setting, despite the regret-optimal guarantee in [26].

Distributed control. Motivated by large-scale cyberphysical systems that are composed of physically distributed subsystems with local dynamical interactions, there is a large body of work on control design for networked systems [73], [207], [217]. Cyberphysical systems such as the power grid are commonly constrained by a communication layer that allows specific structure of information exchange among the subsystems. Such information structure imposes significant challenges for optimal control design, often rendering the problem NP-hard [218]. In [39], it was shown that a large class of practically relevant distributed control problems is convex and tractable to solve. Since then, many works have focused on this class of problems [206], [219]. However, [45] observes that the complexity of computation and implementation of distributed controllers developed under this setting can be prohibitively expensive, and thus not scalable to large-scale systems. The System Level Synthesis (SLS) framework is developed as a scalable alternative to distributed control design [73]. In particular, SLS allows order-constant complexity for synthesis and implementation, due to its special parameterization and implementation of the feedback controller. As a result, many works have adopted SLS as the basis for novel (learning-based) control algorithms in both distributed and centralized setting [61], [210], [220], [221]. We contribute to the literature on SLS by developing a suit of technical results for SLS controllers that can find applications beyond the setting of this work.

Learning distributed controllers. Many learning-based control algorithms for networked systems adopt a centralized learning or computational approach with

the objective of regret minimization, e.g., [202]–[205], [222]. All prior work use the stochastic noise or no-noise model and assume a known stabilizing distributed controller is given [121], [210], [223]–[226]. As far as we are aware, no previous work accommodates communication delay while doing both learning and control. The most related to our work are [227] and [203], where learning-based SLS controllers are designed to control unknown networked systems. Both of the methods require the knowledge of a stabilizing and distributed controller. [227] is only applicable to small-uncertainty scenarios, while [203] requires a stabilizing distributed controller and performs centralized learning. In this work, we focus on stabilization and propose the first distributed learning-based control algorithm that guarantees stability for unknown networked systems under adversarial disturbances.

Online learning. The problem of online stabilization for unknown dynamical systems is an instance of online decision making problems, where an agent makes a sequence of decisions based on the feedback from an unknown environment with the goal of cost minimization. Online decision making is studied extensively in the online learning literature, with a line of work [55], [165], [228], [229] that makes interesting connections between convex function and body chasing [230], [231] and linear control theory. In particular, [13] proposes an online nonlinear robust control method based on convex body chasing that guarantees finite mistakes under adversarial disturbances without the need for system identification. While [13] considers binary cost functions for general nonlinear systems, we present novel technical results that establish the first connection between convex body chasing and stability analysis for both single-agent and networked multi-agent linear dynamical systems.

Notation. Let $\| \cdot \|$ be the ℓ_2 norm and $\| \cdot \|_F$ be the Frobenius norm. We denote the (i, j)th position of a matrix M as $M(i, j)$ and use $M(:, j), M(i, :)$ for the jth column and ith row of M respectively. We use $[N]$ for the set of positive integers up to N. Positive integers are denoted as $\mathbb{N}_+$. Bold face lower cases are reserved for vector signal of the form $\mathbf{x} := \{x(t)\}_{t=0}^{\infty}$ with $x(t) \in \mathbb{R}^n$. We reserve bold face capital letters for causal linear operators/transfer matrices with components $K[0], K[1], \ldots$, such that

$$
\mathbf{K} := \begin{bmatrix} K[0] & 0 & \ldots \\ K[1] & K[0] & 0 & \ldots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}.
$$

We write $\mathbf{y} = \mathbf{Gx}$ to mean that $y(t) = \sum_{k=0}^{t} G[k]x(t-k)$. Given any binary matrix

(a) System with communication graph $\mathcal{G}^C$.

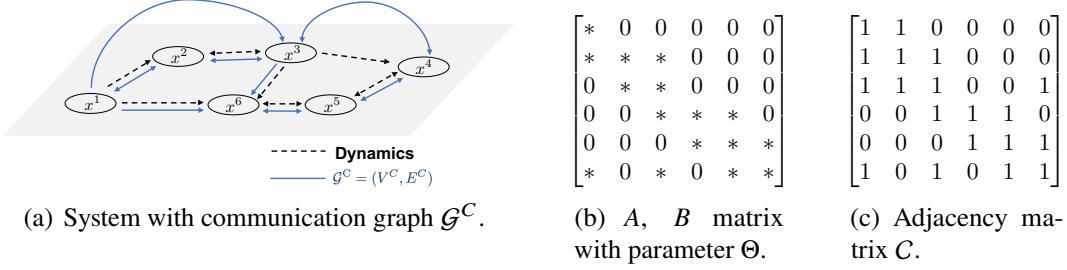(b) A, B matrix with parameter Θ.

(c) Adjacency matrix C.

Figure 6.1: Example networked LTI system with information constraints.

$C \in \{1, 0\}^{N \times N}$, we say $M \in C$ for a matrix $M \in \mathbb{R}^{N \times N}$ if the sparsity of M is C. We use $\{e_j\}_{j=1}^n$ for the standard basis in $\mathbb{R}^n$.

6.2 Preliminaries and Problem Setup

We consider the task of stabilizing an unknown networked system made up of N interconnected, heterogeneous linear time-invariant (LTI) subsystems, illustrated in Figure 6.1(a). For each subsystem $i \in [N]$, let $x^i(t) \in \mathbb{R}^{n_i}$, $u^i(t) \in \mathbb{R}^{m_i}$, $w^i(t) \in \mathbb{R}^{n_i}$ be the local state, control, and disturbance vectors respectively. Each subsystem i has dynamics,

$$x^i(t + 1) = \sum_{j \in \mathcal{N}(i)} \left(A^{ij} x^j(t) + B^{ij} u^j(t) \right) + w^i(t), \tag{6.1}$$

where we write $j \in \mathcal{N}(i)$ if the states or control actions of subsystem j affect those of subsystem i through the open-loop network dynamics ($i \in \mathcal{N}(i)$). Concatenating all the subsystem dynamics, we can represent the global dynamics as

$$x(t + 1) = Ax(t) + Bu(t) + w(t), \tag{6.2}$$

where $x(t) \in \mathbb{R}^{N_x}$, $u(t) \in \mathbb{R}^{N_u}$, $w(t) \in \mathbb{R}^{N_x}$, with $N_x = \sum_{i=1}^N n_i$ and $N_u = \sum_{i=1}^N m_i$, and we define A^{ij}, $B^{ij} \equiv 0$ for all $j \notin \mathcal{N}(i)$. The networked LTI model (6.1) has been extensively studied in the networked control literature for various applications such as robotic swarms [232], voltage control for the distribution network of the power grid [55], and many other large-scale cyber-physical systems [37], [233]. An example is the linearized swing equation for power systems, where the global system is composed of a mesh of interacting buses [234], [235]. In this setting, the states x^i of each bus i is two-dimensional and corresponds to the phase angle relative to some given setpoint and the associated frequency. The input u^i at bus i is the controllable load, while w^i is the bounded load disturbances that are often correlated in space and time.

We assume that the topology among the subsystems is known, *i.e.*, the sets $\mathcal{N}(i)$ for $i \in [N]$ are known. However, the parameters of the dynamics (entries of matrices A^{ij}, B^{ij}) are unknown. Let θ^i denote the unknown *local parameter* for subsystem i, i.e., $\theta^i := \left(A^{ij},\, B^{ij}\right)_{j \in \mathcal{N}(i)}$. Further, let $\Theta := (\theta_1, \ldots, \theta_N)$ be the *global parameter*. We write $A(\Theta)$ and $B(\Theta)$ (equivalently $A^{ij}(\theta^i)$, $B^{ij}(\theta^i)$) to emphasize that A and B are matrices constructed with appropriate zeros according to the network topology (known), and the nonzero entries specified by Θ (unknown).

Example 6. *Consider the networked system in Figure 6.1(a) where each subsystem $i \in [6]$ has $x^i(t) \in \mathbb{R}$ and $u^i(t) \in \mathbb{R}$. For each i, the set $\mathcal{N}(i)$ contains the subsystems that has a dashed arrow pointing towards x^i in the figure. For example, $\mathcal{N}(6) = \{1, 3, 5, 6\}$. Each A^{ij} and B^{ij} for $j \in \mathcal{N}(i)$ is a scalar. The stacked global dynamics has matrix A and B with structure shown in Figure 6.1(b). The unknown local parameter θ^i corresponds to the $*$ entries of the ith row of A and B, while the global parameter Θ is a vector containing $*$ entries in matrix A and B.*

We now introduce three core assumptions needed for our algorithm and analysis. As we highlight below, these are standard assumptions in the learning-based control literature.

Assumption 8 (Adversarial disturbances). $\|w(t)\|_\infty \leq W$ *for* (6.2).

Assumption 9 (Compact Parameter Set). *The network structure $\mathcal{N}(i)$ for $i \in [N]$ is known. The true system parameter $\Theta^\star := \left(\theta^{1,\star}, \ldots, \theta^{N,\star}\right)$ is an element of a known compact convex set $\mathcal{P}_0 = \mathcal{P}_0^1 \times \cdots \times \mathcal{P}_0^N$, which is a product space of local parameter sets where $\theta^{i,*} \in \mathcal{P}_0^i$. The known parameter set is bounded such that there exists a known constant $\kappa > 0$ where $\|[A\,(\Theta)\;\, B\,(\Theta)]\|_F \leq \kappa$ for all $\Theta \in \mathcal{P}_0$.*

Assumption 10 (Controllability). *For all $\Theta \in \mathcal{P}_0$, $(A(\Theta), B(\Theta))$ is controllable.*

Bounded adversarial disturbances is a common model in the adversarial online learning and control problems [12], [23], [221]. Since we make no assumptions on how large the bound on the disturbance W is, Assumption 8 models a variety of disturbance models, such as bounded and correlated stochastic noise or state-dependent disturbances such as the linearization and discretization error for nonlinear continuous dynamics [120]. Moreover, the known bound W can be relaxed to an unknown parameter η with $\eta \leq W$ for a known constant W to reduce conservatism for large W. Assumptions 9 and 10 are standard in the learning-based control

literature, e.g., see [23], [193]. We impose controllability in Assumption 10 for ease of exposition but it can be relaxed to stabilizability by adjusting the choice of model-based controller to an infinite-horizon controller such as the one proposed in [236] for the algorithm.

Stability

One of the fundamental goals for control design is to ensure stability. In this chapter, we aim to learn a stabilizing controller for the networked linear system (6.2) in the sense of input to state stability (ISS) [237]. ISS is one of the main notions of stability for both linear and nonlinear systems [238], [239]. Here we adapt the ISS definition to the ℓ^∞-norm.

Definition 6.2.1 (ISS). *A dynamical system of the form* (6.2) *is said to be input to state stable (ISS) if there exist functions* $\beta : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}_+$ *that is continuous, strictly increasing, and bijective with respect to the second argument with* $\lim_{t\to\infty} \beta(a, t) = 0$ *for all* $a \geq 0$, $t \in \mathbb{N}$, *and* $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ *that is continuous, strictly increasing, and bijective such that for all initial state* $x(0)$, *disturbance sequence* **w**, *and time* $t \geq t_0$ *for* $t_0 \in \mathbb{N}_+$, *we have* $\|x(t)\|_\infty \leq \beta(\|x(t_0)\|_\infty, t - t_0) + \gamma(\sup_{t \geq t_0} \|w(t)\|_\infty)$.

Distributed Design and Information Constraints

For large-scale networks such as the power grid with state dimension in the orders of thousands to millions, it is unrealistic and prohibitively costly for a central agent to learn a global policy online. A promising remedy is to decompose the global policy learning into a *local* one, where each subsystem in the network learns a local policy in a distributed fashion. In this work, we propose a distributed learning-based control algorithm for the networked linear system (6.2) that guarantees stability of the global system.

In addition to distributed design, networks of the form (6.1) are often modelled with additional information constraints that require careful consideration. In this work we consider two common information constraints. The first is *communication delay*, where the dynamical system is endowed with a communication network that specify delayed information transmission among subsystems. The second is *local information*, where each subsystem only computes with (delayed) local information within a specified neighborhood, and discard information outside of the neighborhood. We come back to these information constraints and present definitions in Section 6.4.

Algorithm Preliminaries

Our proposed algorithm makes use of two emerging techniques, one from the learning community, i.e., nested convex body chasing (NCBC), and one from the control community, i.e., system level synthesis (SLS). We provide important background on each below before introducing our algorithm in the next section.

Preliminaries on NCBC

The Nested Convex Body Chasing (NCBC) problem is a well-studied online learning problem [189], [191]. At every round t, the player is presented a convex body $\mathcal{K}_t \subset \mathbb{R}^n$ which is nested in the previous body, e.g., $\mathcal{K}_t \subseteq \mathcal{K}_{t-1}$. The player selects a point $q_t \in \mathcal{K}_t$ with the objective of minimizing the total path length of the selection for T rounds, e.g., $\sum_{t=0}^{T} \|q_{t+1} - q_t\|$. There are many algorithms for the NCBC problem such as greedy projection of the previously selected point onto the current body [240]. Among these, the Steiner point selector has been shown to achieve optimal competitive ratio against the offline optimal selector [191]. The Steiner point of a convex body $\mathcal{K}$ can be interpreted as the average of the extreme points and is defined as

$$\text{St}(\mathcal{K}) := \mathbb{E}_{v:\|v\|\leq 1}\left[g_{\mathcal{K}}(v)\right],$$

where $g_{\mathcal{K}}(v) := \text{argmax}_{x \in \mathcal{K}} v^\top x$ and the expectation is taken with respect to the uniform distribution over the unit ball. The Steiner point selector achieves the following total path length:

$$\sum_{t=0}^{T} \|\text{St}(\mathcal{K}_t) - \text{St}(\mathcal{K}_{t+1})\| \leq n \cdot \text{diam}(\mathcal{K}_0), \text{ for all } T \in \mathbb{N}_+. \tag{6.3}$$

We note that the Steiner point can be approximated with any accuracy by solving sampling based linear programs [189, Algorithm 3].

Distributed control via SLS

Even when the dynamics (6.1) is known, it remains challenging to design distributed and localized control policies that accommodates communication delay and information constraints due to non-convexity and computational scalability issues. In Section 2.3, we introduced the SLS framework that synthesizes distributed controllers by via convex parameterization of feedback controllers with the closed-loop mappings induced by the controllers. For ease of notation, we will denote the CLMs from $\mathbf{w}$ to $\mathbf{x}$ and $\mathbf{u}$ as $\mathbf{\Phi^x} : \mathbf{w} \to \mathbf{x}$ and $\mathbf{\Phi^u} : \mathbf{w} \to \mathbf{u}$.

Recall that by Theorem 1, for any $\mathbf{\Phi^x}, \mathbf{\Phi^u}$ satisfying (2.6), controller $\mathbf{K} = \mathbf{\Phi^u}(\mathbf{\Phi^x})^{-1}$ achieves the prescribed closed-loop responses and internally stabilizes the system with implementation (2.7).

In this chapter, we will restrict our attention to the space of linear causal operators with FIR up to horizon H, instead of the entire space of linear causal operators as presented in Theorem 1. This is because FIR allows temporal localization of disturbances, in addition to the spatial localization which can be encoded as sparsity constraints as shown in Chapter 2. Such temporal localization implies that any introduced errors from asynchronous learning will only have an effect on the system up to finite number of time steps. For completeness, we present the SLS theorem Theorem 1 specialized to linear operators with finite impulse responses (FIR) here.

Theorem 19 (Adapted from [73]). *For system* (6.2)*, any linear causal operators* $\mathbf{\Phi^x}, \mathbf{\Phi^u}$ *with finite impulse response of horizon H and satisfying the following*

$$\Phi^x[0] = I, \quad \Phi^x[k+1] = A\Phi^x[k] + B\Phi^u[k], \quad for\ k = 0, \ldots, H-1 \quad (6.4a)$$

$$\Phi^x[\tau] = 0 \quad for\ \tau \geq H \quad (6.4b)$$

are closed-loop mappings for (6.2) *under a stabilizing linear controller* $\mathbf{K}$*. Moreover, given any linear causal operators* $\mathbf{\Phi^x}, \mathbf{\Phi^u}$ *that satisfy* (6.4)*, the following SLS controller constructed using* $\mathbf{\Phi^x}, \mathbf{\Phi^u}$*,*

$$\widehat{w}(t) = x(t) - \sum_{k=1}^{H-1} \Phi^x[k]\widehat{w}(t-k) \quad (6.5a)$$

$$u(t) = \sum_{k=0}^{H-1} \Phi^u[k]\widehat{w}(t-k) \quad (6.5b)$$

with $\widehat{w}(0) = x(0)$ *achieves the desired closed-loop response prescribed by* $\mathbf{\Phi^x}, \mathbf{\Phi^u}$*.*

The horizon H is a system-dependent design parameter relating to controllability of (6.2). Under Assumption 10, $H \leq N_x$. We note that here we have shifted the indices of the kernels of the CLMs to allow for causal operators rather than strictly causal operators without loss of generality. Moreover, (6.4) provides affine constraints on finite number of nonzero parameters of the closed-loop responses. Therefore, one can tractably optimize the closed-loop responses with respect to a convex cost. A common choice is the Linear Quadratic Regulator (LQR) cost on the state and input expressed in terms of the closed-loop responses, e.g.,

$$\min_{\mathbf{\Phi^x}, \mathbf{\Phi^u}} \sum_{k=0}^{\infty} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \Phi^x[k] \\ \Phi^u[k] \end{bmatrix} \right\|_F^2 \quad \text{s.t. (6.4)}. \quad (6.6)$$

In this work, we leverage the *SLS controllers* (6.5) that is parameterized by and constructed from the operators $\mathbf{\Phi^x}$, $\mathbf{\Phi^u}$. The interpretation of (6.5) is intuitive. When $\mathbf{\Phi^x}$, $\mathbf{\Phi^u}$ satisfy (6.4), they are valid closed-loop responses, mapping $\mathbf{w}$ to $\mathbf{x}$ and $\mathbf{u}$ under (6.2). Then equation (6.5a) estimates the disturbance entering the state in the last time step by computing the difference between the currently observed state $x(t)$ and the counterfactual state $\sum_{k=1}^{H-1} \Phi^x[k]\widehat{w}(t-k)$ that should have been observed according to the closed-loop respsonse $\mathbf{\Phi^x}$ if there were no disturbance. Indeed, a simple calculation using substitution will reveal that $\widehat{w}(t) = w(t-1)$, i.e., that the estimated disturbance from an SLS controller constructed with operators that satisfy (6.4) is the perfect one-step delayed estimation of the true disturbances. Then (6.5b) computes the control action to attenuate the estimated disturbance according to the prescription of the closed-loop responses $\mathbf{\Phi^u}$.

Recall from Chapter 2 that a feature of SLS is that both the closed-loop response synthesis (6.6) and the controller implementation (6.5) can be performed in a *distributed* manner, unlike the commonly adopted optimal LQR control method via the Riccati equation [59]. This is crucial for scalability of the control algorithm for large-scale systems.

In particular (6.6) is a column separable problem, which means that we can partition matrix variables $\Phi^x[k]$, $\Phi^u[k]$ into columns such as $\Phi^x[k](:,i)$, $\Phi^u[k](:,i)$ corresponding to each subsystem i. We refer to [78] for the definition of column separability and the verification of (6.6) as a column separable problem. Thus, subsystem i only needs to solve the column subproblems corresponding to its dynamics (6.1) in the global dynamics (6.2) as follows. Let $\phi^{i,x}$ and $\phi^{i,u}$ denote the ith column of $\mathbf{\Phi^x}$ and $\mathbf{\Phi^u}$ respectively and let ϕ^i collectively stand for $\phi^{i,x}$, $\phi^{i,u}$. The ith column subproblem is

$$
\begin{aligned}
\min_{\phi^i} \quad & \sum_{k=0}^{\infty} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \phi^{i,x}[k] \\ \phi^{i,u}[k] \end{bmatrix} \right\|_F \\
\text{s.t.} \quad & \phi^{i,x}[k+1] = A\phi^{i,x}[k] + B\phi^{i,u}[k] \quad \text{for } k = 0, \ldots, H-1 \\
& \phi^{i,x}[0] = e_i, \quad \phi^{i,x}[H] = 0,
\end{aligned}
\tag{6.7}
$$

where the constraints in (6.7) is the column-wise decomposition of the constraints (6.4) for the closed-loop repsonse synthesis (6.6). It is straightforward to see that stacking the solutions to the column subproblems recovers the optimal solution to (6.6).

When the dynamics interaction among subsystems (6.1) is sparse, additional sparsity

can be imposed on the closed-loop responses during synthesis (6.6). With sparse $\mathbf{\Phi^x}$ and $\mathbf{\Phi^u}$, the implementation of the controller (6.5) can be distributed in a similar decomposition as the synthesis procedure. In particular, each subsystem computes a disjoint subset of coordinates of $\widehat{w}(t)$. Due to sparsity, such local computation for subsystem i only requires the solutions to the column subproblems from the local neighbors of i via communication instead of from the entire network.

6.3 Online Stabilization under Adversarial Sisturbances

In this section, we propose a novel online algorithm presented in Algorithm 4 that stabilizes an unknown networked linear system (6.2) under bounded and potentially adversarial disturbances. The algorithm selects hypothesis models using methods for NCBC and constructs an SLS distributed controller based on the hypothesis model. Our approach is distinguished from prior learning-based control methods in that it does not perform system identification as part of the algorithm.

We first introduce our algorithm without any communication or localization constraints. Then, in Section 6.4, we extend the algorithm to a distributed one that accommodates communication delay and local information (Algorithm 5). Though inspired by the approach in [13], Algorithms 4 and 5 are the first to consider the control goal of stabilization, which can not be subsumed under the framework proposed in [13] where only binary cost functions are considered. To cast stabilization in terms of a binary cost function, one needs to specify the largest norm of the state and control input of the closed-loop system, which is unavailable a priori[1]. Moreover, our algorithms perform both the parameter selection and the model-based control design *distributedly* for each local subsystem based on *delayed* information from other subsystems, whereas [13] is a single-agent algorithm.

Algorithm 4 starts with the construction of a set of candidate models that are consistent with the online data (line 3) after observing the latest state transition (line 2). A hypothesis model is selected from the set of candidate models with NCBC techniques (line 5) if the previously selected hypothesis model is invalidate by the new observation (line 4). Based on the selected hypothesis model, model-based control design is performed using the SLS procedure introduced in Section 2.3 (line 6 - 7). We discuss the details of Algorithm 4 in the following subsections.

---

[1]A crude approximation of the largest norm can be achieved by computing the worst-case state norm over all systems in the initial parameter set $\mathcal{P}_0$, but such approximation results in significant conservatism and requires the knowledge of control theoretical constants of the controller, e.g., SLS controllers, that may not always be available.

Algorithm 4: Online stabilization under adversarial disturbances

---

Input: Parameter set $\mathcal{P}_0$
Initialize: $t = 0$, $u(0) = 0$
1 **for** $t = 1, 2, \ldots$ **do**
2   Observe $x(t)$
     /* CONSIST: Select consistent models */
3   Construct $\mathcal{P}_t$ with (6.8)
4   **if** $\Theta_{t-1} \in \mathcal{P}_t$ **then** $\Theta_t \leftarrow \Theta_{t-1}$
5   **else** $\Theta_t \leftarrow \mathrm{St}(\mathcal{P}_t)$
     /* CONTROL: Perform model-based control with SLS */
6   Synthesize $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$ using (6.6) based on Θ_t
7   Compute $u(t)$ using the SLS controller (6.5) with $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$
8 **end**

CONSIST: Consistent Hypothesis Model Selection

The first component of Algorithm 4 is to select a hypothesis model Θ_t in order to perform model-based control. We name this component CONSIST. Due to the potentially adversarial disturbances such as state-dependent noise, standard identification methods such as linear regression do not guarantee accurate estimation of the model. Instead, we leverage NCBC for hypothesis model selection.

After observing the latest state transition from $x(t-1)$, $u(t-1)$ to $x(t)$, the algorithm constructs the set of all Θ's such that $A(\Theta)$, $B(\Theta)$ satisfy (6.2) with some admissible disturbances defined in Assumption 8. In particular, each observed transition defines a set of linear constraints on Θ and we construct the *consistent parameter set*, $\mathcal{P}_t$ at each time t using the set membership estimation (SME) method introduced in Chapter 4, namely,

$$\mathcal{P}_t := \{\Theta \in \mathcal{P}_{t-1} \ : \ \|x(t) - (A(\Theta)x(t-1) + B(\Theta)u(t-1))\|_\infty \leq W\} \tag{6.8}$$

with $\mathcal{P}_0$ as the local initial parameter set defined in Assumption 9. Recall that by construction, the consistent parameter set $\mathcal{P}_t$ is always convex, and nested within the parameter set $\mathcal{P}_{t-1}$ recursively. Moreover, $\mathcal{P}_t$ is nonempty for all $t \in \mathbb{N}_+$ because the true parameter $\Theta^\star$ belongs to every $\mathcal{P}_t$. The key property of SME is that for all t, any $\Theta_t \in \mathcal{P}_t$ could have generated the observed trajectory up to $x(t)$ and is equally likely to be the true system model. By construction, the observed state trajectory can be written as

$$x(t) = A(\Theta^\star)x(t-1) + B(\Theta^\star)u(t-1) + w^\star(t-1) \tag{6.9a}$$

$$= A(\Theta_t)x(t-1) + B(\Theta_t)u(t-1) + \widetilde{w}(t-1), \tag{6.9b}$$

where $w^\star(t)$ is the true disturbance and $\{\widetilde{w}(k)\}_{t=0}^{\infty}$ is some admissible disturbance sequence such that $\|\widetilde{w}(t)\|_\infty \leq W$. We say a model is *consistent* with observations up to time t if it belongs to $\mathcal{P}_t$. Among all consistent models, we need to select a hypothesis model Θ_t in order to perform model based control. An ideal candidate is one that can remain inside of *future* consistent parameter sets. To see why, consider an extreme case where the first selected parameter Θ_1 stays consistent for the entire online operation as we apply control actions generated based on Θ_1. Since the consistent model (6.9b) generates the same trajectory as the true model (6.9a), any guarantees that the model-based control policy has for Θ_1 will manifest in the observation. Note Θ_1 does not necessarily have to be close to $\Theta^\star$.

This intuition motivates us to select a Θ_t that could remain an element of the (yet unknown) future consistent parameter set. In particular, if the hypothesis model selected at a previous time is consistent for the current observation, we continue to use it. If the previous hypothesis model is invalidated by the new observation, then we want to select a new Θ_t's from the nested and convex body $\mathcal{P}_t$ with the objective of moving as little as possible for future bodies. This is an instance of NCBC introduced in Section 6.2. The total path length cost function in NCBC formalizes a measure of *model consistency* in our case: the less the a selector moves, the longer the selected points stay consistent overall. In Algorithm 4, we select the Steiner point of $\mathcal{P}_t$ as the hypothesis model. The finite path length guarantee of Steiner point in (6.3) can be interpreted as a finite budget for the adversarial disturbances: if the disturbances try to make the state norm large, then the selected (wrong) hypothesis model will be quickly invalidated thanks to the excitation from the disturbances. This will make CONSIST frequently re-select new hypothesis models. However, such inconsistent model selection has bounded occurrences due to the finite path length guarantee (6.3) of the Steiner point, i.e., CONSIST gains information and stops moving eventually.

CONTROL: Model-based Control with SLS

After the selection of a hypothesis model Θ_t from the consistent parameter set, Algorithm 4 performs the SLS closed-loop response synthesis (6.6) and implementation (6.5) based on Θ_t. We name this component of the algorithm CONTROL.

Distributed Implementation of Algorithm 4

Per discussion in Chapter 2, it is straightforward to see that Algorithm 4 can be implemented by each subsystem in a distributed fashion. In particular, in the CONSIST component, subsystem i constructs a local consistent parameter set $\mathcal{P}_t^i$

based on the local observations generated from the local dynamics (6.1). Subsystem i then selects the Steiner point of $\mathcal{P}_t^i$ as its local hypothesis model θ_t^i. In the CONTROL component, all subsystems collects the local hypothesis models from other subsystems and construct a global estimate $\Theta_t = (\theta_t^1, \ldots, \theta_t^N)$ since we assume no communication delay here. Based on Θ_t, each subsystem synthesizes columns of $\mathbf{\Phi_t^x}$ and $\mathbf{\Phi_t^u}$ by solving the subproblems decomposed from (6.6). After collecting and assembles the column solutions via instantaneous communication, each subsystem computes a disjoint subset of coordinates of $\widehat{w}(t)$ and $u(t)$, corresponding to the positions of the local states $x^i(t)$ and input $u^i(t)$ in the global dynamics (6.2) respectively.

Stability Guarantee

The main result in this section is the following ISS guarantee for Algorithm 4.

Theorem 20. *Under Assumption 8-10, Algorithm 4 guarantees the stability of the closed loop of (6.2) in the sense of ISS such that for all $t \geq t_0$*

$$\max\{\|x(t)\|_\infty, \|u(t)\|_\infty\} \leq O\left(e^{N_x{}^{5/2}}\right) \cdot \left(e^{-(t-t_0)/H}x(t_0) + \sup_{t_0 \leq k < t} \|w(k)\|_\infty\right),$$

where $x(t_0)$ is the initial condition, N_x is the total state dimension of the global network (6.2), and H is the finite impulse response horizon for the SLS model-based control synthesis.

We remark that the decay factor $e^{-t/H}$ corroborates the fact that H quantifies the controllability of the parameter set $\mathcal{P}_0$. Intuitively, the smaller H can be for the SLS synthesis (6.18) to be feasible, the easier the systems in the set can be learned and controlled.

Proof. The main idea of the proof is as follows. First, we characterize the closed loop dynamics of (6.2) under *any* SLS controllers constructed with arbitrary linear causal operators (Lemma 28). We then relax the original SLS condition (6.4) in Theorem 19 to a sufficient condition for ISS of the closed-loop dynamics under bounded adversarial disturbances (Lemma 29). Crucially, we show that the bounded path length property (6.3) of the selected hypothesis models in Algorithm 4 implies the satisfaction of the sufficient condition for closed-loop stability. This implication is established through a novel perturbation analysis (Theorem 21) of the SLS closed-loop response synthesis problem (6.6). We defer the proofs of the helper lemmas used here to Section 6.B.

Specifically, we show that given arbitrary $\mathbf{\Phi^x}$, $\mathbf{\Phi^u}$ with FIR horizon H, the closed-loop dynamics of (6.2) under an SLS controller constructed from $\mathbf{\Phi^x}$, $\mathbf{\Phi^u}$ is characterized as follows.

Lemma 28 (Closed-loop characterization). *The closed loop of* (6.2) *under Algorithm 4 is characterized as follows for all time $t \in \mathbb{N}$:*

$$x(t) = \sum_{k=0}^{H-1} \Phi_t^x[k]\widehat{w}(t-k), \quad u(t) = \sum_{k=0}^{H-1} \Phi_t^u[k]\widehat{w}(t-k) \tag{6.10a}$$

$$\widehat{w}(t) = \sum_{k=1}^{H} \left(A\Phi_{t-1}^x[k-1] + B\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right)\widehat{w}(t-k) + w(t-1), \tag{6.10b}$$

where A, B are the true model parameters from (6.2) *while $w(t)$ is the true unknown bounded disturbances with $\|w(t)\|_\infty \leq W$. The linear causal operators $\mathbf{\Phi_t^x}$, $\mathbf{\Phi_t^u}$ are synthesized via* (6.6) *based on the selected hypothesis model at t and $\widehat{w}(t)$ is the estimated disturbance from the SLS controller* (6.5).

This result generalizes Theorem 19 where we characterize the closed loop behaviour of SLS controllers constructed from *any* linear casual operators, not necessarily those satisfying (6.4a). Under Algorithm 4, we can further replace the true model in (6.10b) with the selected hypothesis model (Steiner point of the consistent set) Θ_t, i.e.,

$$(6.10b) = \sum_{k=1}^{H} \left(A(\Theta_t)\Phi_{t-1}^x[k-1] + B(\Theta_t)\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right)\widehat{w}(t-k) + \widetilde{w}(t-1),$$

with admissible disturbances such that $\|\widetilde{w}\|_\infty \leq W$ due to the consistency property (6.9) of Θ_t.

Moreover, Lemma 28 leads to a simple sufficient condition for stability of the closed loops under any SLS controllers. To see this, we first argue that there exist constants that bound the decay rate of the closed loop responses synthesized from (6.6). In particular, due to the finite impulse response property imposed by (6.4b) of the synthesized closed-loop responses, there always exists a large enough $C > 0$ and $\rho \in (0, 1)$ such that

$$\left\| \begin{bmatrix} \phi^{i,x}[k] \\ \phi^{i,u}[k] \end{bmatrix} \right\|_F \leq C\rho^k \quad \text{for all closed-loop responses satisfying (6.4b).}$$

This property is commonly employed in SLS-based analysis [24], [203], [220]. We use C and ρ for the sake of proof here and does not require the knowledge of them for Algorithm 4 to execute.

With the decay property, according to Lemma 28, if $\|\widehat{w}(t)\|_\infty \leq \widehat{W}_\infty$ for some $\widehat{W}_\infty > 0$, then we can bound the global state via (6.10a) as follows:

$$\|x(t)\|_\infty \leq \widehat{W}_\infty \sum_{k=0}^{H-1} \left\|\Phi_t^x[k]\right\|_\infty \leq \widehat{W}_\infty C^{1/2} {N_x}^{1/2} \frac{1}{1-\rho^{1/2}}.$$

The bound on control input $\|u(t)\|_\infty$ follows analogously. Therefore, the stability of the closed loop reduces to the boundedness of $\widehat{w}(t)$ in (6.10b). To show this, we prove the following.

Lemma 29 (Sufficient condition for H-convolution ISS). *Let $H \in \mathbb{N}_+$. For $k \in [H]$, let $\{a_t[k]\}_{t=1}^\infty$ and $\{w_t\}_{t=0}^\infty$ be positive sequences. Let $\{s_t\}_{t=0}^\infty$ be a positive sequence such that*

$$s_t \leq \sum_{k=1}^{H} a_{t-1}[k] \cdot s_{t-k} + w_{t-1}. \tag{6.11}$$

Then $\{s_t\}_{t=0}^\infty$ is ISS if $\sum_{t=0}^\infty \sum_{k=1}^H a_t[k] \leq L$ for some $L \in \mathbb{R}_+$. In particular, for all $t \geq t_0$,

$$s_t \leq e^{-(t-t_0)/H} \cdot e^L s_{t_0} + \frac{\left(e^L + e - 1\right)}{e-1} \sup_{t_0 \leq k < t} w_k. \tag{6.12}$$

The above sufficient condition is suitable for analyzing dynamical evolution under adversarial inputs. Consider taking the norm on both sides of (6.10b). Then Lemma 29 is immediately applicable with $s_t = \|\widehat{w}(t)\|_\infty$, and

$$a_t[k] = \left\|A(\Theta_t)\Phi_{t-1}^x[k-1] + B(\Theta_t)\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right\|_\infty. \tag{6.13}$$

Therefore, a sufficient condition for ISS of (6.2) under Algorithm 4 is the boundedness of (6.13) summing over time $t \in \mathbb{N}_+$ and horizon $k \leq H$. This quantity represents the total error of the implemented closed-loop responses $\mathbf{\Phi_t^x}$, $\mathbf{\Phi_t^u}$ synthesized from the selected hypothesis dynamics model Θ_t, with respect to the *correct* closed-loop responses generated from the true model $\Theta^\star$.

To bound (6.13), we make a crucial connection between the total path length of the Steiner point model selection in Algorithm 4 and (6.13). This is established via the following perturbation result for the SLS closed-loop response synthesis problem (6.6), where the formal statement (Theorem 26) and proof is presented in Appendix 6.D.

Theorem 21 (Informal, Perturbation bound). *Let $\phi^\star(A, B) := [\mathbf{x}^{\star,\top}, \mathbf{u}^{\star,\top}]^\top$ denote the concatenated optimal solution to the following optimization problem*

$$\min_{x,u} \quad \sum_{t=0}^{H} x(t)^\top Q x(t) + u(t)^\top R u(t) \tag{6.14}$$

$$s.t. \quad x(t+1) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad x(H) = 0,$$

with $Q, R > 0$. Let (A_1, B_1) and (A_2, B_2) be two system matrices such that (6.14) *is feasible. Then the corresponding optimal solutions $\phi^\star(A_1, B_1)$ and $\phi^\star(A_2, B_2)$ satisfy*

$$\|\phi^\star(A_1, B_1) - \phi^\star(A_2, B_2)\|_F \leq \Gamma \left\| \begin{bmatrix} A_1 - A_2 \\ B_1 - B_2 \end{bmatrix} \right\|_F,$$

where $\left\|\phi^\star(A, B)\right\|_F := \sum_{k=0}^{H} \|[x(k)^\top, u(k)^\top]\|_F$. Constant $\Gamma > 0$ involves the system theoretical quantities for A_1, A_2, B_1, B_2, Q, R.

The quadratic program (6.14) corresponds to the column-wise decomposed sub-problems of the SLS closed-loop response synthesis (6.6). Therefore, (6.13) can be bounded as follows:

$$(6.13) = \left\| A(\Theta_t)(\Phi_{t-1}^x[k-1] - \Phi_t^x[k-1]) + B(\Theta_t)(\Phi_{t-1}^u[k-1] - \Phi_t^u[k-1]) \right\|_\infty$$

$$\leq 2N_x\kappa \left\| \begin{bmatrix} \Phi_{t-1}^x[k-1] - \Phi_t^x[k-1] \\ \Phi_{t-1}^u[k-1] - \Phi_t^u[k-1] \end{bmatrix} \right\|_F,$$

where the equality is due to the constraint (6.4) during the model-based control step in Line 5 of Algorithm 4. The inequality invokes Assumption 9. Finally we show the total error summing (6.13) over all time step t and horizon $k \leq H$ is bounded by the total path length of the selected hypothesis models via the Steiner point.

$$\sum_{t=0}^{\infty} \sum_{k=1}^{H} (6.13) \leq 2N_x\kappa \sum_{t=0}^{\infty} \sum_{k=1}^{H} \left\| \begin{bmatrix} \Phi_{t-1}^x[k-1] - \Phi_t^x[k-1] \\ \Phi_{t-1}^u[k-1] - \Phi_t^u[k-1] \end{bmatrix} \right\|_F$$

$$\leq 2N_x^{3/2}\kappa\Gamma \sum_{t=0}^{\infty} \|\Theta_{t-1} - \Theta_t\|_F \quad \leq 2N_x^{5/2}\kappa\Gamma\,\mathrm{diam}(\mathcal{P}_0), \tag{6.15}$$

where we use Theorem 21 for the second inequality and the total path length bound (6.3) of the Steiner point selector for the last inequality. Finally, we plug the total bound (6.15) in (6.12) for an ISS bound on $\widehat{w}(t)$, which gives the desired state and control input bound in Theorem 20. $\qquad\square$

Remark 5. *NCBC algorithms other than the Steiner point selector can be substituted in Algorithm 4 as long as the finite path length guarantee (6.3) holds. Therefore, we can use a more computationally efficient algorithm with respect to the number of constraints in (6.8), such as greedy projection, at the expense of a larger worst-case path length bound. Such trade-off is potentially important since the number of constraints in (6.8) grows linearly with time. A topic of continuing work is to find an efficient representation of (6.8) that does not involve linear growth in the number of constraints.*

Comparison of Theorem 20 with previous results   Compared to the state-of-art system identification-based algorithm for online control under adversarial disturbances given in [26], which induces $\Omega(2^n)$ state and control input norm, our algorithm also incurs state norms that are exponential-polynomial in the global dimension. However, our bound is a worst-case guarantee which is on average not achieved during deployment. On the other hand, the exponential bound in [26] is qualitatively obtained, since system identification-based methods require full excitation of the system despite adversarial disturbances [26, Lemma 14]. This is the reason behind the orders of magnitude of performance improvement of our algorithm over system identification-based methods observed in the numerical study shown in Table 6.1.

Comparison of Theorem 21 with previous results   The Lipschitz continuity of optimal control problems, similar to (6.14), has been investigated in learning-based LQR literature, e.g., [241], [242]. However, our perturbation result Theorem 21 (formal statement in Theorem 26) is with regard to a finite-horizon quadratic program with terminal state constraints, whereas previous Lipschitz continuity analysis is performed with respect to the infinite-horizon LQR optimal gain. As a result, we use a different set of tools from matrix theory, unlike the Riccati equation (value function) based analysis for infinite-horizon LQR problems in previous works. In Section 6.4, we further generalize the perturbation result to handle sparsity constraints.

6.4   Adversarial Stabilization with Information Constraints

The implementation of Algorithm 4 assumes that each subsystem has instantaneous access to the information from other subsystems, such as the local consistent hypothesis models, and the column solutions to the subproblems decomposed from (6.6). Such instantaneous information sharing is often unrealistic in large-scale networked control systems. Therefore, in this section we extend the presentation

in Section 6.3 to a fully distributed algorithm, shown in Algorithm 5, that for the first time guarantees the stability of unknown interconnected LTI systems with information constraints under bounded adversarial disturbances. These results are the main contribution of this chapter.

Specifically, we consider two classes of information constraints, namely *communication delay* and *local information*, which we define formally below. After defining these information constraints, we describe the adjustments to Algorithm 4 and present our main result.

Communication Delay

A key feature of large-scale networked systems is that information observed locally at each subsystem cannot be immediately available to the global network. Instead, information sharing among subsystems is constrained by communication limitations. Such limitations often lead to delayed partial observation and pose further challenges for learning-based algorithm design [121], [205], [210]. To formalize the communication constraints, we define a communication graph $\mathcal{G}^C = (V^C, E^C)$ for (6.2), where $V^C = [N]$ and E^C is the set of directed communication link from one subsystem to the other. Self-loops at all vertices are included in E^C and they represent zero delay. The communication graph is demonstrated by the solid blue lines in Figure 6.1(a). We use $C \in \{1, 0\}^{N \times N}$ to denote the adjacency matrix associated with the communication graph $\mathcal{G}^C$. Moreover, we define the information delay induced by $\mathcal{G}^C$ as follows.

Definition 6.4.1 (Information delay)**.** *The information delay from subsystem i to j is defined to be the total distance of the shortest path from i to j according to $\mathcal{G}^C$ and is denoted as $d(i \rightarrow j)$.*

Globally, the kth power of the adjacency matrix C^k has nonzero (i, j)th entry if subsystem i gets k-delayed information from subsystem j. Locally, at time step t, subsystem i has access to subsystem j's full information up to time $t - d(j \rightarrow i)$. Moreover, $d(j \rightarrow i)$ is the smallest integer such that $C^{d(j \rightarrow i)}(i, j) \neq 0$. With slight abuse of notation, we write C^k to mean the support of the matrix so $C^k \in \{1, 0\}^{N \times N}$.

Example 7. *Consider the system in Figure 6.1(a) where the solid blue line denotes the communication among subsystems. The adjacency matrix C is depicted in Figure 6.1(c). Observe that $C(1, 3) = 0$ but $C^2(1, 3) \neq 0$. Therefore, the delay from subsystem 3 to subsystem 1 is $d(3 \rightarrow 1) = 2$.*

Given $\mathcal{G}^C$, we make a mild assumption on the communication delay. This assumption ensures that the graph describing the global dynamics is a subgraph of the communication graph. Such an assumption ensures nested information structure [243] and is commonly adopted [205], [219]. It holds true for systems where communication operates at least as fast as the dynamical propagation.

Assumption 11 (Communication Topology)**.** $C(i, j) = 1$ *for all* $j \in \mathcal{N}(i)$.

The communication delay model considered here is well-established in the distributed control literature [219], [244], [245] and is applicable to many engineering systems [246], [247]. We refer interested readers to [40] for a detailed discussion on information structures and their consequences for distributed control design. While we specify the communication delay to be synchronous with the discrete time dynamics propagation for ease of exposition, our results can be readily applied to systems with faster communication than the dynamics propagation.

Local Information

Even though communication delay causes asynchronous partial information for each subsystem, eventually each subsystem can obtain the delayed global information. However, due to the scale of the global network, it can be prohibitively costly for subsystems to compute their local control actions using such delayed global information. Moreover, a larger delay between subsystems means, intuitively, that they are more dynamically decoupled due to Assumption 11. Therefore, by discarding information from far-away subsystems, each subsystem has a smaller and more up-to-date information set. A common approach is to require each subsystem i to only use delayed information from a local neighborhood. In this work, we define three neighborhoods, $\mathcal{D}_{\text{in}}(i)$, $\mathcal{D}_{\text{out}}(i)$, and $\mathcal{M}(i)$ that subsystem i is allowed to access information from. This is sometimes referred to as localized control in multi-agent reinforcement learning [248]–[250] and distributed control [78], [210] as a method for ensuring a scalable implementation of the control policy in large-scale networked systems. Below we define each of the neighborhoods.

Definition 6.4.2 ($\bar{d}$-incoming/outgoing neighbors)**.** *The $\bar{d}$-incoming and outgoing neighbors of subsystem i according to $\mathcal{G}^C$ are respectively*

$$\mathcal{D}_{in}(i) = \{j \in [N] : d(j \to i) \leq \bar{d}\}, \quad \mathcal{D}_{out}(i) = \{j \in [N] : d(i \to j) \leq \bar{d}\}.$$

The localization parameter $\bar{d}$ is a design choice that is network structure dependent. Here we focus on the cases where the dynamics topology and communication graph

have sparse enough edges that the network structure can be leveraged to design a localization parameter $\bar{d}$ (given) that is much smaller than the size of the global network and scales well with the number of subsystems.

Definition 6.4.3 ($\bar{d}$-interaction neighbors). *The $\bar{d}$-interaction neighbors of subsystem i according to local interaction* (6.1) *and $\mathcal{G}^C$ is defined as*

$$\mathcal{M}(i) = \{\ell \in [N] : j \in \mathcal{N}(\ell) \text{ for some } j \in \mathcal{D}_{out}(i)\}.$$

The intuition behind $\mathcal{M}(i)$ is that any subsystem $\ell \in \mathcal{M}(i)$ is dynamically influence by subsystem j because $j \in \mathcal{N}(\ell)$. Furthermore, j makes local decisions such as $u^j(t)$ based on the information from subsystem i because $j \in \mathcal{D}_{out}(i)$. Therefore, it is sensible for subsystem i to take the information from ℓ into consideration during decision making, since ℓ will be indirectly affected by decisions made at i through information sharing and dynamical interaction via j.

Finally, we make the following feasibility assumption.

Assumption 12 (Feasibility). *For all $\Theta \in \mathcal{P}_0$, there exists a stabilizing controller for $A(\Theta), B(\Theta)$ such that each agent with local dynamics* (6.1) *uses delayed and locally available information from its $\bar{d}$-interaction, incoming, and outgoing neighbors according to $\mathcal{G}^C$.*

Assumption 12 ensures the well-posedness of the distributed controller learning problem and is commonly employed [60], [109], [251]. If a parameter set $\mathcal{P}_0$ has a few singular points where (A, B) loses feasibility such as when $B = 0$, a simple heuristic is to ignore these points in the algorithm since we assume the underlying system is controllable. We discuss the case of non-convex parameter sets in Section 6.E.

A Fully Distributed and Localized Algorithm

We now describe how to extend Algorithm 4 to handle communication delay and localized control constraints. To do this we add additional information exchange steps to Algorithm 4 in each of the two components. The full algorithm is shown in Algorithm 5. For ease of exposition, we let the subsystems have scalar state and fully actuated control actions ($N_x = N_u = N$) in order to minimize notation. It is straightforawrd to generalize the presented algorithm and analysis to vector subsystems.

Algorithm 5: Distributed online stabilization under information constraints

Input: Parameter set $\mathcal{P}_0$

Initialize: $t = 0$, $u(0) = 0$, $\mathcal{I}(i, 0) = \emptyset$ for $i \in [N]$

1 **for** $t = 1, 2, \ldots$ **do**

2   **for** *Subsystem $i = 1, 2, \ldots, N$* **do**

3   Observe $x^i(t)$

         `/* CONSIST: Select consistent models */`

4   Construct $\mathcal{P}_t^i$ with (6.16)

5   **if** $\theta_{t-1}^i \in \mathcal{P}_t^i$ **then** $\theta_t^i \leftarrow \theta_{t-1}^i$

6   **else** $\theta_t^i \leftarrow \mathrm{St}(\mathcal{P}_t^i)$

         `/* CONTROL: Perform model-based control with SLS */`

7   Assemble local estimate of the global model $A\left(\widehat{\Theta}_t^i\right), B\left(\widehat{\Theta}_t^i\right)$ with (6.17)

8   Synthesize closed-loop response columns ϕ_t^i using (6.18) based on $A\left(\widehat{\Theta}_t^i\right), B\left(\widehat{\Theta}_t^i\right)$

9   Assemble delayed local column solutions $\bigcup_{j \in \mathcal{D}_{\mathrm{in}}(i)} \phi_{t-d(j \to i)}^j$

10   Compute local control action $u^i(t)$ using (6.19) with the assembled column solutions

11   **end**

12 **end**

CONSIST

This component of Algorithm 5 is identical to that of the distributed implementation of Algorithm 4 discussed in Line 8. Formally, subsystem i constructs the *local consistent parameter set*, $\mathcal{P}_t^i$ according to local dynamics (6.1) as

$$
\mathcal{P}_t^i := \left\{ \theta^i \in \mathcal{P}_{t-1}^i \ : \ \left\| x^i(t) - \left(\sum_{j \in \mathcal{N}(i)} A^{ij}(\theta^i) x^j(t-1) + B^{ij}(\theta^i) u^j(t-1) \right) \right\|_\infty \leq W \right\}
$$

(6.16)

with $\mathcal{P}_0^i$ as the local initial parameter set defined in Assumption 9. The communication delay pattern allows the construction of $\mathcal{P}_t^i$ because each subsystem i precisely has access to $x^j(t-1)$ and $u^j(t-1)$ from its immediate dynamical interaction neighbors $\mathcal{N}(i)$ by Assumption 11.

Analogous to Algorithm 4, each subsystem i selects the Steiner point of $\mathcal{P}_t^i$ as the *local hypothesis model* if the previous selection is invalidated by the latest observation.

CONTROL

Since the local hypothesis models are no longer shared instantly among subsystems due to the communication delay and local information constraints, we modify the model-based control component of Algorithm 4 and carefully keep track of the available information. To give an overview, at every step t, subsystem i first assembles a local estimate of the "global" model using delayed information from other subsystems (line 7). Based on the estimated global model, subsystem i synthesizes the ith column of the SLS closed-loop responses by solving the column subproblem of (6.6) as discussed in Line 8 (line 8). Then, subsystem i assembles a local SLS controller with the local column solutions ϕ_t^i computed from the previous step and the delayed column solutions from other subsystems (line 9). Finally, the local control action is computed using the locally assembled SLS controller (6.19) (line 10).

Local estimate of the global model (line 7). After selecting a local hypothesis model, Subsystem i assembles a local estimate of the "global" parameter by collecting the available (delayed) local hypothesis models from its neighbors in $\mathcal{M}(i)$,

$$\widehat{\Theta}_t^i := \left(\theta_{t-d(j\to i)}^j \right)_{j\in\mathcal{M}(i)}, \tag{6.17}$$

where the local neighborhood $\mathcal{M}(i)$ (Definition 6.4.3) represents the the set of neighbors whose model information i needs for synthesizing its local column solution later in (6.18).

Local column synthesis (line 8). Analogous to line 6 in Algorithm 4, subsystem i now performs model-based control via SLS by solving the column subproblem (6.7) with additional communication delay and local information constraints based on the locally estimated "global" parameter $\widehat{\Theta}_t^i$. It is well-established that information constraints described in Section 6.4 becomes convex sparsity constraints on $\mathbf{\Phi^x}$ and $\mathbf{\Phi^u}$ [45]. In particular, these information constraints can be represented as binary matrices C^k (for delay) and $C^{\bar{d}}$ (for local information) with $k \in [H]$. Now, the column subproblem for subsystem i changes from (6.7) to

$$\min_{\phi_t^{i,x}, \phi_t^{i,u}} \quad \sum_{k=0}^{\infty} \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \phi_t^{i,x}[k] \\ \phi_t^{i,u}[k] \end{bmatrix} \right\|_F \tag{6.18a}$$

$$\text{s.t.} \quad \phi_t^{i,x}[k+1] = A\left(\widehat{\Theta}_t^i\right)\phi_t^{i,x}[k] + B\left(\widehat{\Theta}_t^i\right)\phi_t^{i,u}[k], \quad \text{for } k \in [H-1] \tag{6.18b}$$

$$\phi_t^{i,x}[0] = e_i, \quad \phi_t^{i,x}[H] = 0 \tag{6.18c}$$

$$\phi_t^{i,x}[k], \ \phi_t^{i,u}[k] \in C^k(:,i) \cap C^{\bar{d}}(:,i), \quad \text{for } k \in [H-1], \tag{6.18d}$$

where (6.18a)-(6.18b) are the same LQR cost and closed-loop response characterization in (6.7). The communication and local information constraints are introduced via (6.18d). We refer interested readers to [235], [236] for a standard derivation on how (6.18d) is equivalent to the information constraints specified in Section 6.4. The problem (6.18) is always feasible due to Assumption 10 and 12.

Delay in the local parameter information results in differently synthesized columns of different $\mathbf{\Phi^x}, \mathbf{\Phi^u}$ for different subsystems. This contrasts Algorithm 4 where all subsystems use the same global model as input to the local synthesis problems and output a column of the same $\mathbf{\Phi^x}, \mathbf{\Phi^u}$.

Asynchronous closed-loop response assembly (line 9). Once local closed-loop columns are synthesized, subsystem i has to assemble other relevant columns from subsystem j from $\mathcal{D}_{\text{in}}(i)$ in order to perform the downstream task of local control action computation via the local version of the SLS controller (6.5), shown in (6.19). In particular, (6.19) requires the ith element of every column j such that $C^{\bar{d}}(i, j) \neq 0$. By definition, $\mathcal{D}_{\text{in}}(i)$ (Definition 6.4.2) is the set of j's such that $C^{\bar{d}}$ has nonzero (i, j)th element. Thus, only closed-loop columns from $j \in \mathcal{D}_{\text{in}}(i)$ are required. The assembled closed-loop responses for each subsystem has asynchronous columns with varying delays.

Local Control Action Computation (line 10). The final step in CONTROL is to compute a local control action, where each subsystem i plugs the assembled closed-loop responses into the SLS controller (6.5). Due to the sparsity constraints (from information constraints) enforced on the column solutions during the synthesis (6.18), the matrix-vector computation in (6.5) does not require the entire network's delayed column solution. Instead, subsystem i computes a local version of (6.5),

$$\widehat{w}^i(t) = x^i(t) - \sum_{j \in \mathcal{D}_{\text{in}}(i)} \sum_{k=1}^{H-1} \phi^{j,x}_{t-d(j \to i)}[k](i) \cdot \widehat{w}^j(t-k) \tag{6.19a}$$

$$u^i(t) = \sum_{j \in \mathcal{D}_{\text{in}}(i)} \sum_{k=0}^{H-1} \phi^{j,u}_{t-d(j \to i)}[k](i) \cdot \widehat{w}^j(t-k), \tag{6.19b}$$

where $x^i(t), u^i(t), \widehat{w}^i(t) \in \mathbb{R}$ are the local state, control action, and estimated disturbance respectively. The local controllers are initiated with $\widehat{w}^i(0) = x^i(0)$. Similar to the global controller (6.5), the intuition behind (6.19) is that each subsystem i *counterfactually* assumes that the global closed loop of (6.2) behaves exactly as the columns $\phi^j_{t-d(j \to i)}$ prescribe. In particular, the ith position of the jth column solution $\boldsymbol{\phi}^j_{t-d(j \to i)}$ maps the jth position of $\mathbf{w}$ ($\mathbf{w}^j$) to the ith position of $\mathbf{x}$

and $\mathbf{u}$ ($\mathbf{x}^i$ and $\mathbf{u}^i$). Therefore, (6.19a) estimates the local disturbances by comparing observed local state $x^i(t)$ and the counterfactual state computed with $\phi^j_{t-d(j\to i)}$'s. Then (6.19b) acts upon the computed disturbance.

In this step, the errors caused by the delayed information propagate further during (6.19) when each subsystem computes control action using the assembled closed-loop column solutions from different sets of sub-controllers in (6.17). This contrasts the setting in Algorithm 5, where without communication delay, all subsystems use the globally agreed closed-loop operators $\mathbf{\Phi^x},\mathbf{\Phi^u}$ to compute the local control action using (6.5).

Thanks to (6.18d), regardless of the delay, all closed-loop columns has the correct sparsity required by the communication and locality constraints. Consequently, any assembled closed loop columns used for (6.19) at each subsystem preserve the required sparsity. Therefore, the SLS controller implemented with these column solutions conforms to the information constraints.

Stability Guarantee

We now present the main result of this chapter. This is the first stabilization result for a distributed policy (Algorithm 5) in a networked setting with unknown dynamics, communication delay, local information constraint, and adversarial disturbances.

Theorem 22 (Stability). *Under Assumptions 8-12, Algorithm 5 guarantees the ISS of the closed loop of (6.2) such that for all $t \geq t_0$,*

$$\max\{\|x(t)\|_\infty, \|u(t)\|_\infty\} \leq O\left(e^{(\bar{n})^{9/2}\bar{d}}\right)\left(e^{-(t-t_0)/H}x(t_0) + \sup_{t_0\leq k\leq t}\|w(k)\|_\infty\right),$$

where $x(t_0)$ is the initial condition, local dimension $\bar{n} = \max\{\|C^{\bar{d}}\|_1, \|C^{\bar{d}}\|_\infty, \max_j |\mathcal{M}(j)|\}$ represents the total state dimension in the $\bar{d}$-neighborhood specified by the dynamics interaction (6.1) and the communication graph $\mathcal{G}^C$. Parameter $\bar{d}$ is the largest local delay each subsystem allows for delayed information, and H is the SLS closed-loop response finite impulse horizon.

Theorem 22 highlights that only the local constants $\bar{d}$ and $\bar{n}$ impact the stability guarantee, in contrast to the dependence on the global network dimension in Algorithm 4 and in system-identification based approaches [26]. Further, the result makes explicit that communication delay adds an exponential factor of error on the state deviation from the desired steady state compared to Theorem 20. When the

network connectivity is sparse, local constants $\bar{n}$ and $\bar{d}$ can remain small even if the number of subsystems in the network is large and growing [235], [252].

Proof Outline. The proof of Theorem 22 follows a similar structure as that of Theorem 20. We defer formal proofs to Section 6.C. The main challenge here is to characterize the error caused by asynchronous information at different subsystems throughout the algorithm due to delay.

To begin, we use Lemma 28 and show that despite the fact that each subsystem in Algorithm 5 uses differently delayed information to compute the local parameter, sub-controller, and control actions, the closed loop for the global system under such distributed policy can be characterized with a simple global representation. In particular, denote the actual closed-loop response implemented by Algorithm 5 as $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$. By observation, each element of $\Phi_t^x[k], \Phi_t^u[k]$ is

$$\Phi_t^x[k](i,j) := \phi_{t-d(j\to i)}^{j,x}[k](i), \quad \Phi_t^u[k](i,j) := \phi_{t-d(j\to i)}^{j,u}[k](i).$$

Therefore, the closed loop of (6.2) under Algorithm 5 can be characterized by (6.10) with $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$. It follows from Lemma 29 that as long as the error term

$$\sum_{t=1}^{\infty} \sum_{k=1}^{H} \left\| A(\Theta_t)\Phi_{t-1}^x[k-1] + B(\Theta_t)\Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right\|_{\infty} \tag{6.20}$$

is bounded, then the closed loop is ISS. Here Θ_t is the consistent global model constructed from the *local* consistent hypothesis models selected by all subsystems at time t. In Section 6.C, we quantify the effect of delay that manifests in $\mathbf{\Phi_t^x}$ and $\mathbf{\Phi_t^u}$.

To bound (6.20), we extend the perturbation bound in Theorem 21 to accommodate the additional sparsity constraints in (6.18) (Corollary 23.1). This result allows us to make a connection between (6.20) and the total path length of each subsystem's local parameter selection. Furthermore, Corollary 23.1 has potential application for a class of SLS-based distributed and localized MPC problems [210], [211].

6.5 Simulation

The main contribution of this work focuses on deriving a stability guarantee for the proposed method under adversarial disturbances and information constraints. In this section, we provide a preliminary numerical exploration of the performance improvement of our approach compared the state-of-the-art adversarial control method in the single-agent case in Section 6.5. We further test our method on a mesh network of discretized swing dynamics for power systems, where we demonstrate

near-optimal performance of Algorithm 4 and Algorithm 5 compared to the offline optimal controller synthesized according to the true dynamics in Section 6.5. Further, we study the effect of the localization parameter and the network size under correlated Gaussian noise.

Single-agent: Double Integrator Dynamics

We consider the classic double integrator dynamics [253],

$$\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}(t+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u(t) + \begin{bmatrix} w^1 \\ w^2 \end{bmatrix}(t),$$

where $x(t) = [x^1, x^2]^\top(t) \in \mathbb{R}^2$, $u(t) \in \mathbb{R}$. Disturbance $w(t) \in \mathbb{R}^2$ is the bounded ($\|w(t)\|_\infty \leq 1$). The system models a unit mass vehicle with position (x^1) and velocity (x^2) as its state under force u.

To the best of our knowledge, the only online algorithm that guarantees stability under bounded adversarial disturbances is [26], where system identification is performed before a certainty-equivalent controller is synthesized based on the estimated dynamics. Therefore, we study the performance of our algorithm and that of [26]. The results are summarized in Table 6.1, where we report the averaged maximum and top 90% state deviation from origin, i.e. $\max_t \|x(t)\|_\infty$ across 10 runs under three different disturbance profiles. In particular, we generate correlated (across coordinates) Gaussian noise projected to -1 and 1, the uniform disturbance, and the projected state-dependent adversarial disturbance, where the adversary chooses $w(t) = \text{sign}\left(A(\Theta^\star)x(t) + B(\Theta^\star)u(t)\right)$.

To instantiate [26], we use exact system theoretical constants required for the algorithm and perform the black-box system identification algorithm in [26, Algorithm 2] with identification accuracy set to be 10^{-2} (largest error tolerable by the algorithm). Then, we generate a stabilizing controller with [26, Algorithm 3]. For the proposed approach, we use the optimal LQR feedback gain in place of the centralized SLS controller (6.6) and (6.5), since under Assumption 10, the SLS controller synthesized under with LQR cost is equivalent to the optimal LQR feedback [73]. We remark that for all disturbance profiles and regardless of the choice of stabilizing controller, the system identification algorithm of [26] always requires control inputs in the order of 10^{11}. Therefore, across all disturbances, the trajectories generated by [26] are nearly identical.
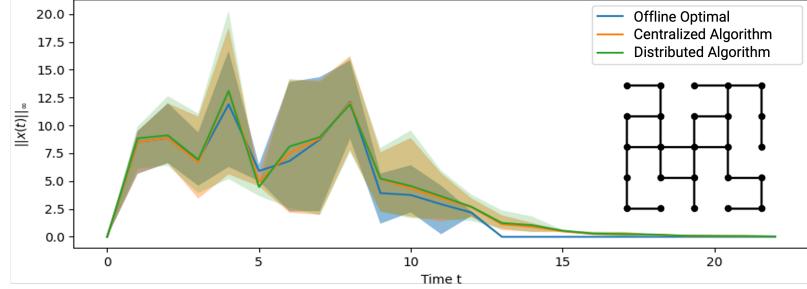
Figure 6.2: State trajectory of the optimal distributed controller, Algorithm 4, and Algorithm 5 for the 5×5 mesh network.

Multi-agent:Discretized Swing Dynamics in Power Systems

We now consider a power network with randomly generated sparse edges representing dynamical interactions over a 5 by 5 mesh, where each vertex represents a bus, illustrated at the right lower corner of Figure 6.2. The local dynamics at bus i is given by the two-state discretized swing equations [73],

$$x^i(t+1) = \begin{bmatrix} 1 & \Delta t \\ -\frac{\sum_{j \in \mathcal{N}(i)} k_{ij}}{m_i} \Delta t & 1 \end{bmatrix} x^i(t) + \sum_{j \in \mathcal{N}(i)} \begin{bmatrix} 0 & 0 \\ -\frac{k_{ij}}{m_i} \Delta t & 0 \end{bmatrix} x^j(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \left(u^i(t) + w^i(t) \right)$$

where the states are the phase angle (first state) and frequency (second state) deviation from the set point (origin), $\Delta t = 0.1$s is the discretization time step, and m_i, k_{ij}, u^i, w^i, are the inertia, line susceptance between bus i and j, control action, and external disturbance respectively. We assume each bus has a phase measurement unit and a frequency sensor to measure x^i.

We randomly generate each $k_{ij} \in [0.1, 1]$ and m_i between $[0.1, 10]$, and assume these parameters are unknown to the algorithm except their bounds. The global network is generated to be open-loop unstable. We use correlated (across buses) Gaussian disturbances with a known bound. In Figure 6.2 we compare the performance of Algorithm 4 (information shared globally and without delay), Algorithm 5, and the offline optimal distributed SLS controller synthesized from (6.6) with the knowledge of k_{ij}'s and m_i's, all subject to the same distributed control design requirements. Specifically, the communication network is assume to be the same as the dynamical interaction mesh graph, and we choose the localization parameter to be $\bar{d} = 3$, which is much smaller compared to the network size of 25. The centralized algorithm where no communication delay is present matches closely with the trajectory generated by the offline optimal controller, whereas the presence of the information constraints for Algorithm 5 degrades the performance. However, we highlight that despite the exponential dependency on the local dimensions in

Theorem 22, the actual performance of Algorithm 5 in this case is significantly better than the theoretical guarantee.

Localization parameter	Mean	Top 95%	Network Size	Mean	Top 95%
$\bar{d} = 3$	3.98	14.02	$N = 9$	2.96	10.28
$\bar{d} = 5$	3.85	14.18	$N = 25$	3.98	14.02
$\bar{d} = 10$	4.19	14.08	$N = 36$	4.27	14.05

Table 6.2: Comparison of the state norm ($\|x(t)\|_\infty$) for different localization parameters d on the 5 by 5 network (left) and comparison for different network sizes with N agents with fixed localization parameter $d = 3$ (right).

Furthermore, we compare the effects of different localization parameter choices. On the one hand, larger $\bar{d}$ results in larger worst-case guarantee in Theorem 22 due to delayed information for local computation. On the other, larger $\bar{d}$ means that each agent in the network can access more (delayed) information. This trade-off manifests on the left of Table 6.2, where $\bar{d} = 5$ appears to achieve lower average state norm over 4 random runs with correlated Gaussian noises, slightly outperforming controllers with $\bar{d} = 3$ (too little information) and $\bar{d} = 10$ (too much delay from far-away neighbors). On the right of Table 6.2, we corroborate Theorem 22 where the stability guarantee only depends on local constants $\bar{d}$ and $\bar{n}$. We randomly generate 3x3, 5x5, and 6x6 mesh networks of similar network structure, and the resulting state norm does not scale with the network size.

6.6 Conclusion

In this chapter, we instantiated the SME uncertainty set-based learning and control framework and propose the first learning-based algorithm that provably achieves online stabilization for networked LTI systems subject to communication delays under adversarial disturbances. We leverage nested convex body chasing and distributed control. The novel approach achieves orders of magnitude of improvement in performance over state-of-the-art methods for single-agent systems and handles information delays in networked multi-agent systems. Since most systems are time-varying in nature, an immediate extension of this work is to combine general convex body chasing and model-based control methods to handle time-varying dynamical systems. Future directions include extending the communication model to incorporate stochastic and time-varying delays among agents.

6.A Notation Summary

Table 6.3: Notations and definitions for the model setup, algorithms, and proofs

Notation	Meaning
$x^i(t), u^i(t), w^i(t)$	State ($\mathbb{R}^{n_i}$), action ($\mathbb{R}^{m_i}$), disturbances ($\mathbb{R}^{n_i}$) at subsystem i;
$\mathcal{N}(i)$	Dynamical neighbors of subsystem i where $x^j(t-1)$ affects $x^i(t)$ for $j \in \mathcal{N}(i)$;
$x(t), u(t), w(t)$	Global state, control action, and disturbance vector concatenated from the local ones in (6.1);
A^{ij}, B^{ij}	Local dynamics matrices describing how states and control action of subsystem j affects subsystem i for $j \in \mathcal{N}(i)$ in (6.1);
A, B	Concatenated global dynamics matrices from A^{ij}'s and B^{ij}'s ;
θ^i	The parameters for the nonzero locations in local dynamics matrices and we write $A^{ij}(\theta^i), B^{ij}(\theta^i)$. In particular, $\theta^i \cap \theta^j = \emptyset$ for all $i \neq j$;
Θ	The concatenated local parameters for the global dynamics with $\Theta := \bigcup_{i \in [N]} \theta^i$;
$\mathcal{P}_0$	The known initial compact convex parameter set where the true dynamics parameter lies;
$\mathcal{G}^C$	Communication graph defined over system (6.2) with vertices V^C corresponding to subsystems and directed edges E^C;
C	The adjacency matrix of $\mathcal{G}^C$;
$d(i \rightarrow j)$	Communication delay from subsystem i to subsystem j defined as the graph distance from i to j according to $\mathcal{G}^C$;
$\mathcal{D}_{\text{in}}(i)$	$\bar{d}$-incoming neighbors of subsystem i where $\mathcal{D}_{\text{in}}(i) := \{j \in [N] : d(j \rightarrow i) \leq \bar{d}\}$. In particular, $j \in \mathcal{D}_{\text{in}}(i)$ if $C^{\bar{d}}(i,j) \neq 0$;
$\mathcal{D}_{\text{out}}(i)$	$\bar{d}$-outgoing neighbors of subsystem i where $\mathcal{D}_{\text{out}}(i) := \{j \in [N] : d(i \rightarrow j) \leq \bar{d}\}$. In particular, $j \in \mathcal{D}_{\text{in}}(i)$ if $C^{\bar{d}}(j,i) \neq 0$;
$\mathcal{M}(i)$	Subsystems whose model information is needed for sub-controller synthesis at subsystem i with Algorithm 5 where $\mathcal{M}(i) = \{\ell \in [N] : j \in \mathcal{N}(\ell) \text{ for some } j \in \mathcal{D}_{\text{out}}(i)\}$;
$\bar{d}$-neighbor of i	The union of all subsystems in $\mathcal{D}_{\text{in}}(i), \mathcal{D}_{\text{out}}(i), \mathcal{M}(i)$;
$\mathcal{P}_t^i$	Local consistent parameter set constructed by subsystem i at time t with (6.16);
θ_t^i	Local consistent parameter for subsystem i for A^{ij} and B^{ij} constructed with Algorithm 5;
$\widehat{\Theta}_t^i$	The assembled local estimate of the "global" parameter where $\widehat{\Theta}_t^i := \bigcup_{j \in \mathcal{M}(i)} \theta_{t-d(j \rightarrow i)}^j$;
ϕ_t^i	Local column solutions generated by subsystem i at time t from (6.18);
$\phi_t^{i,x}, \phi_t^{i,u}$	The x and u components of ϕ_t^i, respectively. They are synthesized from (6.18);
Θ_t	The collection of all local consistent parameters at time t where $\Theta_t = \bigcup_{i=1}^N \theta_t^i$;
$A_t, B_t, \widetilde{w}(t)$	The global consistent matrices $A(\Theta_t), B(\Theta_t)$, and corresponding admissible disturbance;
a_t^i, b_t^i	The ith row of A_t, B_t respectively;
$\widehat{w}(t)$	Concatenated global estimated disturbance from $\widehat{w}^i(t)$ in (6.19);
$\mathbf{\Phi_t^x}$	Concatenated global closed loop operators where $\Phi_t^x[k](i,j) := \phi_{t-d(j \rightarrow i)}^{j,x}[k](i)$ from (6.19) ;
$\mathbf{\Phi_t^u}$	Concatenated global closed loop operators where $\Phi_t^u[k](i,j) := \phi_{t-d(j \rightarrow i)}^{j,u}[k](i)$ from (6.19) ;

6.B Proofs for Section 6.3

Below we restate and prove the auxiliary results needed for the proof of Theorem 20 in Section 6.3.

Table 6.4: Constants used throughout the chapter

Constants	Meaning
N	Number of subsystems in the global dynamics (6.2);
n_i, m_i	Local state and control action dimension for subsystem i in (6.1);
n_x, n_u	Global state and control dimension with $n_x = \sum_{i=1}^{N} n_i$ and $n_u = \sum_{i=1}^{N} m_i$;
W	The known bound on the true disturbances such that $\|w(t)\|_\infty \leq W$;
κ	The bound on all possible system matrices where $\|A(\Theta)\|_2, \|B(\Theta)\|_2 \leq \kappa$ for all $\Theta \in \mathcal{P}_0$;
$\bar{d}$	The localization parameter such that each subsystem is constrained to only use information from its $\bar{d}$-neighbors in Algorithm 5;
$\bar{n}$	The largest total local state dimension for the $\bar{d}$-neighbors of the subsystems where $\bar{n} = \max\{\|C^{\bar{d}}\|_1, \|C^{\bar{d}}\|_\infty, \max_j |\mathcal{M}(j)|\}$;
C, ρ	The decay rate for the closed-loop columns ϕ_t^i synthesized in (6.18) such that $\left\|\phi_t^i[k]\right\|_2 \leq C\rho^k$;

Lemma 30 (Closed loop Dynamics). *The closed loop of* (6.2) *under Algorithm 4 is characterized as follows for all time $t \in \mathbb{N}_+$:*

$$x(t) = \sum_{k=0}^{H-1} \Phi_t^x[k]\widehat{w}(t-k), \quad u(t) = \sum_{k=0}^{H-1} \Phi_t^u[k]\widehat{w}(t-k) \tag{6.21a}$$

$$\widehat{w}(t) = \sum_{k=1}^{H} \left(A(\Theta_t)\Phi_{t-1}^x[k-1] + B(\Theta_t)\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right)\widehat{w}(t-k) + w(t-1), \tag{6.21b}$$

where A, B are the true model parameters from (6.2) *while $w(t)$ is the true unknown bounded disturbances with $\|w(t)\|_\infty \leq W$. The linear causal operators $\mathbf{\Phi_t^x}$, $\mathbf{\Phi_t^u}$ are synthesized via* (6.6) *based on the selected hypothesis model at t and $\widehat{w}(t)$ is the estimated disturbance from the SLS controller* (6.5).

Proof. First, we write out the global closed-loop dynamics of (6.2) under the SLS controller (6.5) with the synthesized closed-loop responses,

$$x(t) = A\left(\Theta^\star\right)x(t-1) + B\left(\Theta^\star\right)u(t-1) + w(t-1) \tag{6.22a}$$

$$\widehat{w}(t) = x(t) - \sum_{k=1}^{H-1} \Phi_t^x[k]\widehat{w}(t-k) \tag{6.22b}$$

$$u(t) = \sum_{k=0}^{H-1} \Phi_t^u[k]\widehat{w}(t-k), \tag{6.22c}$$

where (6.22a) is the global dynamics (6.2) while (6.22b) and (6.22c) are the implemented SLS controller. Now, we use the consistency property of all the consistent

hypothesis model Θ_t selected by Algorithm 4 and represent dynamics (6.22a) in terms of the global consistent parameter $A_t := A(\Theta_t), B_t := B(\Theta_t)$,

$$x(t) = A_t x(t-1) + B_t u(t-1) + \widetilde{w}(t-1), \tag{6.23}$$

with admissible consistent disturbances $\|\widetilde{w}(t)\|_\infty \leq W$ for all time t. The replacement of $(A(\Theta^\star), B(\Theta^\star)), w(t))$ with $(A_t, B_t, \widetilde{w}(t))$ is by definition of the consistent set (6.8). Next, observe that by moving $x(t)$ to the left side, (6.22b) becomes

$$
\begin{aligned}
x(t) &= \sum_{k=1}^{H-1} \Phi_t^x[k]\widehat{w}(t-k) + \widehat{w}(t) \\
&= \sum_{k=0}^{H-1} \Phi_t^x[k]\widehat{w}(t-k),
\end{aligned} \tag{6.24}
$$

where in the last equality we used the fact that each $\Phi_t^x[0] = I$ by the constraint (6.4). Now we substitute (6.23) into (6.22b) to get

$$
\widehat{w}(t) = x(t) - \sum_{k=1}^{H-1} \Phi_t^x[k]\widehat{w}(t-k) \tag{6.25a}
$$

$$
= A_t x(t-1) + B_t u(t-1) - \sum_{k=1}^{H-1} \Phi_t^x[k]\widehat{w}(t-k) + \widetilde{w}(t-1) \tag{6.25b}
$$

$$
= A_t \sum_{k=0}^{H-1} \Phi_{t-1}^x[k]\widehat{w}(t-1-k) + B_t \sum_{k=0}^{H-1} \Phi_{t-1}^u[k]\widehat{w}(t-1-k) - \sum_{k=1}^{H-1} \Phi_t^x[k]\widehat{w}(t-k)
$$
$$
+ \widetilde{w}(t-1) \tag{6.25c}
$$

$$
= \sum_{k=1}^{H-1} \left(A_t \Phi_{t-1}^x[k-1] + B_t \Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right) \widehat{w}(t-k)
$$
$$
+ \left(A_t \Phi_{t-1}^x[H-1] + B_t \Phi_{t-1}^u[H-1] - \Phi_{t-1}^x[H] \right) \widehat{w}(t-H) + \widetilde{w}(t-1) \tag{6.25d}
$$

$$
= \sum_{k=1}^{H} \left(A_t \Phi_{t-1}^x[k-1] + B_t \Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right) \widehat{w}(t-k) + \widetilde{w}(t-1), \tag{6.25e}
$$

where in (6.25c) we substituted (6.24) and (6.22c) into $x(t-1)$ and $u(t-1)$ respectively. In (6.25d), we grouped the terms according to $\widehat{w}(t-k)$ and used the fact that the closed-loop responses are synthesized in (6.6) such that $\Phi_{t-1}^x[H] = 0$ for all t. Together, (6.22c),(6.24), and (6.25e) are as requested. $\qquad\square$

Lemma 31 (Sufficient condition for H-convolution ISS). *Let $H \in \mathbb{N}$. For $k \in [H]$, let $\{a_t[k]\}_{t=1}^{\infty}$ and $\{w_t\}_{t=1}^{\infty}$ be positive sequences. Let $\{s_t\}_{t=0}^{\infty}$ be a positive sequence such that*

$$s_t \leq \sum_{k=1}^{H} a_{t-1}[k] \cdot s_{t-k} + w_{t-1}. \tag{6.26}$$

Then $\{s_t\}_{t=0}^{\infty}$ is bounded if $\sum_{t=0}^{\infty} \sum_{k=1}^{H} a_t[k] \leq L$ for some $L \in \mathbb{R}_+$. In particular, for all $t \geq t_0$,

$$s_t \leq e^{-(t-t_0)/H} \cdot e^L s_{t_0} + \frac{(e^L + e - 1)}{e - 1} \sup_{t_0 \leq k < t} w_k.$$

Proof. Fix t_0 and $t \geq t_0$. Denote $\{z_{t_i}\}$ as a finite subsequence of $\{s_\tau\}_{\tau=t_0}^{t}$ such that

$$z_{t_N} = s_t$$

$$z_{t_{i-1}} = \max_{t_i - H \leq \tau \leq t_i - 1} s_\tau, \quad \text{for } i = N, N-1, \ldots, 1,$$

with $t_N = t$ and $z_{t_i} = s_{t_i}$. This construction of the $\{z_{t_i}\}$ has to terminate at $z_{t_0} = s_{t_0}$. Therefore, N is at least $\frac{(t-t_0)}{H}$ and at most $t - t_0$. By the recursive relationship of s_t in (6.26), we have for any i,

$$\begin{aligned}
z_{t_i} = s_{t_i} &\leq \sum_{k=1}^{H} a_{t_i-1}[k] s_{t_i-k} + w_{t_i-1} \\
&\leq \left(\sum_{k=1}^{H} a_{t_i-1}[k] \right) z_{t_i-1} + w_{t_i-1} \\
&= \widehat{a}_{t_i-1} \cdot z_{t_i-1} + w_{t_i-1},
\end{aligned} \tag{6.27}$$

where we use the fact that $a_t[k] \geq 0$ for all t and k. We also denote $\widehat{a}_{t_i-1} = \left(\sum_{k=1}^{H} a_{t_i-1}[k] \right)$ for the last equality. By the recursion (6.27), we have

$$s_t = z_{t_N} \leq \prod_{i=1}^{N} \widehat{a}_{t_i-1} \cdot z_{t_0} + \left(\sup_{t_0 \leq k < t} w_k \right) \left(1 + \sum_{j=1}^{N} \prod_{i=j}^{N} \widehat{a}_{t_i-1} \right). \tag{6.28}$$

Now, $\prod_{i=j}^{N} \widehat{a}_{t_i-1} = \prod_{i=j}^{N} \left((\widehat{a}_{t_i-1} - 1) + 1 \right) \leq \prod_{i=j}^{N} e^{\widehat{a}_{t_i-1}-1} = e^{\sum_{i=j}^{N}(\widehat{a}_{t_i-1}-1)} \leq e^{L-(N-j+1)}$, where the last inequality is due to the hypothesis that $\sum_{t=0}^{\infty} \widehat{a}_t \leq L$. Plug this inequality for $\prod_{i=j}^{N} \widehat{a}_{t_i-1}$ back to (6.28), we continue with

$$s_t \leq e^{-(t-t_0)/H} \cdot s_{t_0} e^L + \left(\sup_{t_0 \leq k < t} w_k \right) \left(1 + \sum_{j=1}^{N} e^{L-(N-j)} \right)$$

$$\leq e^{-(t-t_0)/H} \cdot s_{t_0} e^L + \left(\sup_{t_0 \leq k < t} w_k \right) \left(1 + e^L \sum_{j=0}^{N-1} e^{-j} \right)$$

$$\leq e^{-(t-t_0)/H} \cdot s_{t_0} e^L + \left(\sup_{t_0 \leq k < t} w_k \right) \left(1 + e^L \frac{1}{e-1} \right) ,$$

where we used $z_{t_0} = s_{t_0}$ and that N is at least $(t - t_0)/H$. This is the required bound, which holds for any $t, t_0 \in \mathbb{N}$. $\square$

6.C Proof of Theorem 22

Theorem 23 (Stability, Scalar Subsystems). *Under Assumptions 8-12, Algorithm 5 guarantees the ISS of the closed loop of (6.2) with*

$$\max\{\|x(t)\|_\infty, \|u(t)\|_\infty\} \leq O\left(e^{(\bar{n})^{9/2} \bar{d}} \right) \left(e^{-(t-t_0)/H} x(t_0) + \sup_{t_0 \leq k \leq t} \|w(k)\|_\infty \right) ,$$

where $x(t_0)$ is the initial condition, local dimension $\bar{n} = \max\{\|C^{\bar{d}}\|_1, \|C^{\bar{d}}\|_\infty, \max_j |\mathcal{M}(j)|\}$ represents the total state dimension in the $\bar{d}$-neighborhood specified by the dynamics interaction (6.1) and the communication graph $\mathcal{G}^C$. Parameter $\bar{d}$ is the largest local delay each subsystem allows for delayed information, and H is the SLS closed-loop response finite impulse horizon.

Proof. We first characterize the closed loop dynamics of (6.2) under Algorithm 5. In particular, despite the fact that each subsystem uses differently delayed information to compute the local parameter, column solutions to the closed-loop responses, and control actions, the closed loop for the global system under such distributed policy can be simply characterized as

$$x(t) = \sum_{k=0}^{H-1} \Phi_t^x[k] \widehat{w}(t-k), \quad u(t) = \sum_{k=0}^{H-1} \Phi_t^u[k] \widehat{w}(t-k) \tag{6.29a}$$

$$\widehat{w}(t) = \sum_{k=1}^{H} \left(A_t \Phi_{t-1}^x[k-1] + B_t \Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right) \widehat{w}(t-k) + \widetilde{w}(t-1), \tag{6.29b}$$

by Lemma 30. Here $u(t)$, $\widehat{w}(t)$ are concatenated control action and estimated disturbance from (6.19). A_t, B_t are the global consistent parameter concatenated with the local consistent parameters $A^{ij}(\theta_t^i), B^{ij}(\theta_t^i)$. Vector $\widetilde{w}(t)$ are the admissible consistent disturbances corresponding to A_t, B_t with the property that $\|\widetilde{w}(t)\|_\infty \leq W$

for all time t. Operators $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$ are shorthand for global closed-loop operators when (6.19) is implemented, with

$$\Phi_t^x[k](i,j) := \phi_{t-d(j\to i)}^{j,x}[k](i), \quad \Phi_t^u[k](i,j) := \phi_{t-d(j\to i)}^{j,u}[k](i).$$

We follow similar procedure in the proof of Theorem 20 and bound $\|\widehat{w}(t)\|_\infty$ from (6.29b) by examining the following dynamical evolution,

$$\|\widehat{w}(t)\|_\infty \le \sum_{k=1}^{H} \left\|A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right\|_\infty \|\widehat{w}(t-k)\|_\infty + \|\widetilde{w}(t-1)\|_\infty.$$

$$(6.30)$$

By Lemma 31, as long as $\sum_{t=1}^{\infty}\sum_{k=1}^{H} \left\|A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right\|_\infty \le L$ for some positive constant L, then we can bound (6.30) with

$$\|\widehat{w}(t)\|_\infty \le e^{-(t-t_0)/H} \cdot e^L x(t_0) + \sup_{t_0 \le k < t} \|\widetilde{w}(t)\|_\infty \frac{\left(e^L + e - 1\right)}{e-1}.$$

Therefore, what's left is to show

$$\sum_{t=1}^{\infty}\sum_{k=1}^{H} \left\|A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right\|_\infty \le L,$$

which is proved in Proposition 32 where $L = O\left(\text{poly}\left(\bar{n}\right)\bar{d}\right)$. This concludes the proof. $\qquad\square$

Lemma 32 (Bounded error for closed loop operators). *Let* $\mathbf{\Phi_t^x}, \mathbf{\Phi_t^u}$ *denote the global closed loop operators concatenated from sub-controllers generated with Algorithm 5 where* $\Phi_t^x[k](i,j) := \phi_{t-d(j\to i)}^{j,x}[k](i)$ *and* $\Phi_t^u[k](i,j) := \phi_{t-d(j\to i)}^{j,u}[k](i)$. *Denote matrices* A_t, B_t *as the global consistent parameter concatenated with local consistent parameters* $A^{ij}(\theta_t^i), B^{ij}(\theta_t^i)$. *Then we have*

$$\sum_{t=1}^{\infty}\sum_{k=1}^{H} \left\|A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k]\right\|_\infty \qquad (6.31)$$

$$\le (\bar{d}+3)\bar{n}^3 diam(\mathcal{P}_0)\left(\kappa\bar{n}^{\frac{3}{2}}\Gamma H + \frac{C}{1-\rho}\right),$$

where $\bar{n} = \max\{\|C^{\bar{d}}\|_1, \|C^{\bar{d}}\|_\infty, \max_j |\mathcal{M}(j)|\}$, *and* $\bar{d}$ *is the largest local delay each subsystem considers for the algorithm, while H is SLS controller horizon. Here, Γ is a system-theoretical constant that does not depend on the global dynamics properties detailed in Theorem 26.*

Proof. To ease notation, we use a_t^i and b_t^i to denote the ith row of A_t and B_t respectively.

Our strategy is to bound each term in (6.31) for a fixed t and k. We will see that the summation of these terms over all k and t remain bounded. Each term in (6.31) can be bounded as follows:

$$\left\| A_t \Phi_{t-1}^x[k-1] + B_t \Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right\|_\infty$$

$$= \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left| \left(a_t^i\right)^T \Phi_{t-1}^x[k-1](:,j) + \left(b_t^i\right)^T \Phi_{t-1}^u[k-1](:,j) - \underbrace{\Phi_t^x[k](i,j)}_{\text{Defined to be } \phi_{t-d(j\to i)}^{j,x}[k](i)} \right| . \tag{6.32}$$

Due to the sparsity constraints that correspond to the information constraints placed on the closed-loop responses during synthesis (6.18), the only nonzero elements in a particular row i of $\Phi_t^x[k]$ are the positions at $j \in \mathcal{D}_{\text{in}}(i)$. Hence, we can write sum of each row i as sum of the elements in position (i,j) where $j \in \mathcal{D}_{\text{in}}(i)$ in (6.32). Recall that $\phi_{t-d(j\to i)}^{j,x}$ are synthesized in (6.18) such that

$$\phi_{t-d(j\to i)}^{j,x}[k](i) = \left(a_{t-d(j\to i)-d(i\to j)}^i\right)^T \phi_{t-d(j\to i)}^{j,x}[k-1]$$
$$+ \left(b_{t-d(j\to i)-d(i\to j)}^i\right)^T \phi_{t-d(j\to i)}^{j,u}[k-1] \tag{6.33}$$

because $\phi_{t-d(j\to i)}^{j,x}$ is synthesized by j at time $t - d\,(j \to i)$. The ith position of $\phi_{t-d(j\to i)}^{j,x}$ in particular uses model information from subsystem i, which is transmitted to j from i with delay $d\,(i \to j)$. Therefore, we substitute (6.33) into (6.32) to get

$$(6.32) = \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left| \left(a_t^i\right)^T \Phi_{t-1}^x[k-1](:,j) - \left(a_{t-d(j\to i)-d(i\to j)}^i\right)^T \phi_{t-d(j\to i)}^{j,x}[k-1] \right.$$

$$\left. + \left(b_t^i\right)^T \Phi_{t-1}^u[k-1](:,j) - \left(b_{t-d(j\to i)-d(i\to j)}^i\right)^T \phi_{t-d(j\to i)}^{j,u}[k-1] \right| . \tag{6.34}$$

Adding and subtracting $\left(a_t^i\right)^T \phi_{t-d(j\to i)}^{j,x}[k-1]$ and $\left(b_t^i\right)^T \phi_{t-d(j\to i)}^{j,u}[k-1]$ in (6.34), we can group terms and get

$$(6.34) \leq \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left| \left(a_t^i\right)^T \left(\Phi_{t-1}^x[k-1](:,j) - \phi_{t-d(j\to i)}^{j,x}[k-1] \right) \right.$$

$$\left. + \left(b_t^i\right)^T \left(\Phi_{t-1}^u[k-1](:,j) - \phi_{t-d(j\to i)}^{j,u}[k-1] \right) \right| \tag{6.35a}$$

$$+ \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left| \left(a_t^i - a_{t-d(j \to i)-d(i \to j)}^i \right)^T \phi_{t-d(j \to i)}^{j,x}[k-1] \right.$$

$$\left. + \left(b_t^i - b_{t-d(j \to i)-d(i \to j)}^i \right)^T \phi_{t-d(j \to i)}^{j,u}[k-1] \right|. \qquad (6.35\text{b})$$

We now consider (6.35a) and (6.35b) separately. For the remainder of the proof, we use $\phi_t^{j,x}$ and $\phi_t^{j,u}$ as shorthand for the jth column of $\mathbf{\Phi_t^x}$ and $\mathbf{\Phi_t^u}$ respectively. Apply Cauchy-Schwarz,

$$(6.35\text{a}) \le \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left\| a_t^i \right\|_2 \left\| \phi_{t-1}^{j,x}[k-1] - \phi_{t-d(j \to i)}^{j,x}[k-1] \right\|_2$$

$$+ \left\| b_t^i \right\|_2 \left\| \phi_{t-1}^{j,u}[k-1] - \phi_{t-d(j \to i)}^{j,u}[k-1] \right\|_2$$

$$(6.36\text{a})$$

$$(\text{by Assumption 9}) \quad \le \kappa \cdot \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left\| \phi_{t-1}^{j,x}[k-1] - \phi_{t-d(j \to i)}^{j,x}[k-1] \right\|_2$$

$$+ \left\| \phi_{t-1}^{j,u}[k-1] - \phi_{t-d(j \to i)}^{j,u}[k-1] \right\|_2 \qquad (6.36\text{b})$$

$$= \kappa \cdot \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left(\sum_{\ell \in \mathcal{D}_{\text{out}}(j)} \left| \phi_{t-1}^{j,x}[k-1](\ell) - \phi_{t-d(j \to i)}^{j,x}[k-1](\ell) \right|^2 \right)^{1/2}$$

$$+ \left(\sum_{\ell \in \mathcal{D}_{\text{out}}(j)} \left| \phi_{t-1}^{j,u}[k-1](\ell) - \phi_{t-d(j \to i)}^{j,u}[k-1](\ell) \right|^2 \right)^{1/2}$$

$$(6.36\text{c})$$

$$= \kappa \cdot \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left(\sum_{\ell \in \mathcal{D}_{\text{out}}(j)} \left| \phi_{t-1-d(j \to \ell)}^{j,x}[k-1](\ell) - \phi_{t-d(j \to i)}^{j,x}[k-1](\ell) \right|^2 \right)^{1/2}$$

$$+ \left(\sum_{\ell \in \mathcal{D}_{\text{out}}(j)} \left| \phi_{t-1-d(j \to \ell)}^{j,u}[k-1](\ell) - \phi_{t-d(j \to i)}^{j,u}[k-1](\ell) \right|^2 \right)^{1/2},$$

$$(6.36\text{d})$$

where to arrive at (6.36c) we used the fact that the nonzero elements in any column/sub-controller synthesized or assembled at subsystem j corresponds to the elements in $\mathcal{D}_{\text{out}}(j)$. The last equality comes from the definition of $\mathbf{\Phi_{t-1}^x}, \mathbf{\Phi_{t-1}^u}$. Continuing, we bound any sum using the largest summand multiplied by the number of summands:

$$(6.35\text{a}) \le (6.36\text{d})$$

$$\le \kappa \cdot \max_{i \in [N]} \sum_{j \in \mathcal{D}_{\text{in}}(i)} \left(\bar{n} \cdot \max_{\ell \in \mathcal{D}_{\text{out}}(j)} \left\| \phi_{t-1-d(j \to \ell)}^{j,x}[k-1] - \phi_{t-d(j \to i)}^{j,x}[k-1] \right\|_2^2 \right)^{1/2}$$

$$+ \left(\bar{n} \cdot \max_{\ell' \in \mathcal{D}_{\text{out}}(j)} \left\| \phi^{j,u}_{t-1-d(j\to\ell')}[k-1] - \phi^{j,u}_{t-d(j\to i)}[k-1] \right\|_2^2 \right)^{1/2}$$

$$(6.37\text{a})$$

$$= \kappa \bar{n}^{3/2} \max_{i\in[N]} \max_{j\in\mathcal{D}_{\text{in}}(i)} \left(\left(\max_{\ell\in\mathcal{D}_{\text{out}}(j)} \left\| \phi^{j,x}_{t-1-d(j\to\ell)}[k-1] - \phi^{j,x}_{t-d(j\to i)}[k-1] \right\|_2^2 \right)^{1/2} \right.$$

$$\left. + \left(\max_{\ell' \in \mathcal{D}_{\text{out}}(j)} \left\| \phi^{j,u}_{t-1-d(j\to\ell')}[k-1] - \phi^{j,u}_{t-d(j\to i)}[k-1] \right\|_2^2 \right)^{1/2} \right). \quad (6.37\text{b})$$

Recall that $\phi^j_{t-1-d(j\to\ell)}$ are generated by subsystem j using model information $\widehat{\Theta}^j_{t-1-d(j\to\ell)}$ during synthesis procedure ((6.17), Algorithm 5). Similarly, $\phi^j_{t-d(j\to i)}$ are generated using $\widehat{\Theta}^j_{t-d(j\to i)}$. Therefore, we can invoke Corollary 23.1 and arrive at

$$(6.35\text{a}) \leq (6.36\text{d}) \leq (6.37\text{b})$$

$$\leq \kappa \bar{n}^{3/2} \Gamma \max_{i\in[N]} \max_{j\in\mathcal{D}_{\text{in}}(i)} \left(\left(\max_{\ell\in\mathcal{D}_{\text{out}}(j)} \left\| \widehat{\Theta}^j_{t-1-d(j\to\ell)} - \widehat{\Theta}^j_{t-d(j\to i)} \right\|_F^2 \right)^{1/2} \right.$$

$$\left. + \max_{\ell' \in \mathcal{D}_{\text{out}}(j)} \left(\left\| \widehat{\Theta}^j_{t-1-d(j\to\ell')} - \widehat{\Theta}^j_{t-d(j\to i)} \right\|_F^2 \right)^{1/2} \right).$$

$$(6.38)$$

For any fixed i, j, ℓ, ℓ', the following holds true:

$$(6.38) = \kappa \bar{n}^{3/2} \Gamma \left(\left\| \widehat{\Theta}^j_{t-1-d(j\to\ell)} - \widehat{\Theta}^j_{t-d(j\to i)} \right\|_F + \left\| \widehat{\Theta}^j_{t-1-d(j\to\ell')} - \widehat{\Theta}^j_{t-d(j\to i)} \right\|_F \right)$$

$$= \kappa \bar{n}^{3/2} \Gamma \sum_{m\in\mathcal{M}(j)} \left\| \theta^m_{t-1-d(j\to\ell)-d(m\to j)} - \theta^m_{t-d(j\to i)-d(m\to j)} \right\|_F$$

$$+ \sum_{m\in\mathcal{M}(j)} \left\| \theta^m_{t-1-d(j\to\ell')-d(m\to j)} - \theta^m_{t-d(j\to i)-d(m\to j)} \right\|_F$$

$$\leq \kappa \bar{n}^{3/2} \Gamma \sum_{m\in\mathcal{M}(j)} \left(\sum_{p=\min(t_1,t_2)}^{\min(t_1,t_2)+\delta_t+1} \left\| \theta^m_{t-p+1} - \theta^m_{t-p} \right\|_F + \sum_{p=\min(t_1',t_2)}^{\min(t_1',t_2)+\delta_t'+1} \left\| \theta^m_{t-p+1} - \theta^m_{t-p} \right\|_F \right),$$

$$(6.39)$$

where we define $t_1 = 1 + d(j \to \ell) + d(m \to j)$, $t_1' = 1 + d(j \to \ell') + d(m \to j)$, $t_2 = 1 + d(j \to i) + d(m \to j)$, and $\delta_t = |d(j \to i) - d(j \to \ell) - 1|$, $\delta_t' = |d(j \to i) - d(j \to \ell') - 1|$. We stop at (6.39) for the moment for our bound for (6.35a) and change course to bound the other term (6.35b) in (6.35). We start with cauchy-schwarz for (6.35b).

$$(6.35\text{b}) \leq \max_{i\in[N]} \sum_{j\in\mathcal{D}_{\text{in}}(i)} \left\| a^i_t - a^i_{t-d(j\to i)-d(i\to j)} \right\|_2 \left\| \phi^{j,x}_{t-d(j\to i)}[k-1] \right\|_2$$

$$+ \left\| b_t^i - b_{t-d(j\to i)-d(i\to j)}^i \right\|_2 \left\| \phi_{t-d(j\to i)}^{j,u}[k-1] \right\|_2$$

$$\leq C\rho^{k-1}\bar{n} \cdot \max_{i\in[N]} \max_{j\in\mathcal{D}_{\text{in}}(i)} \left\| a_t^i - a_{t-d(j\to i)-d(i\to j)}^i \right\|_2 + \left\| b_t^i - b_{t-d(j\to i)-d(i\to j)}^i \right\|_2$$

$$= C\rho^{k-1}\bar{n} \cdot \max_{i\in[N]} \max_{j\in\mathcal{D}_{\text{in}}(i)} \left\| \theta_t^i - \theta_{t-d(j\to i)-d(i\to j)}^i \right\|_2 . \tag{6.40}$$

Here we have used the decay property of the finite-impulse-response closed-loop responses to bound the decay rate of the sub-controllers. The last equality holds by recalling that we have defined a_t^i and b_t^i to be the ith row of the A_t and B_t respectively, which is constructed from the global consistent parameter $\Theta_t = \cup_{i=1}^N \theta_t^i$. Therefore, by definition, $[a_t^i, b_t^i] = \theta_t^i$.

We now return to bound $\sum_{t=0}^{\infty} \sum_{k=1}^{H} \left\| A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right\|_{\infty}$. In particular, we have so far showed that

$$\sum_{t=0}^{\infty} \sum_{k=1}^{H} \left\| A_t\Phi_{t-1}^x[k-1] + B_t\Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right\|_{\infty} \leq \sum_{t=0}^{\infty} \sum_{k=1}^{H} (6.39) + (6.40).$$

$$\tag{6.41}$$

Therefore, our goal is to bound each component of the right hand side. Specifically,

$$\sum_{t=0}^{\infty} \sum_{k=1}^{H} (6.39)$$

$$\leq \sum_{t=0}^{\infty} \sum_{k=1}^{H} \kappa\bar{n}^{3/2}\Gamma \sum_{m\in\mathcal{M}(j)} \left(\sum_{p=\min(t_1,t_2)}^{\min(t_1,t_2)+\delta_t+1} \left\| \theta_{t-p+1}^m - \theta_{t-p}^m \right\|_F + \sum_{p=\min(t_1',t_2)}^{\min(t_1',t_2)+\delta_t'+1} \left\| \theta_{t-p+1}^m - \theta_{t-p}^m \right\|_F \right),$$

$$\tag{6.42}$$

for a different tuple of $(i \in [N], j \in \mathcal{D}_{\text{in}}(i), \ell \in \mathcal{D}_{\text{out}}(j), \ell' \in \mathcal{D}_{\text{out}}(j))$ at each t. However, for any (i, j, ℓ, ℓ'), the following holds:

$$\sum_{t=0}^{\infty} \sum_{k=1}^{H} (6.39)$$

$$\leq \kappa\bar{n}^{3/2}\Gamma \sum_{k=1}^{H} \sum_{m\in\mathcal{M}(j)} \left(\sum_{p=\min(t_1,t_2)}^{\min(t_1,t_2)+\delta_t+1} \sum_{t=0}^{\infty} \left\| \theta_{t-p+1}^m - \theta_{t-p}^m \right\|_F + \sum_{p=\min(t_1',t_2)}^{\min(t_1',t_2)+\delta_t'+1} \sum_{t=0}^{\infty} \left\| \theta_{t-p+1}^m - \theta_{t-p}^m \right\|_F \right)$$

$$\leq 2\kappa\bar{n}^{9/2}\Gamma H\text{diam}(\mathcal{P}_0) \left(\max_{i\in[N], j\in\mathcal{D}_{\text{in}}(i), \ell\in\mathcal{D}_{\text{out}}(j)} (1 + 1 + |d(j\to i)) - d(j\to\ell) - 1| \right)$$

$$\tag{6.43a}$$

$$\leq 2\kappa\bar{n}^{9/2}\Gamma H\text{diam}(\mathcal{P}_0)(\bar{d}+3). \tag{6.43b}$$

Here we have used in the competitiveness of each local Steiner point selector via (6.3) in (6.43a) with competitive ratio of $\bar{n}/2$. Furthermore, by definition of $\mathcal{D}_{\text{in}}(i)$ and $\mathcal{D}_{\text{out}}(j)$, we know that the largest delay for $d(j \to i)$ and $d(j \to \ell)$ for any choice of i, j, ℓ is less than $\bar{d}$.

Finally, we investigate the second component of the right hand side of (6.41).

$$\sum_{t=0}^{\infty} \sum_{k=1}^{H} (6.40) = \sum_{t=0}^{\infty} \sum_{k=1}^{H} C\rho^{k-1}\bar{n} \cdot \max_{i \in [N]} \max_{j \in \mathcal{D}_{\text{in}}(i)} \left\| \theta_t^i - \theta_{t-d(j\to i)-d(i\to j)}^i \right\|_2 \tag{6.44a}$$

$$\leq \sum_{k=1}^{H} C\rho^{k-1}\bar{n} \max_{i \in [N]} \max_{j \in \mathcal{D}_{\text{in}}(i)} \sum_{p=0}^{d(j\to i)+d(i\to j)+1} \sum_{t=0}^{\infty} \left\| \theta_{t-p+1}^i - \theta_{t-p}^i \right\|_2 \tag{6.44b}$$

$$\leq C\bar{n}^3 \text{diam}(\mathcal{P}_0)(\bar{d}+1)/(1-\rho), \tag{6.44c}$$

where we once again used the competitive ratio of the local Steiner point selector (6.3). Moreover, by definition of $\mathcal{D}_{\text{in}}(i)$, the largest delay $d(i \to j)$ for any $j \in \mathcal{D}_{\text{in}}(i)$ is less than $\bar{d}$.

Finally, we have the bound on the target quantity with (6.43b) and (6.44c) and conclude

$$\sum_{t=1}^{\infty} \sum_{k=1}^{H} \left\| A_t \Phi_{t-1}^x[k-1] + B_t \Phi_{t-1}^u[k-1] - \Phi_t^x[k] \right\|_\infty \leq (6.43b) + (6.44c)$$

$$\leq 2(\bar{d}+3)\bar{n}^3 \text{diam}(\mathcal{P}_0) \left(\kappa \bar{n}^{\frac{3}{2}} \Gamma H + \frac{C}{1-\rho} \right).$$

$\square$

Corollary 23.1 (of Theorem 26, Structured SLS sensitivity). *Consider the optimal solutions ϕ, ϕ' to (6.18) with two different parameters input Θ, Θ' respectively. Then we have*

$$\|\phi - \phi'\|_2 \leq \Gamma \|\Theta - \Theta'\|_2,$$

with $\Gamma = O(\Gamma_A + \Gamma_B)$ where Γ_A and Γ_B are constants in Theorem 26.

Proof. The SLS synthesis problem that we consider in (6.18) has one additional sparsity constraints than general SLS synthesis presented in (6.49) to which Theorem 26 apples. Therefore, we need to de-constrain the synthesis problem (6.18) and turn it into a problem of the form (6.49) in order to apply Theorem 26. To do so, we follow

the procedure in Section 2.4 of Chapter 2, where a re-parameterization of $\phi_t^{j,u}$ is used to characterize all sparse $\phi_t^{j,u}$ which will result in sparse $\phi_t^{j,x}$ according to the dynamical evolution (6.18b). First, we rewrite (6.18b) with the nonzero variables grouped together as follows:

$$\begin{bmatrix} \widetilde{\phi}^{j,x} \\ \widetilde{\phi}_b^{j,x} \end{bmatrix} [k+1] = \begin{bmatrix} A_{nn}^{(j)} & A_{nb}^{(j)} \\ A_{bn}^{(j)} & A_{bb}^{(j)} \end{bmatrix} \begin{bmatrix} \widetilde{\phi}^{j,x} \\ \widetilde{\phi}_b^{j,x} \end{bmatrix} [k] + \begin{bmatrix} B_n^{(j)} \\ B_b^{(j)} \end{bmatrix} \widetilde{\phi}^{j,u}[k], \qquad (6.45)$$

where $\widetilde{\phi}^{j,x}$ denotes the vector of nonzero entries in $\phi_t^{j,x}$ and $\widetilde{\phi}_b^{j,x}$ denotes the "boundary" positions of $\widetilde{\phi}^{j,x}$. The "boundary" positions of $\widetilde{\phi}^{j,x}$ corresponds to the positions in the vector that would become nonzero from zero due to the dynamical evolution (6.18b) in one time step. We refer Chapter 2 for detailed setup/derivation for (6.45). We also partition A,B in (6.18b) to correspond the entries that are associated with $\widetilde{\phi}^{j,x}$ and $\widetilde{\phi}_b^{j,x}$. $\widetilde{\phi}^{j,u}$ denote the reduced vector with only nonzero entries of $\phi_t^{i,u}$. For completeness, we re-state a key lemma from Chapter 2.

Lemma 33. *If $B_b^{(j)} B_b^{(j)\dagger} = I$, then the vectors $\{v^j[k]\}$ characterize all $\widetilde{\phi}^{j,u}[k]$ via*

$$\widetilde{\phi}^{j,u}[k] = -B_b^{(j),\dagger} A_{bn}^{(j)} \widetilde{\phi}^{j,x}[k] + \left(I - B_b^{(j),\dagger} B_b^{(j)} \right) v^j[k]. \qquad (6.46)$$

We remark that the pseudo-inverse condition in Lemma 33 is equivalently to Assumption 12, as observed in [236] and [254].

We can now substitute (6.46) into the synthesis problem (6.18) and obtain an SLS synthesis problem in the same form as (6.49) with transformed dynamical evolution in terms of the new variables $\widetilde{\phi}^{j,x}[k]$ and $v^j[k]$. Consider the optimal solutions $\widetilde{\phi}$ and $\widetilde{\phi}'$ (concatenated from $\widetilde{\phi}^{j,x}$ and v^j) computed from the de-constrained problem with two different model input Θ and Θ'. By Theorem 26, we have

$$\left\| \widetilde{\phi} - \widetilde{\phi}' \right\|_2 \le (\Gamma_A + \Gamma_B) \left\| \Theta - \Theta' \right\|_F. \qquad (6.47)$$

Observe that

$$\widetilde{\phi}^{j,u} = \begin{bmatrix} -B_b^{(j),\dagger} A_{bn}^{(j)} & \left(I - B_b^{(j),\dagger} B_b^{(j)} \right) \end{bmatrix} \widetilde{\phi}.$$

Therefore, we could bound the sensitivity of the solution to (6.18) via

$$\|\phi - \phi'\|_2 \le \left\| \begin{bmatrix} I & 0 \\ -B_b^{(j),\dagger} A_{bn}^{(j)} & \left(I - B_b^{(j),\dagger} B_b^{(j)} \right) \end{bmatrix} \widetilde{\phi} - \begin{bmatrix} I & 0 \\ -B_b^{'(j),\dagger} A_{bn}^{'(j)} & \left(I - B_b^{'(j),\dagger} B_b^{'(j)} \right) \end{bmatrix} \widetilde{\phi}' \right\|_2$$

$$\le \left\| \left(\begin{bmatrix} 0 & 0 \\ -B_b^{(j),\dagger} A_{bn}^{(j)} + B_b^{'(j),\dagger} A_{bn}^{'(j)} & -B_b^{(j),\dagger} B_b^{(j)} + B_b^{'(j),\dagger} B_b^{'(j)} \end{bmatrix} \right) \widetilde{\phi} \right\|_2$$

$$+ \left\| \begin{bmatrix} I & 0 \\ -B_b^{'(j),\dagger} A_{bn}^{'(j)} & \left(I - B_b^{'(j),\dagger} B_b^{'(j)} \right) \end{bmatrix} \left(\widetilde{\phi} - \widetilde{\phi}' \right) \right\|_2$$

$$\leq \frac{4C\kappa}{\sigma_{\min}(1 - \rho)} + \left(2 + \frac{2\kappa}{\sigma_{\min}} \right) (\Gamma_A + \Gamma_B) \left\| \Theta - \Theta' \right\|_F$$

$$= O\left(\Gamma_A + \Gamma_B \right) \left\| \Theta - \Theta' \right\|_F ,$$

where $\sigma_{\min}$ denotes the minimum singular value of the matrix B_b for all $B(\theta^i)$ with $\theta^i \in \mathcal{P}_0^i$. Note that the left pseudo-inverse has the largest singular value of $1/\sigma_{\min}$ with $\sigma_{\min}$ the smallest singular value of the original matrix. Due to Assumption 10 and Assumption 12, we know that B_b has to be bounded from below so that (6.18) is feasible. We have also used the fact that the norm of an lower triangular block matrix is upper bounded by the sum of the norm of each component block. We invoke the exponential decay property of the closed-loop responses to bound the decay rate of $\widetilde{\phi}$ by relating the nonzero component of the solution to (6.18) and $\widetilde{\phi}$ via (6.47). □

6.D Perturbation Analysis of H2-optimal SLS Synthesis

From $\mathcal{H}_2$-optimal control to Least Squares

This section presents results about general SLS synthesis. Due to notation overhead, we will drop time indices and suppress the horizon index $k \in [H]$ in closed-loop operators $\Phi^x[k]$, $\Phi^x[k]$ and write Φ_k^x, Φ_k^u instead. Let $\Phi_k^x \in \mathbb{R}^{n \times n}$ and $\Phi_k^u \in \mathbb{R}^{m \times n}$ and consider the following canonical SLS synthesis problem with LQR cost for system matrices $[A, B]$ and weighting matrices $C \in \mathbb{R}^{n \times n}, D \in \mathbb{R}^{m \times m}$:

$$S = \min \left\| \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} \Phi_1^x & \Phi_2^x & \dots & \Phi_T^x \\ \Phi_1^u & \Phi_2^u & \dots & \Phi_T^u \end{bmatrix} \right\|_F^2 \tag{6.48}$$

$$\text{s.t.: } \Phi_1^x = I$$

$$\Phi_{k+1}^x = A\Phi_k^x + B\Phi_k^u, \quad \forall k : 1 \leq k \leq H$$

$$\Phi_{H+1}^x = 0 .$$

The objective in (6.48) is equivalent to weighted $\mathcal{H}_2$ norm on the closed-loop operators $\mathbf{\Phi^x}$ and $\mathbf{\Phi^u}$, as well as the LQR cost on the state and control input weighed by C^2 and D^2. Denote $\phi_k^{j,x} \in \mathbb{R}^n$, $\phi_k^{j,u} \in \mathbb{R}^m$ as the jth column of $\Phi_k^x \in \mathbb{R}^{n \times n}$, $\Phi_k^u \in \mathbb{R}^{m \times n}$ and e_j the unit vector in the j-th coordinate axis. As described in Section 2.3, we can separate the problem by columns and can equivalently restate (6.48) in terms of each column $\phi_k^{j,x}$ and $\phi_k^{j,u}$:

$$S_j := \min \left\| \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} \phi_1^{j,x} & \phi_2^{j,x} & \dots & \phi_H^{j,x} \\ \phi_1^{j,u} & \phi_2^{j,u} & \dots & \phi_H^{j,u} \end{bmatrix} \right\|_F^2 \tag{6.49}$$

$$\text{s.t.} \quad \phi_1^{j,x} = e_j$$
$$\phi_{k+1}^{j,x} = A\phi_k^{j,x} + B\phi_k^{j,u}, \quad \forall\, 1 \le k \le H$$
$$\phi_{H+1}^{j,x} = 0 \,.$$

We will now fix j and rewrite (6.49) further and introduce new variables to avoid tedious notation. Define $u_k = \phi_k^{j,u}, \forall 1 \le k \le H$, $\boldsymbol{u} = [u_1^\top, \ldots, u_H^\top]^\top$ and the block-lower-triangular matrix $\boldsymbol{G}_u \in \mathbb{R}^{Hn \times Hm}$, the vector $\xi_j \in \mathbb{R}^{Hn}$ and the lifted weight matrices $\boldsymbol{C}, \boldsymbol{D}$ as

$$\boldsymbol{G}_u = \begin{bmatrix} B & 0 & 0 & \ldots & 0 \\ AB & B & 0 & \ldots & 0 \\ A^2 B & AB & B & \ldots & 0 \\ & \ldots & \ldots & & \\ A^{H-1}B & A^{H-2}B & A^{H-3}B & \ldots & B \end{bmatrix} \quad \xi_j = \begin{bmatrix} -Ae_j \\ -A^2 e_j \\ \ldots \\ -A^H e_j \end{bmatrix} \quad \boldsymbol{C} = I_H \otimes C \quad \boldsymbol{D} = I_H \otimes D,$$

$$(6.50)$$

where I_k is the identity matrix for $\mathbb{R}^k$. Denote by P_i, $1 \le i \le H$ the i-th block-row of $\boldsymbol{G}_u$:

$$P_i = [A^{i-1}B, A^{i-2}B, \ldots, B, 0, \ldots, 0] \,. \tag{6.51}$$

Observe that with these definitions, it holds that for any feasible $\phi_k^{j,u}, \phi_k^{j,x}$ and for all $\forall 1 \le k \le H$:

$$\phi_{k+1}^{j,x} = -\xi_{j,k} + P_k \boldsymbol{u}$$

due to the constraints in (6.49). Now we can rewrite the subproblem S_j as

$$S_j = \min_{\boldsymbol{u}} \quad \left\| \begin{bmatrix} \boldsymbol{C}\boldsymbol{G}_u \\ \boldsymbol{D} \end{bmatrix} \boldsymbol{u} - \begin{bmatrix} \boldsymbol{C}\xi_j \\ 0 \end{bmatrix} \right\|_2^2 + (C^\top C)_{jj} \tag{6.52a}$$

$$\text{s.t.:} \quad 0 = A^\top e_j + P_H \boldsymbol{u} \,. \tag{6.52b}$$

For large systems which consist of many interconnected (sparsely) small systems, it is often the case that the overall system is H-controllable for some suitable choice of $H \ll n$ where n is the global state dimension.

Representation as a Least-Squares problem

We now rewrite (6.52) as a least square problem. Define $\boldsymbol{u}_c^* := P_H^\top (P_H P_H^\top)^{-1} A^\top e_j$, which is the solution to the optimization problem

$$\min_{\boldsymbol{u}} \quad \|\boldsymbol{u}\|_2^2$$

$$\text{s.t. } 0 = -A^\top e_j + P_H u.$$

We can interpret u_c^* as the smallest control action, measured in ℓ_2, that drives the system from the origin to $-A^\top e_j$ in H time-steps. This relates to controllability grammians as described in [5]. Using $M^\dagger$ to denote the Moore-Penrose Inverse of a matrix M, we can also write $u_c^* := P_H^+ A^\top e_j = P_H^\top W_H^{-1} A^\top e_j$, where $W_H = P_H P_H^\top$.

Let H denote the FIR-Horizon of the problem, then define the matrices

$$G_w(A) = \begin{bmatrix} I & 0 & 0 & \dots & 0 \\ A & I & 0 & \dots & 0 \\ A^2 & A & I & \dots & 0 \\ & \dots & \dots & & \\ A^{H-1} & A^{H-2} & A^{H-3} & \dots & I \end{bmatrix}, \quad G_u(A,B) = \begin{bmatrix} B & 0 & 0 & \dots & 0 \\ AB & B & 0 & \dots & 0 \\ A^2 B & AB & B & \dots & 0 \\ & \dots & \dots & & \\ A^{H-1}B & A^{H-2}B & A^{H-3}B & \dots & B \end{bmatrix} \tag{6.53}$$

and denote $P_i(A, B)$ as the ith block matrix row of $G_u(A, B)$:

$$P_i(A, B) = [A^{i-1}B, A^{i-2}B, \dots, B, 0, \dots, 0]. \tag{6.54}$$

$G_u(A, B)$ can be written as $G_u(A, B) = G_w(A)(I_H \otimes B)$, where I_H is the identity matrix in $\mathbb{R}^H$. Let $Z \in \mathbb{R}^{H \times H}$ be defined as the nilpotent matrix

$$Z = \begin{bmatrix} 0_{H-1 \times 1} & I_{H-1} \\ 0 & 0_{1 \times H-1} \end{bmatrix}, \tag{6.55}$$

and notice its psuedo-inverse is $Z^\dagger = Z^\top$. Using Z, it is easy to verify that $G_w(A)$ can be expressed as

$$G_w(A) = \left(I_H - Z^\dagger \otimes A \right)^{-1}. \tag{6.56}$$

Ignoring the constant terms in (6.52a), we can reparametrize $u = -u_c^* + u'$ where $u' \in \text{null}(P_H)$ and describe (6.52) as the optimization problem:

$$S_j := \min_{u' \in \text{null}(P_H(A,B))} \left\| \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A,B) \\ I \end{bmatrix} (u' - u_c^*(A,B)) \right\|_2^2. \tag{6.57}$$

Let $u^*(A, B)$ be a minimizer of the above problem for fixed A, B, we are interested in the SLS solutions

$$\phi^{*j}(A, B) := \begin{bmatrix} \phi_x^{*j}(A, B) \\ \phi_u^{*j}(A, B) \end{bmatrix} = \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix} (u^*(A, B) - u_c^*(A, B))$$

and how these solutions are perturbed with changes in A, B.

For the rest of the discussion, we will drop mentioning the explicit dependence on (A, B) and the column index j to reduce the notational burden. First, we (over-)parametrize u as

$$u = (I - P_H^\dagger P_H)\eta,$$

to cast the above problem into an unconstrained one:

$$S_j := \min_{\eta} \left\| \underbrace{\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix} (I - P_H^\dagger P_H)}_{\mathcal{F}} \eta - \underbrace{\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix} u_c^*(A, B)}_{g} \right\|_2^2.$$

(6.58)

The unique min-norm solution η^* to the above problem is $\eta^* = \mathcal{F}^\dagger g$ and therefore the optimal solution ϕ^* takes the form

$$\phi^* = \begin{bmatrix} C^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} (\mathcal{F}\mathcal{F}^\dagger g - g) = \begin{bmatrix} C^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} \underbrace{(\mathcal{F}\mathcal{F}^\dagger - I)g}_{v^*} =: \begin{bmatrix} C^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} v^*$$

(6.59)

Local lipshitzness of $\mathcal{H}_2$-optimal closed-loop operators

Here, we perform perturbation analysis on the term $v^* = (\mathcal{F}\mathcal{F}^\dagger - I)g$. Throughout the discussion, we will make frequent use of the following identities:

Lemma 34. *For arbitrary matrices $X, Y \in \mathbb{R}^{n \times m}$ and $A, B \in \mathbb{R}^{n \times n}$, it holds that*

1. $A_1^k - A_2^k = \sum_{j=0}^{k-1} A_1^{k-1-j}(A_1 - A_2)A_2^j$

2. $XX^\dagger - YY^\dagger = (I - XX^\dagger)(X - Y)Y^\dagger + \left[(I - YY^\dagger)(X - Y)X^\dagger\right]^\top$

3. *If A and B are invertible, then $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$.*

The following is a corollary from Theorem 4.1 in [255]:

Theorem 24. *Let X and Y be matrices with equal rank, let $\| \cdot \|_2$ denote the induced 2-norm and $\| \cdot \|_F$ denote the Frobenius norm. The following inequalities hold:*

$$\|X^\dagger - Y^\dagger\|_2 \le \varphi \|X^\dagger\|_2 \|Y^\dagger\|_2 \|X - Y\|_2$$
$$\|X^\dagger - Y^\dagger\|_F \le \sqrt{2} \|X^\dagger\|_2 \|Y^\dagger\|_2 \|X - Y\|_F$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ denotes the golden ratio constant.

Next we present the core theorem of the perturbation analysis: Given two *arbitrary* controllable systems (A_1, B_1) and (A_2, B_2), Theorem 25 bounds the worst-case difference in solutions $\|\phi_1^* - \phi_2^*\|_2$ in terms of the differences in parameters space $\|A_1 - A_2\|_2$ and $\|B_1 - B_2\|_2$ between both systems. This result is the first perturbation bound for $\mathcal{H}_2$-optimal control with SLS (considering arbitrary pairs of A_1, A_2 and B_1, B_2) .

Theorem 25. *Let $C, D > 0$, let (A_1, B_1) and (A_2, B_2) be two controllable pairs of system matrices with FIR horizon H and let ϕ_1^{*j} and ϕ_2^{*j} be the corresponding SLS-solutions to the subproblem S_j. Then, it holds that*

$$\|\phi_1^{j*} - \phi_2^{j*}\|_2 \leq \Gamma_A \|A_1 - A_2\|_F + \Gamma_B \|B_1 - B_2\|_F \tag{6.60}$$

where the Lipshitz-constants Γ_A, Γ_B stand for

$$\Gamma_A = \kappa_{CD}\Gamma_1' + \kappa_{CD}\Gamma_2'\|B_1\|_2\|G_w(A_1)\|_2, \quad \kappa_{CD} = \frac{\max\{\sigma_{max}(C), \sigma_{max}(D)\}}{\min\{\sigma_{min}(C), \sigma_{min}(D)\}}$$

$$\Gamma_B = \kappa_{CD}\Gamma_2'\|G_w(A_2)\|_2$$

and Γ_1' and Γ_2' are defined as

$$\Gamma_1' = \alpha_{H,1}\alpha_{H,2}H(1 + \|G_{u,2}\|_2)\|P_{H,2}^\dagger\|_2$$

$$\Gamma_2' = \alpha_{H,1}\|P_{H,1}^\dagger\|_2 \left(1 + \varphi\|P_{H,2}^\dagger\|_2 + \varphi\|P_{H,2}^\dagger\|_2 \|G_{u,2}\|_2\right) + \|g_2\|_2(\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2)$$

$$+ \varphi\|g_2\|_2(\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2)\|P_{H,1}^\dagger\|_2\|P_{H,2}^\dagger\|_2(\|P_{H,1}\|_2 + \|P_{H,2}\|_2)(1 + \|G_{u,1}\|_2)$$

and $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

Proof. Recall the identities of Lemma 34. Write $v_1^* - v_2^*$ where v_i^* is from (6.59) for (A_i, B_i) as

$$v_1^* - v_2^* = (\mathcal{F}_1\mathcal{F}_1^\dagger - I)(g_1 - g_2) + (\mathcal{F}_1\mathcal{F}_1^\dagger - \mathcal{F}_2\mathcal{F}_2^\dagger)g_2$$

$$\|v_1^* - v_2^*\|_2 \leq \|g_1 - g_2\|_2 + \|\mathcal{F}_1\mathcal{F}_1^\dagger - \mathcal{F}_2\mathcal{F}_2^\dagger\|_2\|g_2\|_2, \tag{6.61}$$

where we used the fact that $(\mathcal{F}_1\mathcal{F}_1^\dagger - I)$ is a projection and therefore $\|\mathcal{F}_1\mathcal{F}_1^\dagger - I\|_2 = 1$. Rewrite $\mathcal{F}_1\mathcal{F}_1^\dagger - \mathcal{F}_2\mathcal{F}_2^\dagger$ as

$$(I - \mathcal{F}_1\mathcal{F}_1^\dagger)(\mathcal{F}_1 - \mathcal{F}_2)\mathcal{F}_2^\dagger + \left[(I - \mathcal{F}_2\mathcal{F}_2^\dagger)(\mathcal{F}_1 - \mathcal{F}_2)\mathcal{F}_1^\dagger\right]^\top$$

to conclude that

$$\|\mathcal{F}_1\mathcal{F}_1^\dagger - \mathcal{F}_2\mathcal{F}_2^\dagger\|_2 \leq \|\mathcal{F}_1 - \mathcal{F}_2\|_2(\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2). \tag{6.62}$$

Substitution into (6.61) yields

$$\|v_1^* - v_2^*\|_2 \leq \|g_1 - g_2\|_2 + \|\mathcal{F}_1 - \mathcal{F}_2\|_2(\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2)\|g_2\|_2, \tag{6.63}$$

1. <u>Bounding $\|\mathcal{F}_1 - \mathcal{F}_2\|_2$</u>: Rewrite $\mathcal{F}_1 - \mathcal{F}_2$ as

$$
\begin{bmatrix} C^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} (\mathcal{F}_1 - \mathcal{F}_2)
$$

$$
= \begin{bmatrix} G_{u,1} \\ I \end{bmatrix} (I - P_{H,1}^{\dagger} P_{H,1}) - \begin{bmatrix} G_{u,2} \\ I \end{bmatrix} (I - P_{H,2}^{\dagger} P_{H,2}) \tag{6.64}
$$

$$
= \begin{bmatrix} G_{u,1} \\ I \end{bmatrix} (P_{H,2}^{\dagger} P_{H,2} - P_{H,1}^{\dagger} P_{H,1}) + \begin{bmatrix} G_{u,1} - G_{u,2} \\ 0 \end{bmatrix} (I - P_{H,2}^{\dagger} P_{H,2}). \tag{6.65}
$$

From the above we can derive the inequality:

$$
\frac{\|\mathcal{F}_1 - \mathcal{F}_2\|_2}{\max\{\|C\|_2, \|D\|_2\}}
$$
$$
\leq (1 + \|G_{u,1}\|_2)\|P_{H,2}^{\dagger} - P_{H,1}^{\dagger}\|_2(\|P_{H,1}\|_2 + \|P_{H,2}\|_2) + \|G_{u,1} - G_{u,2}\|_2.
$$
$$
\tag{6.66}
$$

Now we will use the result Theorem 24 to bound $\|P_{H,2}^{\dagger} - P_{H,1}^{\dagger}\|_2$ as

$$
\|P_{H,2}^{\dagger} - P_{H,1}^{\dagger}\|_2 \leq \varphi \|P_{H,1}^{\dagger}\|_2 \|P_{H,2}^{\dagger}\|_2 \|P_{H,2} - P_{H,1}\|_2. \tag{6.67}
$$

Furthermore, noticing $P_{H,2} - P_{H,1} = [0, \ldots, 0, I_n](G_{u,2} - G_{u,1})$ we can conclude

$$
\|P_{H,2}^{\dagger} - P_{H,1}^{\dagger}\|_2 \leq \varphi \|P_{H,1}^{\dagger}\|_2 \|P_{H,2}^{\dagger}\|_2 \|G_{u,2} - G_{u,1}\|_2. \tag{6.68}
$$

We combine this into (6.66) to obtain

$$
\frac{\|\mathcal{F}_1 - \mathcal{F}_2\|_2}{\max\{\|C\|_2, \|D\|_2\}}
$$
$$
\leq \left(1 + \varphi \|P_{H,1}^{\dagger}\|_2 \|P_{H,2}^{\dagger}\|_2 (1 + \|G_{u,1}\|_2)(\|P_{H,1}\|_2 + \|P_{H,2}\|_2)\right) \|G_{u,1} - G_{u,2}\|_2.
$$
$$
\tag{6.69}
$$

2. <u>Bounding $\|g_1 - g_2\|_2$</u>: Introduce the constant $\alpha_H := \max_{0 \leq k \leq H} \|A^k\|_2$ and observe that $\|A_1^H - A_2^H\|_2$ can be bounded as

$$
\|A_1^H - A_2^H\|_2 = \|\sum_{j=0}^{H-1} A_1^{H-1-j}(A_1 - A_2)A_2^j\| \leq H\alpha_{H,1}\alpha_{H,2}\|A_1 - A_2\|_2.
$$
$$
\tag{6.70}
$$

We can rewrite $g_1 - g_2$ as

$$
\begin{bmatrix} C^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} (g_1 - g_2) = \begin{bmatrix} G_{u,1} \\ I \end{bmatrix} P_{H,1}^{\dagger} A_1^H e_j - \begin{bmatrix} G_{u,2} \\ I \end{bmatrix} P_{H,2}^{\dagger} A_2^H e_j \tag{6.71}
$$

$$= \begin{bmatrix} (G_{u,1} - G_{u,2}) \\ 0 \end{bmatrix} P_{H,1}^\dagger A_1^H e_j + \begin{bmatrix} G_{u,2} \\ I \end{bmatrix} (P_{H,1}^\dagger - P_{H,2}^\dagger) A_1^H e_j \tag{6.72}$$

$$+ \begin{bmatrix} G_{u,2} \\ I \end{bmatrix} P_{H,2}^\dagger (A_1^H - A_2^H) e_j$$

and obtain the bound

$$\frac{\|g_1 - g_2\|_2}{\max\{\|C\|_2, \|D\|_2\}} \leq \alpha_{H,1} \|G_{u,1} - G_{u,2}\|_2 \|P_{H,1}^\dagger\|_2 + \alpha_{H,1}(1 + \|G_{u,2}\|_2) \|P_{H,1}^\dagger - P_{H,2}^\dagger\|_2 \tag{6.73}$$

$$+ \alpha_{H,1}\alpha_{H,2} H(1 + \|G_{u,2}\|_2) \|P_{H,2}^\dagger\|_2 \|A_1 - A_2\|_2 \tag{6.74}$$

$$\leq \alpha_{H,1} \|P_{H,1}^\dagger\|_2 \left(1 + \varphi \|P_{H,2}^\dagger\|_2 + \varphi \|P_{H,2}^\dagger\|_2 \|G_{u,2}\|_2\right) \|G_{u,1} - G_{u,2}\|_2 \tag{6.75}$$

$$+ \alpha_{H,1}\alpha_{H,2} H(1 + \|G_{u,2}\|_2) \|P_{H,2}^\dagger\|_2 \|A_1 - A_2\|_2 . \tag{6.76}$$

We get the bound

$$\frac{\|\nu_1^* - \nu_2^*\|_2}{\max\{\|C\|_2, \|D\|_2\}} \leq \Gamma_1' \|A_1 - A_2\|_2 + \Gamma_2' \|G_{u,1} - G_{u,2}\|_2 \tag{6.77}$$

where Γ_1' and Γ_2' are the constants

$$\Gamma_1' = \alpha_{H,1}\alpha_{H,2} H(1 + \|G_{u,2}\|_2) \|P_{H,2}^\dagger\|_2 \tag{6.78}$$

$$\Gamma_2' = \alpha_{H,1} \|P_{H,1}^\dagger\|_2 \left(1 + \varphi \|P_{H,2}^\dagger\|_2 + \varphi \|P_{H,2}^\dagger\|_2 \|G_{u,2}\|_2\right) + \|g_2\|_2 (\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2) + \dots \tag{6.79}$$

$$+ \varphi \|g_2\|_2 (\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2) \|P_{H,1}^\dagger\|_2 \|P_{H,2}^\dagger\|_2 (\|P_{H,1}\|_2 + \|P_{H,2}\|_2)(1 + \|G_{u,1}\|_2)$$

Using Lemma 35, we obtain the final bound:

$$\|\phi_1^* - \phi_2^*\|_2 \leq \kappa_{CD} \|\nu_1^* - \nu_2^*\|_2 \leq \Gamma_A \|A_1 - A_2\|_2 + \Gamma_B \|B_1 - B_2\|_2 \tag{6.80}$$

with the constants Γ_A, Γ_B defined as:

$$\Gamma_A = \kappa_{CD}\Gamma_1' + \kappa_{CD}\Gamma_2' \|B_1\|_2 \|G_w(A_1)\|_2 \|G_w(A_2)\|_2 \tag{6.81}$$

$$\Gamma_B = \kappa_{CD}\Gamma_2' \|G_w(A_2)\|_2 \tag{6.82}$$

$\square$

Global lipshitzness of $\mathcal{H}_2$-optimal closed-loop operators over compact sets $\mathcal{S}$

This section derives a global Lipshitz bound for $\mathcal{H}_2$-optimal SLS solutions over a compact set of controllable systems $\mathcal{S}$. As a starting point we consider the previous theorem Theorem 25. Our main proof strategy is to derive global bounds on the constants Γ_A and Γ_B instead of for a fixed pair of systems. We proceed with a collection lemmas bounding individual terms in the equations (6.80) and (6.81) for $\mathcal{S}$.

Auxiliary Lemmas

Lemma 35. *For any pair of system matrices (A_1, B_1) and (A_2, B_2) (with compatible dimensions) holds*

$$\|\boldsymbol{G}_w(A_1) - \boldsymbol{G}_w(A_2)\|_2 \leq \|\boldsymbol{G}_w(A_1)\|_2 \|\boldsymbol{G}_w(A_2)\|_2 \|A_1 - A_2\|_2 \tag{6.83}$$

$$\|\boldsymbol{G}_u(A_1, B_1) - \boldsymbol{G}_u(A_2, B_2)\|_2 \leq \|B_1\|_2 \|\boldsymbol{G}_w(A_1)\|_2 \|\boldsymbol{G}_w(A_2)\|_2 \|A_1 - A_2\|_2 + \|\boldsymbol{G}_w(A_2)\|_2 \|B_1 - B_2\|_2$$

Proof. Using Lemma 34 we can write Using $\boldsymbol{G}_u(A, B) = \boldsymbol{G}_w(A)(I_H \otimes B)$ and Lemma 34 we can write $\boldsymbol{G}_{u,1} - \boldsymbol{G}_{u,2}$ as

$$\boldsymbol{G}_{u,1} - \boldsymbol{G}_{u,2} = \boldsymbol{G}_w(A_1)(I_H \otimes B_1) - \boldsymbol{G}_w(A_2)(I_H \otimes B_2) \tag{6.84}$$

$$= (\boldsymbol{G}_w(A_1) - \boldsymbol{G}_w(A_2))(I_H \otimes B_1) + \boldsymbol{G}_w(A_2)(I_H \otimes (B_1 - B_2)) \tag{6.85}$$

It holds that

$$\boldsymbol{G}_w(A_1) - \boldsymbol{G}_w(A_2) = \boldsymbol{G}_w(A_1)(\boldsymbol{G}_w(A_2)^{-1} - \boldsymbol{G}_w(A_1)^{-1})\boldsymbol{G}_w(A_2) \tag{6.86}$$

$$= \boldsymbol{G}_w(A_1)(Z^\dagger \otimes (A_1 - A_2))\boldsymbol{G}_w(A_2) \tag{6.87}$$

which leads to the bound

$$\|\boldsymbol{G}_w(A_1) - \boldsymbol{G}_w(A_2)\|_2 \leq \|\boldsymbol{G}_w(A_1)\|_2 \|A_1 - A_2\|_2 \|\boldsymbol{G}_w(A_2)\|_2. \tag{6.88}$$

$\square$

In total, we need to global bounds on the quantities $\|\boldsymbol{G}_u\|_2, \|\boldsymbol{G}_w\|_2, \|P_H^\dagger\|_2, \|P_H\|_2, \|\mathcal{F}^\dagger\|_2, \|\boldsymbol{g}\|_2$.

Lemma 36. *Let (A, B) be pair of fixed system matrices, let $G_u(A, B)$, $G_w(A)$ be the matrices defined in (6.53), and let $W_H^u = \sum_{i=0}^{H-1} A^i BB^\top A^{i\top}$, $W_H^w = \sum_{i=0}^{H-1} A^i A^{i\top}$ be the Hth controllability grammian w.r.t to the input u and the distrubance w, respectively. Then it holds:*

$$\|G_u(A, B)\|_2 \leq \sqrt{H\sigma_{max}(W_H^u(A, B))} \quad \|G_w(A)\|_2 \leq \sqrt{H\sigma_{max}(W_H^w(A))} \quad (6.89)$$

Proof. $\|G_u\|_2$ is defined as $\|G_u\|_2^2 := \max_{\|u\|_2=1} \|G_u u\|_2^2$, by decomposing $u = [u_0^\top, \ldots, u_{H-1}^\top]^\top$ we can rewrite this as

$$\|G_u\|_2^2 = \max_{\|u\|_2=1} \left\| \begin{bmatrix} Bu_0 \\ ABu_0 + Bu_1 \\ \ldots \\ A^{H-1}Bu_0 + + Bu_{H-1} \end{bmatrix} \right\|_2^2 = \max_{\|u\|_2=1} \sum_{k=1}^{H} \|P_k u\|_2^2 \quad (6.90)$$

$$\leq \sum_{k=1}^{H} \max_{\|u\|_2=1} \|P_k u\|_2^2 = \sum_{k=1}^{H} \|P_k\|_2^2 \leq H\|P_H\|_2^2 \leq H\|W_H^u\|_2. \quad (6.91)$$

where we used the fact that $\|P_k\|_2^2$ increases in k and that $\|P_k\|_2^2$ is equal to the induced 2-norm of the corresponding controllabillity grammian $W_k^u = \sum_{i=0}^{k-1} A^i BB^\top A^{i\top}$. Thus, we obtain the bound

$$\|G_u(A, B)\|_2 \leq \sqrt{H\sigma_{max}(W_H^u(A, B))},$$

and the bound on $\|G_w(A)\|_2$ follows in the same way. □

Lemma 37. *Let (A, B) be pair of H-controllable fixed system matrices, let $P_H(A, B)$ be the matrix defined in (6.54), and let $W_H^u = \sum_{i=0}^{H-1} A^i BB^\top A^{i\top}$ be the Hth controllability grammian w.r.t to the input u. Then, the induced 2 norm of $P_H(A, B)$ and its Moore-Penrose Inverse $P_H^\dagger(A, B)$ can be written as:*

$$\|P_H(A, B)\|_2 = \left(\sigma_{max}(W_H^u(A, B))\right)^{\frac{1}{2}} \quad \|P_H^\dagger(A, B)\|_2 = \left(\sigma_{min}(W_H^u(A, B))\right)^{-\frac{1}{2}} \quad (6.92)$$

Proof. Because we assume a sufficient degree of controllability, $P_H(A, B)$ is full row-rank. This implies that

$$\|P_H(A, B)\|_2 = \sqrt{\lambda_{max}(P_H(A, B)P_H^\top(A, B))} = \sqrt{\sigma_{max}(W_H^u(A, B))} \quad (6.93)$$

$$\left(\|P_H^\dagger(A, B)\|_2\right)^{-1} = \sqrt{\lambda_{min}(P_H(A, B)P_H^\top(A, B))} = \sqrt{\sigma_{min}(W_H^u(A, B))} \quad (6.94)$$

□

Lemma 38. *Let (A, B) be a fixed pair of H-controllable system matrices, and let $\mathcal{F}(A, B)$ denote the matrix*

$$\mathcal{F}(A, B) = \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix} (I - P_H^\dagger(A, B) P_H(A, B)). \qquad (6.95)$$

Then, $\|\mathcal{F}^\dagger(A, B)\|_2 \le \sigma_{min}^{-1}(D)$.

Proof. For an arbitrary matrix M, $(\|M^\dagger\|_2)^{-1}$ is equal to the smallest <u>non-zero</u> singular eigenvalue of M (we will denote this quantity as $\sigma_{-1}(M)$). Thus, in order to bound $\|M^\dagger\|_2$ from above, we have to bound $\sigma_{-1}(M)$ from below. Denote L as the matrix

$$L := \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix}$$

and notice that it is full column rank and has rank of $H \times n_u$. The projection $\Pi_{\mathcal{N}(P_H)} := (I - P_H^\dagger(A, B) P_H(A, B))$ has rank $H \times n_u - n_x$ due to the assumption of H-controllability. Hence, $F = L\Pi_{\mathcal{N}(P_H)}$ is full column rank with rank $r_{\mathcal{F}} := H \times n_u - n_x$ and has a null space $\mathcal{N}(\mathcal{F})$ of dimension n_x. From these observations, we can equivalently say that $\sigma_{-1}(\mathcal{F})$ is the $r_{\mathcal{F}}$th largest (or equivalently $n_x + 1$ smallest) singular eigenvalue of $\mathcal{F}$. Using the Minimax principle, we can therefore write

$$\sigma_{-1}(\mathcal{F}) = \max_{\text{proj.}\Pi, \text{ s.t.: } \text{rank}(\Pi)=r_{\mathcal{F}}} \min_{x \text{ s.t.: } \|\Pi x\|=1} x^\top \Pi \mathcal{F}^\top \mathcal{F} \Pi x \qquad (6.96)$$

$$= \max_{\text{proj.}\Pi, \text{ s.t.: } \text{rank}(\Pi)=r_{\mathcal{F}}} \min_{x \text{ s.t.: } \|\Pi x\|=1} x^\top \Pi \Pi_{\mathcal{N}(P_H)} L^\top L \Pi_{\mathcal{N}(P_H)} \Pi x . \qquad (6.97)$$

Now recall that $\Pi_{\mathcal{N}(P_H)}$ is of rank $r_{\mathcal{F}}$, hence it is a feasible choice for the variable Π of the outer optimization problem. This leads to the bound

$$\sigma_{-1}(\mathcal{F}) \ge \min_{x \text{ s.t.: } \|\Pi_{\mathcal{N}(P_H)}x\|=1} x^\top \Pi_{\mathcal{N}(P_H)} L^\top L \Pi_{\mathcal{N}(P_H)} x \qquad (6.98)$$

$$\ge \min_{z \text{ s.t.: } \|z\|=1} z^\top L^\top L z = \sigma_{min}(L) . \qquad (6.99)$$

We obtain a simple, but possibly conservative, lower bound on $\sigma_{min}(L)$ as follows:

$$\sigma_{min}^2(L) = \min_{z \text{ s.t.: } \|z\|=1} \|Lz\|_2^2 = \min_{z \text{ s.t.: } \|z\|=1} \|CG_u(A, B)z\|_2^2 + \|Dz\|_2^2 \ge \sigma_{min}^2(CG_u(A, B)) + \sigma_{min}^2(D)$$

$$\implies \quad \sigma_{min}(L) \ge \sigma_{min}(D) .$$

Finally, this provides us with the final result: $\|\mathcal{F}^\dagger(A, B)\|_2 = \sigma_{-1}^{-1}(F) \le \sigma_{min}^{-1}(L) \le \sigma_{min}^{-1}(D) .$ $\qquad \square$

We obtain an upper bound for $\|g\|_2$, as a corollary of the previous three Lemmas:

Lemma 39. *Let* (A, B) *be a fixed pair of H-controllable system matrices. Let* $g = Lu_c^*$, *where* L *and* u_c^* *are defined as:*

$$L := \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} G_u(A, B) \\ I \end{bmatrix} \qquad u_c^* := P_H^+ A^H e_j = P_H^\top W_H^{-1} A^H e_j. \qquad (6.100)$$

Then, it holds:

$$\|g\|_2 \leq \left(\|C\|_2 \sqrt{H} \sigma_{max}^{\frac{1}{2}}(W_H^u) + \|D\|_2 \right) \sigma_{min}^{-\frac{1}{2}}(W_H^u) \alpha_H,$$

where $\alpha_H := \max_{0 \leq k \leq H} \|A^k\|_2$.

The final bound

With the results of the last section, we can now bound the constants Γ_A and Γ_B used in Theorem 25. Rather than writing the explicit form of the constants we shall only analyze how they scale with system parameters. Recall Γ_A, Γ_B are defined as

$$\Gamma_A = \kappa_{CD} \Gamma_1' + \kappa_{CD} \Gamma_2' \|B_1\|_2 \|G_w(A_1)\|_2 \|G_w(A_2)\|_2$$
$$\Gamma_B = \kappa_{CD} \Gamma_2' \|G_w(A_2)\|_2,$$

where Γ_1', Γ_2' are dominated by the terms

$$\Gamma_1' \sim O\left(\alpha_{H,1} \alpha_{H,2} H \left\| G_{u,2} \right\|_2 \|P_{H,2}^\dagger\|_2 \right)$$
$$\Gamma_2' \sim O\left(\|g_2\|_2 (\|\mathcal{F}_1^\dagger\|_2 + \|\mathcal{F}_2^\dagger\|_2) \|P_{H,1}^\dagger\|_2 \|P_{H,2}^\dagger\|_2 (\|P_{H,1}\|_2 + \|P_{H,2}\|_2)(1 + \|G_{u,1}\|_2) \right).$$

Let us first revisit the collection of bounds we have derived:

1. $\|G_u(A, B)\|_2 \leq \sqrt{H \sigma_{max}(W_H^u(A, B))}$, $\|G_w(A)\|_2 \leq \sqrt{H \sigma_{max}(W_H^w(A))}$

2. $\|P_H(A, B)\|_2 = \left(\sigma_{max}(W_H^u(A, B)) \right)^{\frac{1}{2}}$, $\quad \|P_H^\dagger(A, B)\|_2 = \left(\sigma_{min}(W_H^u(A, B)) \right)^{-\frac{1}{2}}$

3. $\|\mathcal{F}^\dagger(A, B)\|_2 \leq \sigma_{min}^{-1}(D)$

4. $\|g\|_2 \leq \left(\|C\|_2 \sqrt{H} \sigma_{max}^{\frac{1}{2}}(W_H^u) + \|D\|_2 \right) \sigma_{min}^{-\frac{1}{2}}(W_H^u) \alpha_H$

5. $\alpha_H := \max_{0 \leq k \leq H} \|A^k\|_2$

6. $\kappa_{CD} = \frac{\max\{\sigma_{max}(C), \sigma_{max}(D)\}}{\min\{\sigma_{min}(C), \sigma_{min}(D)\}}$

Before we state the final bound, we require the following standard controllability result [5].

Lemma 40. *Let $\mathcal{S}$ be a compact set of matrices where each element $(A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}) \in \mathcal{S}$ represents a controllable linear dynamical system with equations $x(t+1) = Ax(t) + Bu(t) + w(t)$, state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}^m$ and disturbance $w(t) \in \mathbb{R}^n$. Then, there exists an FIR Horizon $H \leq n$, and positive scalar constants $\overline{\sigma}^w, \underline{\sigma}^w, \overline{\sigma}^u, \underline{\sigma}^u$ such that the following statements hold:*

- *For any $(A, B) \in \mathcal{S}$ and any initial state, ζ_0, there exists an input $u(0), \ldots, u(H-1)$, such that the system trajectory $x(t+1) = Ax(t) + Bu(t), \forall t \leq H-1, x(0) = \zeta_0$ satisfies $x(H) = 0$ at time H.*

- *For any $(A, B) \in \mathcal{S}$, the matrix $P_H = [A^{H-1}B, A^{H-2}B, \ldots, B] \in \mathbb{R}^{n\times Hm}$ is full column rank.*

- *For any $(A, B) \in \mathcal{S}$, the following FIR-SLS-constraint is feasible: There exist $\Phi^x[1], \ldots, \Phi^x[H] \in \mathbb{R}^{n\times n}$ and $\Phi^u[0], \ldots, \Phi^u[H-1] \in \mathbb{R}^{m\times n}$ such that:*

$$\Phi^x[0] = I, \quad \forall k = 0, ..., H-1: \quad \Phi^x[k+1] = A\Phi^x[k] + B\Phi^u[k], \; and \; \Phi^x[H] = 0$$

- *For any $(A, B) \in \mathcal{S}$, the corresponding grammians $W_H^u(A, B)$ and $W_H^w(A)$ are positive-definite and their singularvalues satisfy the inequalities:*

$$\underline{\sigma}^u \leq \sigma_{min}(W_H^u(A, B)), \qquad \sigma_{max}(W_H^u(A, B)) \leq \overline{\sigma}^u$$
$$\underline{\sigma}^w \leq \sigma_{min}(W_H^w(A)), \qquad \sigma_{max}(W_H^w(A)) \leq \overline{\sigma}^w.$$

We can not use $\underline{\sigma}_u, \overline{\sigma}_u, \underline{\sigma}_w, \overline{\sigma}_w$ in Lemma 40 in conjuncture of the bounds derived above to obtain

$$\Gamma_2' = O\left(\alpha_H \, \kappa_{CD} \, H \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{3}{2}}\right) \qquad \Gamma_1' = O\left(\alpha_H^2 H^{\frac{3}{2}} \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{1}{2}}\right) \tag{6.101}$$

and finally

$$\Gamma_A = O\left(\alpha_H^2 \, \kappa_{CD}^2 \, \|B_1\|_2 \, H^2 \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{3}{2}} \overline{\sigma}_w\right) \quad \Gamma_B = O\left(\alpha_H \, \kappa_{CD}^2 \, H^{\frac{3}{2}} \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{3}{2}} \overline{\sigma}_w^{\frac{1}{2}}\right).$$
$$\tag{6.102}$$

Theorem 26. *Let $C, D > 0$, and let $\mathcal{S}$ be a compact set of controllable systems with known FIR horizon H and constants $\underline{\sigma}_u, \overline{\sigma}_u, \underline{\sigma}_w, \overline{\sigma}_w$ as defined in Lemma 40. Then there are fixed constants Γ_A, Γ_B, such that for any two pairs of system matrices $(A_1, B_1), (A_2, B_2) \in \mathcal{S}$ the corresponding $\mathcal{H}_2$ optimal SLS-solutions of problem S_j (j arbitrary), denoted ϕ_1^{*j} and ϕ_2^{*j}, satisfy the following inquality:*

$$\|\phi_1^{j*} - \phi_2^{j*}\|_2 \leq \Gamma_A \|A_1 - A_2\|_F + \Gamma_B \|B_1 - B_2\|_F. \tag{6.103}$$

Furthermore, Γ_A and Γ_B satisfy

$$\Gamma_A = O\left(\alpha_H^2 \, \kappa_{CD}^2 \, \beta \, H^2 \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{3}{2}} \overline{\sigma}_w\right) \quad \Gamma_B = O\left(\alpha_H \, \kappa_{CD}^2 \, H^{\frac{3}{2}} \left(\frac{\overline{\sigma}_u}{\underline{\sigma}_u}\right)^{\frac{3}{2}} \overline{\sigma}_w^{\frac{1}{2}}\right), \tag{6.104}$$

where $\beta := \max\limits_{(A,B)\in\mathcal{S}} \|B\|_2$ and κ_{CD} stands for

$$\kappa_{CD} = \frac{\max\{\sigma_{max}(C), \sigma_{max}(D)\}}{\min\{\sigma_{min}(C), \sigma_{min}(D)\}}.$$

6.E Extensions to Non-Convex Parameter Set Setting

Representing model uncertainty as convex compact parameter sets is not always practical, sometimes potentially even impossible. Our approach can be readily extended to compact non-convex parameter sets $\mathcal{S}$, if those can be written as a finite union of convex sets $\bigcup_{i=1}^N \mathcal{P}_i$. This class of non-convex sets covers a large range of practical scenarios and the presented approach can be extended without losing stability guarantees. We can ensure by wrapping the proposed algorithm in a high-level routine SETSELECT, which runs the algorithm on the smaller convex sets $\mathcal{P}_i$ until they become entirely inconsistent:

1. At $t = 0$, we select an arbitrary convex set $\mathcal{P}_{k_0}$ and perform consistent model chasing with CONSIST as before.

2. If at some point $\mathcal{P}_{k_0}$ becomes entirely inconsistent, we select an arbitrary set $\mathcal{P}_{k_1}$ from the remaining collection $\{\mathcal{P}_1, \ldots, \mathcal{P}_N\} \setminus \mathcal{P}_{k_0}$ and restart CONSIST with that set $\mathcal{P}_{k_1}$. If $\mathcal{P}_{k_1}$ is also entirely inconsistent, repeat that selection process.

Per definition, the above algorithm never violates consistency. Because there are finitely many convex sets $\mathcal{P}_{k_i}$, the cost accrued due to restarting CONSIST scales up the total movement cost of the convex counterpart by a fixed constant. Overall, the stability proof is not impacted.

Part III

Application in Sustainable Energy Systems

Chapter 7

SAFE CONTROL FOR VOLTAGE REGULATION WITH AN UNKNOWN GRID TOPOLOGY

So far, we have focused on analyzing the theoretical underpinning of the uncertainty set-based learning and control framework. From the convergence and performance guarantees in stochastic settings to the stability guarantees in adversarial settings, the framework provides flexible and principled integration of data-driven learning methods and model-based controllers.

In this chapter, we will explore how these algorithmic ideas can be applied to sustainable energy systems. In particular, we will study the voltage control problem in the distribution network. Voltage control generally requires accurate information about the grid's topology in order to guarantee network stability. However, accurate topology identification is a challenging problem for existing methods, especially as the grid is subject to increasingly frequent reconfiguration due to the adoption of renewable energy. Further, running existing control mechanisms with incorrect network information may lead to unstable control. To address this challenge, we instantiate the framework presented in Part II and combine nested convex body chasing algorithms with a robust predictive controller to achieve provably finite-time convergence to safe voltage limits where there is uncertainty in both the network topology as well as load and generation variations. Even though we develop the theoretical results under the assumption of linear system dynamics, our experiments show that the algorithm continues to stabilize the voltage in realistic nonlinear simulations with real-world data from the Southern California Edison utility under the partial observation and partial control settings. This chapter is mainly based on the following papers:

[1]   C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Robust online voltage control with an unknown grid topology," *Proceedings of the Thirteenth ACM international conference on future energy systems (e-Energy)*, pp. 240–250, 2022. DOI: 10.1145/3538637.3538853.

[1]   C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Online learning for robust voltage control under uncertain grid topology," *IEEE Transactions on Smart Grid*, 2024. DOI: 10.1109/TSG.2024.3383804.

7.1 Introduction

Operators of electricity distribution grids must maintain voltages at each bus within certain operating limits, as deviations from such limits may damage electrical equipment and cause power outages [63], [64]. This "voltage control" or "voltage regulation" problem has been well-studied, *e.g.*, [256]–[258] and the references therein. Voltage control devices and algorithms aim to guarantee grid stability and minimize the costs associated with control inputs. While classic voltage regulation devices such as tap-changing transformers are effective in dealing with *slow* voltage variations [259], [260], increasing penetration of renewables leads to faster variations, and a growing body of literature has focused on inverter-based controllers that can respond quickly by adjusting their active and reactive power set-points. Most of these works cast voltage control as an optimization problem and then propose different centralized or decentralized algorithms depending on the communication infrastructure.

Typically, voltage control algorithms assume *exact knowledge* of the underlying grid topology. This includes centralized controllers such as algorithms based on model predictive control (MPC) which optimize control decisions for a short-term horizon. [261] uses MPC to manage distributed generation and energy storage systems, whereas [262] proposes a robust MPC controller that is robust to uncertainty in the forecasts of future loads and solar generation.

However, the exact grid topology and line parameters are often not known, and using existing voltage control algorithms with incorrect grid information may lead to problems with grid stability [263], [264]. For example, parts of the grid may undergo reconfiguration due to load balancing or unplanned maintenance, as frequently as every hour of the day [66], [265]–[267]. This problem is exacerbated by the increasing integration of distributed energy resources (DERs), such as photovoltaic (PV) and storage devices. Especially in distribution grids, where DERs are not owned or operated by the electricity utility, the grid operator may lack up-to-date information about the grid topology [65]. While a grid operator can install sensors to help identify the current network topology, unless such sensors are densely deployed (at great cost), uncertainty about the topology remains. Thus, distribution grid operators cannot expect to operate with perfect topology information and the design of voltage control algorithms robust to unknown grid topology is crucial.

There are several families of existing algorithms that do not require knowing the network topology: decentralized controllers, model-free controllers, and controllers

that first try to infer the network topology. While decentralized voltage control algorithms are generally efficient to implement, such controllers lack voltage stability guarantees when the load is time-varying [257], [268]–[271]. Likewise, model-free controllers based on deep reinforcement learning do not require knowing the network topology, but they generally have no performance or voltage stability guarantees and are therefore not suitable for safety-critical infrastructure [272]–[276]. Some recent works [277]–[279] have proposed methods for introducing stability guarantees for model-free deep reinforcement learning approaches. Their main tool is Lyapunov stability theory, from which a structural constraint for stable controllers is derived, and policy optimization with the constraint is performed. However, their stability guarantees are only valid over an infinite time horizon, and achieving good performance with deep reinforcement learning generally requires large amounts of historical training data. In contrast, our proposed framework jointly learns the system model (consistent with data) and stable controller in an online fashion, achieving a finite-mistake guarantee and good performance without relying on historical data.

Another standard approach for handling uncertainty about network topology is to first estimate the topology and line parameters using a form of system identification with data and then apply a standard voltage control algorithm using the identified network topology. There is a growing literature of such data-driven methods, *e.g.*, [65], [263], [264], [280]–[288]. A common approach is to leverage least squares for system model estimation. The estimation and therefore control guarantees depend on statistical modeling of measurement noise (*e.g.*, Gaussian). In contrast, we leverage online learning in order to be robust against any bounded disturbances, such as modeling errors and adversarial noise. While least squares-based algorithms focus on *asymptotic* estimation convergence, *e.g.* [289], [290], we present a *finite* mistake guarantee that is crucial for safe *transient* system behavior.

Another prominent approach is to use graphical models for topology reconstruction [291], via maximum likelihood methods while enforcing other structural restrictions like low-rank and sparsity. However, these methods that first perform some form of system identification have drawbacks. First, the estimated topology and/or system dynamics may be imperfect [292], and applying standard voltage control algorithms using these imperfect estimates may still lead to system instability. Second, these methods either assume access to historical data or require acquiring data online over hundreds of time steps, during which the stability of the system is ignored [65], [291]. In contrast, our proposed approach does not perform system

identification separately from control; the joint operation of our robust controller with the system dynamics estimation gives rise to our stability guarantee.

Contributions

We propose a new approach for voltage control over an *uncertain grid topology* with *uncertain maximum variability of load and generation* entities in the grid that does not perform system identification and voltage control separately. Instead, our approach robustly learns to stabilize voltage within the desired limits directly, *without any prior knowledge of the topology and without needing to precisely learn the topology*.

Our approach takes ideas from online nested convex body chasing (CBC) [240] and robust predictive control and combines them using an uncertainty set-based learning and control framework [13], [99], [111], [293] to apply them to voltage control for the first time. Intuitively, we use a nested CBC algorithm in order to track the set of topologies that are consistent with the observed voltage measurements—as more measurements are taken the set of consistent topologies shrinks (and so the sets are nested). As these measurements are taken, a form of robust predictive control is used for voltage control, where the robustness guarantee is used to ensure the uncertainty about the topology can be handled. Our main result (Theorem 27) provides a finite error stability bound for the overall controller, which is summarized in Algorithm 6. This represents the first voltage control algorithm that is provably robust to uncertainty about network topology.

In addition to providing theoretical guarantees, we demonstrate the effectiveness of our proposed approach using a case study of a 56-bus distribution grid from the Southern California Edison (SCE) utility [294]. In this setting, we give the controller no prior information about the topology of the grid, yet the controller quickly narrows down the set of topologies and line parameters that are consistent with its observations and adjusts reactive power generation to keep voltages within desired safety limits when faced with disturbance. In fact, our controller's performance nearly matches that of controllers which assume perfect knowledge of the topology, even when given only partial observations of bus voltages.

Beyond the linear model, we test the performance of the proposed algorithm with a more realistic nonlinear power flow model with partial control and partial observation. Even though the design of our method is based on a linear approximation to the power flow model, our method still performs well for the nonlinear system. We also

demonstrate how to incorporate existing partial knowledge of the grid topology and network line parameters into the algorithm. We show that incorporating such prior knowledge can improve the performance of our algorithm.

7.2 Model

We study voltage control on an unknown grid topology. We consider a radial (tree-structured) power distribution network represented as a connected directed graph $G = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N} = \{0, 1, 2, \ldots, n\}$ is the set of buses (nodes) and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ is the set of lines (directed edges). Let the network be rooted at bus 0 (the substation or slack bus), and let other buses be branch buses. Let $C \subseteq \mathcal{N}$ denote the subset of buses with controllable reactive power injection. Because the network is radial and rooted at bus 0, there is a unique path $\mathcal{P}_i$ from bus 0 to any other bus i. For branch buses, let $v \in \mathbb{R}^n$ be their squared voltage magnitudes and $p + \mathbf{i}q$ be their complex power injection, where $p \in \mathbb{R}^n$ (units W) is the net active power injection, and $q \in \mathbb{R}^n$ (units Var) is the net reactive power injection. The DistFlow branch equations [295] for a distribution grid are as follows, for all $j \in \mathcal{N}$ and $(i, j) \in \mathcal{E}$:

$$-p_j = P_{ij} - r_{ij}l_{ij} - \sum_{k:(j,k)\in\mathcal{E}} P_{jk} \tag{7.1a}$$

$$-q_j = Q_{ij} - x_{ij}l_{ij} - \sum_{k:(j,k)\in\mathcal{E}} Q_{jk} \tag{7.1b}$$

$$v_j = v_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)l_{ij} \tag{7.1c}$$

$$l_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{v_i} \tag{7.1d}$$

where P_{ij} and Q_{ij} represent the active power and reactive power flow on line (i, j), and $r_{ij}, x_{ij} > 0$ are the real-valued line resistance and reactance (units Ω). Equations (7.1a) and (7.1b) represent the real and reactive power conservation at bus j, and (7.1c) represents the voltage drop from bus i to bus j.

Assuming the branch power losses $(r_{ij}l_{ij}, x_{ij}l_{ij})$ are negligible yields the simplified DistFlow equations [296], which can be rearranged into

$$v = R^\star p + X^\star q + v^0 \mathbb{1}_n \tag{7.2}$$

where $v^0 \in \mathbb{R}^n$ is the known, constant squared voltage magnitude at the substation, and $R^\star, X^\star \in \mathbb{S}^n$ are computed from the network topology and line parameters

$$R_{ij}^\star := 2 \sum_{(h,k)\in\mathcal{P}_i\cap\mathcal{P}_j} r_{hk}, \quad X_{ij}^\star := 2 \sum_{(h,k)\in\mathcal{P}_i\cap\mathcal{P}_j} x_{hk}, \quad i, j \in [n] \tag{7.3}$$

with $[n] := \{1, \ldots, n\}$ [269]. ($\mathbb{S}^n$ is the set of symmetric $n \times n$ matrices.) $R^\star, X^\star$ are positive definite with nonnegative entries [297], and the largest entry of each row of these matrices is along the diagonal, since

$$X_{ij}^\star = 2 \sum_{(h,k)\in\mathcal{P}_i\cap\mathcal{P}_j} x_{hk} \leq 2 \sum_{(h,k)\in\mathcal{P}_i} x_{hk} = X_{ii}^\star \tag{7.4}$$

and likewise for $R_{ij}^\star \leq R_{ii}^\star$.

We assume that the active power injection p is exogenous but that reactive power at each bus can be decomposed as $q = q^c + q^e$, where q^c is the "controllable" component and q^e is the "exogenous" (*i.e.*, uncontrollable) component. Following [269], we define $v^{\mathrm{par}} = R^\star p + X^\star q^e + v^0 \mathbb{1}_n \in \mathbb{R}^n$ ("par" stands for "partial") representing the exogenous effects on voltage. Then, $v = X^\star q^c + v^{\mathrm{par}}$, which can be modeled as a discrete-time linear system

$$v(t+1) = X^\star q^c(t) + v^{\mathrm{par}}(t). \tag{7.5}$$

Substituting $u(t) = q^c(t) - q^c(t-1)$ (change in controllable reactive power injection) and $w(t) = v^{\mathrm{par}}(t) - v^{\mathrm{par}}(t-1)$ (change in exogenous noise) yields the linear dynamical system

$$v(t+1) = v(t) + X^\star u(t) + w(t). \tag{7.6}$$

The voltage control problem [294] is to drive the squared voltage magnitudes of each bus from an initial state $v(1) \in \mathbb{R}^n$ into a given multi-dimensional interval $[\underline{v}, \overline{v}] \subset \mathbb{R}^n$; it is possible that $v(1)$ does not start within the interval due to some large initial disturbance. For all $t \geq 2$, the voltage control algorithm aims to maintain $v(t)$ within $[\underline{v}, \overline{v}]$, ideally as close as possible to a "nominal" value $v^{\mathrm{nom}} \in [\underline{v}, \overline{v}]$, typically $v^{\mathrm{nom}} = (\underline{v} + \overline{v})/2$. The cost for deviating from v^{nom} is measured by $\|v(t) - v^{\mathrm{nom}}\|_{P_v}^2$ for some positive semidefinite matrix P_v, where $\|x\|_A^2 := x^\top A x$.

At each time step, buses may change their reactive power injection $q^c(t)$ in order to regulate the voltage close to v^{nom}. The reactive power injection (including $q^c(0)$) is limited within a given bound $[\underline{q}, \overline{q}] \subset \mathbb{R}^n$. Buses not in C do not have any ability to control the reactive power injection: $\forall i \notin C. \underline{q}_i = \overline{q}_i = 0$. We do not place any hard "ramp constraints" on $u(t)$. However, we impose a quadratic ramping cost $\|u(t)\|_{P_u}^2$ where P_u is a positive semidefinite matrix.

In summary, the voltage control problem is to determine an online sequence of reactive power injections $q^c(1), q^c(2), \ldots$ to drive voltages $v(t)$ to a desired interval $[\underline{v}, \overline{v}]$ while minimizing voltage violation and control costs $\|v(t) - v^{\mathrm{nom}}\|_{P_v}^2 + \|u(t)\|_{P_u}^2$. In this work, we solve the voltage control problem in the setting where $X^\star$ is unknown.
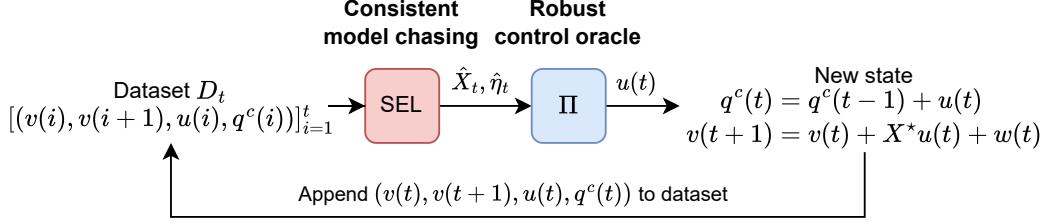
Figure 7.1: Online robust voltage control

7.3 Robust Online Voltage Control

In this section we introduce our robust online voltage control algorithm (Algorithm 6) and its performance bound (Theorem 27), which is the main result of this chapter.

Algorithm

As shown in Figure 7.1, the algorithm has two main components: a consistent model chasing algorithm SEL (Algorithm 6, step 1) and a robust control oracle Π (Algorithm 6, step 2). SEL and Π are combined by adapting ideas from [13].

The model chasing algorithm SEL selects a consistent model for the robust control oracle Π out of all plausible models that are consistent with the online observations and prior knowledge of the grid, where the model uncertainty set is constructed with set membership estimation [115], [293]. The selection may use any competitive NCBC algorithm, which is the online problem of choosing a sequence of points within sequentially nested convex sets, with the aim of minimizing the sum of distances between the chosen points [240]. In our experiments, we use a simple projection-based NCBC algorithm, detailed in Section 7.5.

The robust control oracle Π is a novel robust predictive controller (Theorem 29). The robustness guarantee of Π is necessary for the analysis which integrates SEL with Π to provide the finite mistake guarantee of the overall algorithm. We remark that other choices for either component are possible, as long as they provide the guarantees needed in the analysis in Section 7.4.

Intuitively, SEL and Π are combined in a way such that SEL always reduces the uncertainty about the unknown model whenever Π outputs an action that causes a voltage limit violation. This means that Π cannot take too many "bad" actions before the system uncertainty is small.

Algorithm 6: Online Robust Voltage Controller

Input:

- desired nominal squared voltage magnitude: $v^{\text{nom}} \in \mathbb{R}^n$
- limits on the squared voltage magnitude: $[\underline{v}, \overline{v}] \subset \mathbb{R}^n$
- limits on the reactive power injection: $[\underline{q}, \overline{q}] \subset \mathbb{R}^n$
- initial state: $v(1), q^c(0) \in \mathbb{R}^n$
- state and action cost matrices: $P_v, P_u \in \mathbb{S}_+^n$
- compact convex uncertainty set for the model parameter: $\mathcal{X} \subset \mathbb{S}_+^n \cap \mathbb{R}_+^{n \times n}$
- compact convex uncertainty set for exogenous voltage quantities: $\mathcal{V}^{\text{par}} \subset \mathbb{R}^n$
- upper bound for noise: $\overline{\eta} > 0$
- robustness padding: $\epsilon > 0$
- weight for slack variable: $\beta > 0$
- weight for noise accuracy: $\delta > 0$

Initialize: Initialize an empty trajectory $D_0 = [\ ]$. Set $t = 1$

Procedure:

(1) If $t = 1$, initialize estimate of model parameters $\widehat{X}_1 \in \mathcal{X}$.
Otherwise, query the model chasing algorithm for a new consistent parameter estimate: $(\widehat{X}_t, \widehat{\eta}_t) \leftarrow \text{SEL}[D_t]$.

$$\text{SEL}[D_t] := \text{NCBC}(P_t, \widehat{X}_{t-1}, \widehat{\eta}_{t-1}) \tag{7.7a}$$

$$P_t := \left\{ (\widehat{X}, \widehat{\eta}) \ \middle| \ \begin{array}{l} \widehat{X} \in \mathcal{X}, \ \widehat{\eta} \in [0, \overline{\eta}] \\ \forall (v_i, v_{i+1}, u_i, q_i^c) \in D_t : \\ \|v_{i+1} - v_i - \widehat{X}_t u_i\|_\infty \leq \widehat{\eta} \\ v_{i+1} - \widehat{X} q_i^c \in \mathcal{V}^{\text{par}} \end{array} \right\} \tag{7.7b}$$

(2) Query the robust control oracle for the next control action: $u(t) \leftarrow \Pi_{\widehat{X}_t, \widehat{\eta}_t}(v(t))$.

$$\Pi_{\widehat{X}_t, \widehat{\eta}_t} : \ \min_{u, \xi} \ \|\widehat{v}' - v^{\text{nom}}\|_{P_v}^2 + \|u\|_{P_u}^2 + \beta \xi^2 \tag{7.8a}$$

$$\text{s.t. } u \in \mathbb{R}^n, \ \xi \in \mathbb{R}_+ \tag{7.8b}$$

$$\underline{q} \leq q^c(t-1) + u \leq \overline{q} \tag{7.8c}$$

$$\widehat{v}' = v(t) + \widehat{X}_t u \tag{7.8d}$$

$$k = \widehat{\eta}_t + \rho \left(\frac{1}{\delta} + \|u\|_2 \right) \tag{7.8e}$$

$$\underline{v} + (k - \xi)\mathbb{1} \leq \widehat{v}' \leq \overline{v} - (k - \xi)\mathbb{1} \tag{7.8f}$$

where $\rho = \delta\epsilon / (1 + \delta\|\overline{q} - \underline{q}\|_2)$.

(3) Apply the control action $u(t)$. Observe the system transition to $v(t+1) = v(t) + X^\star u(t) + w(t)$ and $q^c(t) = q^c(t-1) + u(t)$.

(4) Append $(v(t), v(t+1), u(t), q^c(t))$ to the trajectory:

$$D_t = [(v(i), v(i+1), u(i), q^c(i))]_{i=1}^t .$$

(5) Increment $t \leftarrow t + 1$. Repeat from Step (1).

Assumptions

Before presenting the main results, we introduce three assumptions that underlie our analysis and discuss why they are both needed and practical.

Assumption 13. *The change in noise is bounded as*

$$\forall t : \quad \|w(t)\|_\infty \leq \eta^\star,$$

where $w(t) = v^{par}(t) - v^{par}(t-1)$. $\eta^\star \in [0, \overline{\eta}]$ is a constant (possibly unknown), while $\overline{\eta}$ is a known upper-bound.

This first assumption is standard and bounds the noise in the dynamics. It represents realistic behavior in power systems where the active and exogenous reactive power injections do not vary dramatically between time steps, as can be seen by expanding $w(t)$:

$$
\begin{aligned}
w(t) &= v^{\mathrm{par}}(t) - v^{\mathrm{par}}(t-1) \\
&= R^\star(p(t) - p(t-1)) + X^\star(q^e(t) - q^e(t-1)).
\end{aligned}
$$

For example, if the net active and exogenous reactive power injection is the same at time steps t and $t-1$, then $w(t) = 0$.

An unknown $\eta^\star$ indicates uncertainty in the maximum variability of the exogenous power injections. Our inclusion of both an unknown $\eta^\star$ and a known upper-bound $\overline{\eta}$ allows more flexibility in the algorithmic design and the incorporation of prior knowledge.

Assumption 14. *The true model $X^\star$ lies within a known compact, convex uncertainty set $\mathcal{X} \subset \mathbb{S}_+^n \cap \mathbb{R}_+^{n \times n}$. ($\mathbb{S}_+^n$ is the set of $n \times n$ positive semidefinite matrices, and $\mathbb{R}_+^{n \times n}$ is the set of $n \times n$ matrices with nonnegative entries.)*

Our second assumption bounds the uncertainty about the network topology and line parameters. It ensures that the unknown true model parameters $X^\star$ belong to a compact, convex set $\mathcal{X}$, which is a minimal assumption necessary for proving an analytic guarantee. $P_1 = \mathcal{X} \times [0, \overline{\eta}]$ forms the initial "consistent set" (see Definition 7.4.1) for our consistent model chasing algorithm SEL.

This assumption is realistic, as a grid operator should have at least some prior knowledge about the distribution grid topology and the range of possible line parameters, even if they do not have the exact values. In cases where the grid has

multiple possible topologies due to switches, $\mathcal{X}$ could be set to the convex hull of the corresponding X matrices.

Definition 7.3.1 ($\|\cdot\|_{\triangle}$ and $\|\cdot\|_{\triangle,\delta}$)**.** *For any matrix $X \in \mathbb{S}^n$ and scalars $\eta, \delta \geq 0$, define*

$$\|X\|_{\triangle} := \|\mathrm{vech}(X)\|_2 = \sqrt{\sum_{i=1}^{n} \sum_{j=i}^{n} X_{ij}^2}$$

$$\|(X,\eta)\|_{\triangle,\delta} := \sqrt{\delta^2\eta^2 + \sum_{i=1}^{n} \sum_{j=i}^{n} X_{ij}^2} = \sqrt{\delta^2\eta^2 + \|X\|_{\triangle}^2}.$$

For any sets $\mathcal{X} \subseteq \mathbb{S}^n$ and $A \subseteq \mathbb{R}$, we define diameters $\mathrm{diam}(\mathcal{X})$ and $\mathrm{diam}(\mathcal{X} \times A)$ with respect to the norms $\|\cdot\|_{\triangle}$ and $\|\cdot\|_{\triangle,\delta}$, respectively.

These norms isometrically map our parameter space to Euclidean space, enabling us to take advantage of known results on NCBC within Euclidean space. For the norm $\|\cdot\|_{\triangle,\delta}$, the hyperparameter δ trades off the weight between X and η in the norm. The choice of δ is discussed in Section 7.5.

In practice, we consider uncertainty sets of the form

$$\mathcal{X}_\alpha = \left\{ X \in \mathbb{S}^n_+ \cap \mathbb{R}^{n \times n}_+ \;\middle|\; \begin{array}{l} \|X - X^\star\|_{\triangle} \leq \alpha \|X^\star\|_{\triangle}, \\ \forall i, j \in [n] : X_{ij} \leq X_{ii} \end{array} \right\}$$

with $\mathrm{diam}(\mathcal{X}_\alpha) = 2\alpha \|X^\star\|_{\triangle}$. A larger α yields a larger uncertainty set. From Section 7.2 (*e.g.*, (7.4)), we know that $X^\star \in \mathcal{X}_\alpha$.

Furthermore, we can incorporate partial knowledge we may have of the network topology and/or line parameters by adding constraints to the description of $\mathcal{X}$. For example, if we know that the lowest common ancestor between buses i, j in the network is bus k, then we can add the following constraint on X, which is a consequence of (7.3):

$$X_{ij} = \begin{cases} 0, & k = 0 \\ X_{kk}, & \text{otherwise.} \end{cases} \tag{7.9}$$

If we additionally know the values for some line parameters x_{ij}, we may be able to further constrain some entries of X, again by applying (7.3).

Assumption 15. *There exists a compact, convex set $\mathcal{V}^{par} \subset \mathbb{R}^n$ such that $\forall t \geq 0 :$ $v^{par}(t) \in \mathcal{V}^{par}$. Furthermore, for some known $\epsilon > 0$,*

$$\forall v^{par} \in \mathcal{V}^{par}, X \in \mathcal{X}.$$

$$\exists q^c \in [\underline{q}, \overline{q}] \ s.t. \ Xq^c + v^{par} \in [\underline{v} + (\overline{\eta} + \epsilon)\mathbb{1}, \ \overline{v} - (\overline{\eta} + \epsilon)\mathbb{1}].$$

Our final assumption is about the existence of feasible control actions for the robust control oracle. This assumption can be interpreted as either a bound on the noise, or a requirement that the controllable reactive power injection be flexible enough to satisfy the demand of any admissible noise. It represents the reasonable assumption that a grid operator should have installed enough controllable reactive power injection capability to perform voltage control. Intuitively, the $\overline{\eta}$ padding is required for robustness to the noise $w(t)$, while the ϵ padding is required for robustness to model uncertainty (*i.e.*, uncertainty about $X^\star$).

Main result

We now state our main result, which is a finite-error bound for Algorithm 6.

Theorem 27 (Main Result)**.** *Under Assumptions 13 to 15, Algorithm 6 ensures that the voltage limits will be violated at most $\frac{2\gamma(m)}{\rho} \operatorname{diam}(\mathcal{X} \times [0, \overline{\eta}]) + 1$ times, where $\rho = \frac{\delta\epsilon}{1+\delta\|\overline{q}-\underline{q}\|_2}$ and $\gamma(m)$ is the competitive ratio of the NCBC algorithm in m-dimensional Euclidean space, where $m = 1 + \frac{n(n+1)}{2}$.*

Furthermore, if $\eta^\star$ is known, then the voltage limits will be violated at most $\frac{2\gamma(m)}{\rho} \operatorname{diam}(\mathcal{X}) + 1$ times, where $\rho = \frac{\epsilon}{\|\overline{q}-\underline{q}\|_2}$ and $m = \frac{n(n+1)}{2}$.

To the best of our knowledge, this result is the first provable stability bound for voltage control in a setting where the network topology is unknown. It highlights that Algorithm 6 can ensure stability even after *unknown* changes to the network topology, *e.g.*, due to maintenance, failures, etc., without the need to perform system identification while remaining robust to any bounded and potentially adversarial perturbations satisfying Assumptions 13 and 15.

Intuitively, this result guarantees that the model chasing algorithm SEL will learn a "good enough" model for voltage regulation quickly. When the robust controller Π makes a mistake, the model chasing algorithm will learn from that mistake and significantly reduce the set of consistent models. Because the initial set of consistent models is bounded, and this set shrinks a significant amount after each mistake, the total number of mistakes is bounded. Note that this finite mistake bound implies finite-time convergence to safe voltage limits without an explicit finite-time bound.

To interpret the error bounds in Theorem 27, we notice that they are proportional to the diameter of the parameter space and the competitive ratio $\gamma(m)$ of the NCBC

algorithm, and inversely proportional to the oracle robustness margin ρ. Because of computational tractability concerns, our experiments implement SEL with a greedy projection-based NCBC algorithm with $\gamma_{\text{proj}}(m) = \pi(m-1)m^{m/2}$ [240], rather than the state-of-the-art Steiner point method which can achieve $\gamma_{\text{Steiner}}(m) = m/2$ [191]. As our case studies show, in practice the projection-based NCBC algorithm performs much better than the worst-case bound. We note that any other NCBC algorithm with a finite competitive ratio can be used in (7.7a) in Algorithm 6. Investigating whether widely-used estimation methods, like least squares, have a finite competitive ratio would be an interesting avenue for future research.

Note that for Theorem 27 to hold, the optimization problem for the robust control oracle Π should first be solved without the slack variable ξ in Algorithm 6. This ensures that if $(\widehat{X}_t, \widehat{\eta}_t)$ is sufficiently close enough to the true model, then the algorithm will not make a mistake. In the case that Π is infeasible initially (*e.g.*, when the initial model estimate is far from the true model), it should be solved again with a slack variable, which ensures feasibility. However, solving Π twice is unnecessary in practice, and so we have written Algorithm 6 to reflect its practical implementation.

We outline a proof of Theorem 27 in the next section. We want to highlight one piece of that proof that is of independent interest. In particular, a major step in the proof is to provide a feasibility guarantee for the robust control oracle component Π of the algorithm, which is done in Theorem 29.

7.4 Proofs

We now prove our main result Theorem 27. Our proof builds on and adapts the approach of [13], which outlines a general framework for integrating model chasing and robust control via uncertainty sets constructed from set membership estimation. To explain the general framework, we first consider a discrete-time nonlinear dynamical system

$$x_{t+1} = f_*(x_t, u_t) + w_t, \qquad x_0 \text{ given}, \qquad (f_*, \mathbf{w}) \in \mathcal{F},$$

where $x \in \mathcal{S} \subseteq \mathbb{R}^n$ is the system state and $u \in \mathcal{U} \subseteq \mathbb{R}^m$ is the control input. The unknown function f_* and disturbance sequence $\mathbf{w} \in \ell^\infty(\mathbb{Z}_+; \mathbb{R}^n)$ belong to an uncertainty set $\mathcal{F}$, and the disturbance is bounded as $\|\mathbf{w}\|_\infty \leq \overline{\eta}$. Assume that $\mathcal{F}$ has a *compact parametrization* $(\mathbb{T}, \mathsf{K}, d)$, where $\mathbb{T} : \mathsf{K} \to \wp(\mathcal{F})$ is a mapping from a parameter space K to a set of functions and disturbances such that $\mathcal{F} \subseteq \bigcup_{\theta \in \mathsf{K}} \mathbb{T}[\theta]$. $\wp(\mathcal{F})$ denotes the powerset of $\mathcal{F}$. Let d denote a metric on K, so (K, d) is a compact metric space.

The control objective is specified as a sequence of indicator "goal" functions $\mathcal{G} = (\mathcal{G}_0, \mathcal{G}_1, \dots)$. Each $\mathcal{G}_t : \mathcal{X} \times \mathcal{U} \rightarrow \{0, 1\}$ encodes a desired condition per time step t:

$$\mathcal{G}_t(x_t, u_t) = \mathbb{1}[x_t, u_t \text{ violate desired condition at time } t].$$

The main result of [13] specifies a set of sufficient conditions for a finite-mistake guarantee—*i.e.*, $\sum_{t=0}^{\infty} \mathcal{G}_t(x_t, u_t) < \infty$. These conditions decouple online robust control into separate online learning and robust control components. The online learning component requires a consistent model chasing algorithm SEL, which takes as input the current observed trajectory $D_t = [(x_i, x_{i+1}, u_i)]_{i=1}^t$ and outputs an estimated parameter $\theta_t \in \mathsf{K}$ which must be *consistent* with D_t.

Definition 7.4.1 (Consistent Parameter). *We say $\theta \in \mathsf{K}$ is consistent with D_t if there exists $(f, \mathbf{w}) \in \mathbb{T}[\theta]$ such that*

$$\forall (x_t, x_{t+1}, u_t) \in D_t : \ x_{t+1} = f(x_t, u_t) + w_t.$$

Let P_t denote the set of all parameters consistent with D_t; P_t is called the *consistent set*. We say SEL is γ-competitive if $\sum_{t=1}^{\infty} d(\theta_t, \theta_{t-1}) \leq \gamma \max_{\theta \in \mathsf{K}} d(P_\infty, \theta)$ holds for a fixed constant $\gamma > 0$, which we call the *competitive ratio*.

The robust control component requires a control oracle Π, which given the current state x_t and a parameter θ_t, outputs a control action $u_t = \Pi_{\theta_t}(x_t)$ that is robust for all systems that are close to θ_t. In particular, we call a control oracle ρ-*robust* for control objective $\mathcal{G}$, if all trajectories in $S^\Pi[\rho; \theta]$ achieve $\mathcal{G}$ after finitely many mistakes. $S^\Pi[\rho; \theta]$ is defined as the set of all possible trajectories generated by $\Pi_{\widehat{\theta}}$ for all $\widehat{\theta}$ such that $d(\theta, \widehat{\theta}) \leq \rho$:

$$S^\Pi[\rho; \theta] = \left\{ \begin{array}{l} D_\infty = [(x_t, x_{t+1}, u_t)]_{t=1}^\infty : \\ u_t = \Pi_{\widehat{\theta}}(x_t) \\ x_{t+1} = f(x_t, u_t) + w_t \end{array} \left| \begin{array}{l} (f, \mathbf{w}) \in \mathbb{T}[\theta], \\ d(\widehat{\theta}, \theta) \leq \rho \end{array} \right. \right\}$$

Due to the page limit, we refer readers to [13] for a more detailed discussion of consistent model chasing algorithms and ρ-robust control oracles. As a summary, if SEL chases consistent models and Π is a robust oracle for $\mathcal{G}$, then the resulting $A_\Pi(\text{SEL})$ algorithm achieves a finite mistake guarantee, which is stated in the following.

Theorem 28. *[13, Theorem 2.5] Assume that SEL chases consistent models and Π is a robust oracle for objective $\mathcal{G}$. Then for any starting point x_0 and trajectory*

$[(x_t, u_t)]_{t=0}^{\infty}$ *generated by* $\mathcal{A}_{\Pi}(\mathsf{SEL})$ *(illustrated in Figure 7.1), the following mistake guarantees hold: (i) If* Π *is robust, then* $\sum_{t=0}^{\infty} \mathcal{G}_t(x_t, u_t) < \infty$; *(ii) If* Π *is uniformly* ρ*-robust and* SEL *is* γ*-competitive, then*

$$\sum_{t=0}^{\infty} \mathcal{G}_t(x_t, u_t) < \max\left\{1, M_{\rho}^{\Pi}\right\} \left(\frac{2\gamma}{\rho} \operatorname{diam}(\mathsf{K}) + 1\right)$$

where M_{ρ}^{Π} *denotes the worst case total mistakes of the* ρ*-robust control oracle* Π.

To apply Theorem 28 to prove Theorem 27, we need to prove that (i) the proposed algorithm (7.7) chases consistent models and has a bounded competitive ratio, and (ii) the proposed robust algorithm in (7.11) is a ρ-robust control oracle, for bounded disturbance in the system topology. In particular, the correspondence of the definitions is as follows. We have $\theta = (X, \eta)$, and

$$\mathsf{K} = \mathcal{X} \times [0, \overline{\eta}], \quad v(1), q^c(0) \text{ given}$$

$$d((X, \eta), (X', \eta')) = \|(X, \eta) - (X', \eta')\|_{\triangle, \delta}$$

$$\mathbb{T}[(X, \eta)] = \left\{ (f, \mathbf{w}) \left| \begin{array}{l} f(v, u) = v + Xu, \ \|\mathbf{w}\|_{\infty} \le \eta, \\ \forall t \ge 0 : \widehat{v_0^{\mathrm{par}}} + \sum_{\tau=1}^{t} w(\tau) \in \mathcal{V}^{\mathrm{par}} \\ \text{where } \widehat{v_0^{\mathrm{par}}} := v(1) - Xq^c(0) \end{array} \right. \right\}$$

$$\mathcal{F} = \bigcup_{(X, \eta) \in \mathcal{X} \times [0, \overline{\eta}]} T[(X, \eta)]$$

$$\mathcal{G}_t(v(t)) = \mathbb{1}[v(t) \in [\underline{v}, \overline{v}]].$$

We begin by proving that the set P_t defined in (7.7b) in Algorithm 6 is consistent with the trajectory D_t.

Lemma 41 (SEL is consistent). *Suppose* D_T *is a trajectory of voltage measurements and control actions taken up to time* T:

$$D_T = [(v(t), v(t+1), u(t), q^c(t))]_{t=1}^{T}.$$

The set

$$P_T := \left\{ (\widehat{X}, \widehat{\eta}) \left| \begin{array}{l} \widehat{X} \in \mathcal{X}, \ \widehat{\eta} \in [0, \overline{\eta}], \\ \forall (v(t), v(t+1), u(t), q^c(t)) \in D_T : \\ \|v(t+1) - v(t) - \widehat{X}u(t)\|_{\infty} \le \widehat{\eta} \\ v(t+1) - \widehat{X}q^c(t) \in \mathcal{V}^{par} \end{array} \right. \right\} \tag{7.10}$$

is a consistent set for D_T, i.e., $(\widehat{X}, \widehat{\eta})$ is consistent (Definition 7.4.1) if and only if $(\widehat{X}, \widehat{\eta}) \in P_T$.

Proof. Consider any $(\widehat{X}, \widehat{\eta}) \in P_T$. For $t \in [T]$, define

$$\widehat{f}(v, u) := v + \widehat{X}u, \qquad \widehat{w}(t) := v(t+1) - v(t) - \widehat{X}u(t)$$

so $\|\widehat{w}(t)\| \leq \widehat{\eta}$ and $v(t+1) = \widehat{f}(v(t), u(t)) + \widehat{w}(t)$. Define $\widehat{v_0^{\mathrm{par}}} := v(1) - \widehat{X}q^c(0)$, so for all $t \geq 0$,

$$\widehat{v_0^{\mathrm{par}}} + \sum_{\tau=1}^{t} w(\tau) = v(t+1) - \widehat{X}q^c(t) \in \mathcal{V}^{\mathrm{par}}.$$

Thus, $(\widehat{f}, \widehat{w}) \in \mathbb{T}[(\widehat{X}, \widehat{\eta})]$, so $(\widehat{X}, \widehat{\eta})$ is consistent with D_T.

Conversely, suppose $(\widehat{X}, \widehat{\eta})$ is consistent with D_T, which implies the existence of $\widehat{f}(v, u) := v + \widehat{X}u$ and $\widehat{\mathbf{w}}$ satisfying $\|\widehat{\mathbf{w}}\|_\infty \leq \widehat{\eta}$ such that $v(t+1) = \widehat{f}(v(t), u(t)) + \widehat{w}(t)$. Rearranging yields $\widehat{w}(t) = v(t+1) - v(t) - \widehat{X}u(t)$, so $(\widehat{X}, \widehat{\eta})$ satisfies the norm constraint in (7.10). Now define

$$\forall t \geq 0 : \widehat{v^{\mathrm{par}}}(t) := v(t+1) - \widehat{X}q^c(t) = \widehat{v_0^{\mathrm{par}}} + \sum_{\tau=1}^{t} \widehat{w}(t)$$

so $\widehat{v^{\mathrm{par}}}(t) \in \mathcal{V}^{\mathrm{par}}$ satisfies the remaining constraint in (7.10). $\qquad\square$

Observe that each P_t is a closed, bounded, and convex set. Furthermore, P_t is non-empty, since $(X^\star, \eta^\star) \in P_t$. Intuitively, P_t is the smallest set containing all parameters that could generate the observed trajectory D_t along with a corresponding admissible sequence of noise compatible with Assumptions 13 to 15.

The consistent sets are nested $P_t \subseteq P_{t-1}$, and we use our particular choice of norm $\|\cdot\|_{\triangle, \delta}$ to establish a linear bijection between $(\mathbb{S}^n \times \mathbb{R}, \|\cdot\|_{\triangle, \delta})$ and Euclidean space $(\mathbb{R}^m, \|\cdot\|_2)$. This allows us to take advantage of any $\gamma(m)$-competitive NCBC algorithm in Euclidean space [191], [240], where m is the dimension of the space, to prove that SEL is $\gamma(m)$-competitive. This is formalized in the following lemma.

Lemma 42 (SEL is competitive)**.** *If the NCBC algorithm used in SEL has competitive ratio $\gamma(m)$, then SEL is $\gamma(m)$-competitive.*

Proof. There exists a norm-preserving linear bijection between $(\mathbb{S}^n \times \mathbb{R}, \|\cdot\|_{\triangle, \delta})$ and Euclidean space $(\mathbb{R}^m, \|\cdot\|_2)$. In particular, the mapping between the two spaces is the

vectorization of the upper-triangle of the symmetric matrix, concatenated with an additional dimension corresponding to learning $\widehat{\eta}$. Therefore, any NCBC algorithms with a $\gamma(m)$ competitive ratio will result in $\gamma(m)$ competitive ratio for SEL. □

Finally, we show that our controller Π is ρ-robust. In particular, we prove that $\Pi_{\widehat{X}}$ makes no mistakes ($M_\rho^\Pi = 0$) given consistent parameters $(\widehat{X}, \widehat{\eta}) \in P_t$.

Theorem 29 (Π is ρ-robust)**.** *Under Assumptions 13 to 15, suppose $(\widehat{X}, \widehat{\eta}) \in P_t$, where P_t is given in (7.10) for $t \geq 1$. Define $\rho = \frac{\delta\epsilon}{1+\delta\|\overline{q}-\underline{q}\|_2}$. Then, the following optimization problem is feasible:*

$$\min_{u \in \mathbb{R}^n} \quad \|\widehat{v}' - v^{nom}\|_{P_v}^2 + \|u\|_{P_u}^2 \tag{7.11a}$$

$$s.t. \quad \underline{q} \leq q^c(t-1) + u \leq \overline{q} \tag{7.11b}$$

$$\widehat{v}' = v(t) + \widehat{X}u \tag{7.11c}$$

$$k = \widehat{\eta} + \rho\left(\frac{1}{\delta} + \|u\|_2\right) \tag{7.11d}$$

$$\underline{v} + k\mathbb{1} \leq \widehat{v}' \leq \overline{v} - k\mathbb{1}. \tag{7.11e}$$

Further, the solution of (7.11), $u(t)$, guarantees voltage stability for all $(X, \eta) \in \mathcal{X} \times [0, \overline{\eta}]$ such that $\|(X, \eta) - (\widehat{X}, \widehat{\eta})\|_{\triangle, \delta} \leq \rho$. That is, $v(t) + Xu(t) + w(t) \in [\underline{v}, \overline{v}]$ for all $w(t)$ such that $\|w(t)\|_\infty \leq \eta$.

Observe that (7.11) corresponds to (7.8) in Algorithm 6 with the slack variable set to zero. We note that the robustness margin ρ decreases as $[\underline{q}, \overline{q}]$ increase. The intuitive reason is that the voltage is more sensitive to changes in $\widehat{X}$ when the range of possible u's expands. Therefore, a fixed voltage buffer of ϵ in constraints (7.8e) and (7.11d) affords less robustness to changes in $\widehat{X}$ as $[\underline{q}, \overline{q}]$ gets larger.

Proof of Theorem 29. First, we will show that the following two conditions are sufficient for feasibility of the optimization problem and ρ-robustness for the solution.

- Feasibility: $k \leq \overline{\eta} + \epsilon$
- Robustness: $k \geq \widehat{\eta} + \rho\sqrt{\frac{1}{\delta^2} + \|u\|_2^2}$

Then, we will show that our choices of k and ρ satisfy these sufficient conditions.

To derive the sufficient condition for feasibility, define

$$\widehat{v^{par}}(t-1) := v(t) - \widehat{X}q^c(t-1)$$

as the conjectured noise when we assume the underlying parameter is $\widehat{X}$. Since $\widehat{X} \in P_t$ and $P_t \subseteq P_{t-1}$, we have $\widehat{v^{\mathrm{par}}}(t-1) \in \mathcal{V}^{\mathrm{par}}$. Then, by Assumption 15, there exists $q^c \in [\underline{q}, \overline{q}]$ such that

$$\underline{v} + (\overline{\eta} + \epsilon)\mathbb{1} \le \widehat{X}q^c + \widehat{v^{\mathrm{par}}}(t-1) \le \overline{v} - (\overline{\eta} + \epsilon)\mathbb{1}.$$

Set $u = q^c - q^c(t-1)$ (which satisfies (7.11b)) and define

$$\widehat{v}'(u) := v(t) + \widehat{X}u = v(t) + \widehat{X}[q^c - q^c(t-1)]$$
$$= \widehat{X}q^c + \widehat{v^{\mathrm{par}}}(t-1).$$

Recalling (7.6), we can interpret $\widehat{v}'(u)$ as the one-step voltage prediction (without disturbance) under the model $\widehat{X}$ given control action u and the current voltage $v(t)$. We thus have

$$\underline{v} + (\overline{\eta} + \epsilon)\mathbb{1} \le \widehat{v}'(u) \le \overline{v} - (\overline{\eta} + \epsilon)\mathbb{1}.$$

Therefore, as long as $k \le \overline{\eta} + \epsilon$, u will satisfy constraint (7.11e).

Next, we derive the sufficient condition for robustness. Let u be a solution of (7.11), so it satisfies (7.11e). Let $(X, \eta) \in \mathcal{X} \times [0, \overline{\eta}]$ be arbitrary parameters satisfying $\|(X, \eta) - (\widehat{X}, \widehat{\eta})\|_{\triangle, \delta} \le \rho$. Define $\rho_X := \|X - \widehat{X}\|_{\triangle}$. By Lemma 43,

$$-\rho_X \|u\|_2 \mathbb{1} \le (X - \widehat{X})u \le \rho_X \|u\|_2 \mathbb{1}. \tag{7.12}$$

Furthermore, suppose

$$-\eta\mathbb{1} \le w(t) \le \eta\mathbb{1}. \tag{7.13}$$

Adding together the 3 inequalities (7.11e), (7.12), (7.13) yields

$$\underline{v} + (k - \rho_X \|u\|_2 - \eta)\mathbb{1} \le v(t) + Xu + w(t)$$
$$\le \overline{v} - (k - \rho_X \|u\|_2 - \eta)\mathbb{1}.$$

Clearly, if $k - \rho_X \|u\|_2 - \eta \ge 0$, then the desired robustness condition is satisfied. Since

$$\|(X, \eta) - (\widehat{X}, \widehat{\eta})\|_{\triangle, \delta}^2 = \rho_X^2 + \delta^2 |\eta - \widehat{\eta}|^2 \le \rho^2,$$

we have $|\eta - \widehat{\eta}| \le \frac{1}{\delta}\sqrt{\rho^2 - \rho_X^2}$. This implies $\eta \le \widehat{\eta} + \frac{1}{\delta}\sqrt{\rho^2 - \rho_X^2}$. Therefore, we can express the robustness condition in terms of $\widehat{\eta}$:

$$k \ge \widehat{\eta} + \frac{1}{\delta}\sqrt{\rho^2 - \rho_X^2} + \rho_X \|u\|_2 =: f(\rho_X).$$

For $\rho > 0$, $f(\rho_X)$ is strictly concave and twice-differentiable and therefore achieves its maximum when $f'(\rho_X) = 0$. This maximum value is $\widehat{\eta} + \rho\sqrt{\frac{1}{\delta^2} + \|u\|_2^2}$. Thus, if k is at least this value, then we achieve robustness.

Finally, we show that our choices of k and ρ satisfy the sufficient conditions. Since $a + b \geq \sqrt{a^2 + b^2}$ for all $a, b \geq 0$, our choice of k satisfies the robustness condition:

$$k = \widehat{\eta} + \rho\left(\frac{1}{\delta} + \|u\|_2\right) \geq \widehat{\eta} + \rho\sqrt{\frac{1}{\delta^2} + \|u\|_2^2}.$$

Note that while setting $k = \widehat{\eta} + \rho\sqrt{\frac{1}{\delta^2} + \|u\|_2^2}$ would also satisfy the robustness condition, this expression would make (7.11) a non-convex optimization problem.

The remaining step is to satisfy the feasibility condition. We must choose ρ such that $\widehat{\eta} + \rho\left(\frac{1}{\delta} + \|u\|_2\right) \leq \overline{\eta} + \epsilon$. Since $\widehat{\eta} \leq \overline{\eta}$, it suffices to find ρ such that $\rho\left(\frac{1}{\delta} + \|u\|_2\right) \leq \epsilon$. As $\|u\|_2 \leq \|\overline{q} - \underline{q}\|_2$, setting $\rho = \frac{\delta\epsilon}{1 + \delta\|\overline{q} - \underline{q}\|_2}$ satisfies the inequality. $\square$

In the case where $\eta^\star$ is known, a similar proof shows that $k = \eta^\star + \rho\|u\|_2$ and $\rho = \frac{\epsilon}{\|\overline{q} - \underline{q}\|_2}$ satisfy feasibility and robustness. (This can be seen as the $\delta \to \infty$ limiting case of Theorem 29 such that consistent model chasing only updates $\widehat{X}$ and keeps $\widehat{\eta} = \eta^\star$ fixed.)

Lemma 43. *For all $A \in \mathbb{S}^n$, $b \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}_+$,*

$$\|A\|_\triangle \leq \alpha \quad \text{implies} \quad -\alpha\|b\|_2\,\mathbb{1} \leq Ab \leq \alpha\|b\|_2\,\mathbb{1}.$$

Proof. Let A_i denote the ith row of A. By symmetry of A,

$$\|A_i\|_2^2 = \sum_{j=1}^n A_{i,j}^2 = \sum_{k=1}^{i-1} A_{k,i}^2 + \sum_{j=i}^n A_{i,j}^2$$

$$\leq \sum_{k=1}^n \sum_{j=k}^n A_{k,j}^2 = \|A\|_\triangle^2 \leq \alpha^2,$$

so $\|A_i\|_2 \leq \alpha$. Then

$$-\alpha\|b\|_2 \leq -\|A_i\|_2\|b\|_2 \leq (Ab)_i \leq \|A_i\|_2\|b\|_2 \leq \alpha\|b\|_2.$$

This holds for all $i \in \{1, \ldots, n\}$, which yields the desired result. $\square$

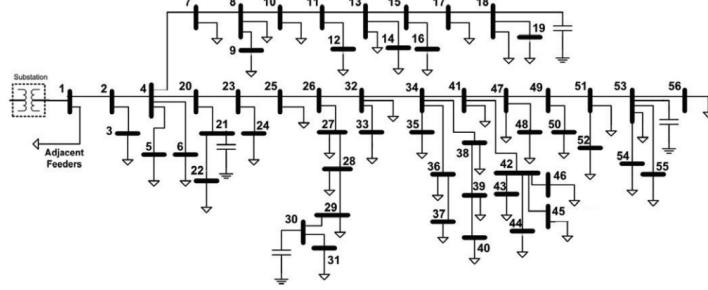Finally, combining Theorem 29 with Lemma 42 and applying Theorem 28 completes the proof of Theorem 27.

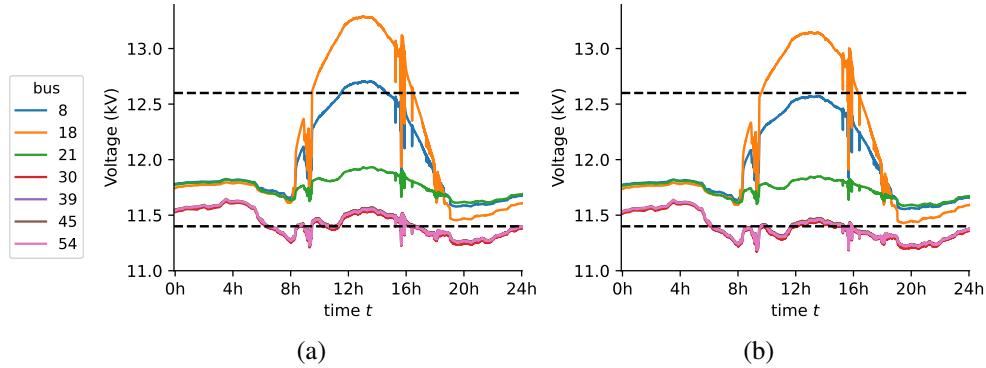Figure 7.2: Schematic diagram of SCE 56 bus distribution system.

Figure 7.3: Voltage profile of 7 buses without control, simulated with **(a)** linear dynamics (7.2) and **(b)** nonlinear balanced AC dynamics (7.1).

7.5   Case Study

We demonstrate the effectiveness of Algorithm 6 using a case study based on a single-phase 56-bus network ($n = 55$) from the Southern California Edison (SCE) utility (Figure 7.2), with line parameters r_{ij}, x_{ij} from [294, Table 1]. In our experiments, we use both the linear power model in Equation (7.5) to solve for voltages as well as the more realistic nonlinear DistFlow model (7.1). Even though our algorithm only has guarantees for the linear power flow model (7.2), we show that our algorithm works well on both the linear and nonlinear model.

Experimental Setup

Following [271], we adapt real-world load and PV data from [298] for the 56-bus network by adding power injection (scaled by the PV generation) at buses $C = \{2, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 19, 20, 23, 25, 26, 32\}$. Exogenous active and reactive power injection measurements are taken at each bus at 6-second intervals over a 24-hour period. Figure 7.3 plots these values for several buses to illustrate the setting considered. We assume that controllers with reactive power injection capacity are available at every node. The network parameters used in our experiments are:

- nominal squared voltage magnitude at the substation
  $v^0 = v^{\text{nom}} = (12\text{kV})^2$

- squared voltage magnitude limits
  $[\underline{v}, \overline{v}] = [0.95, 1.05]\text{pu} = [11.4^2, 12.6^2]\text{kV}^2$

- reactive power injection limits
  $[\underline{q}, \overline{q}] = [-0.24, 0.24]\text{MVar}$

- state and input cost matrices $P_v = 0.1I$, $P_u = 10I$

- initial state $v(1) = R^\star p(0) + X^\star q^e(0) + v^0 \mathbb{1}$, $q^c(0) = \mathbf{0}$

In comparison to previous work in the voltage control literature, our reactive power injection limits $[\underline{q}, \overline{q}]$ are slightly more generous than ± 0.2 MVar used in, *e.g.*, [271]. We choose ± 0.24 MVar because even a controller with perfect knowledge of the future would need reactive power injection capabilities of at least ± 0.238 MVar in order to maintain $v(t) \in [\underline{v}, \overline{v}]$ (if $\underline{q} = -\overline{q}$) under linear dynamics (7.2).

We set $\overline{\eta} = 10$, which upper-bounds the maximum change in exogenous noise observed in our data, which is ≈ 8.6:

$$\eta^\star = \max_t \left\| R^\star(p(t) - p(t-1)) + X^\star(q^e(t) - q^e(t-1)) \right\|_\infty.$$

We fix $\epsilon = 0.1$. In order to satisfy the requirement in Assumption 15 that $v(t) \in [\underline{v} + (\overline{\eta} + \epsilon), \overline{v} - (\overline{\eta} + \epsilon)]$, the reactive power injection capabilities must exceed ± 0.528 MVar. As we show in experiments with only ± 0.24 MVar range of control, though, Assumption 15 does not need to be fully satisfied in order for our method to still provide strong empirical results.

For the robust controller Π, we set slack variable weight $\beta = 100$ and $\mathcal{V}^{\text{par}} = [\underline{v^{\text{par}}}, \overline{v^{\text{par}}}]$ to be a rectangle around the true noise. Under linearized system dynamics, $v^{\text{par}}(t)$ is calculated as described in Section 7.2, and then we set

$$\forall i \in [n] : \quad \underline{v^{\text{par}}}_i = \min_t v_i^{\text{par}}(t), \quad \overline{v^{\text{par}}}_i = \max_t v_i^{\text{par}}(t).$$

Under nonlinear system dynamics, we approximate $v^{\text{par}}(t)$ as the nodal squared voltage magnitudes when $q^c(t) = 0$ (as shown in Figure 7.3), and we add 0.5kV^2 padding which empirically suffices as a convex outer approximation of $\mathcal{V}^{\text{par}}$:

$$\underline{v^{\text{par}}}_i = \min_t v_i^{\text{par}}(t) - 0.5, \quad \overline{v^{\text{par}}}_i = \max_t v_i^{\text{par}}(t) + 0.5.$$

As mentioned previously, we use a greedy projection-based NCBC algorithm [240] in SEL that minimizes the movement distance $\|(\widehat{X}_t, \widehat{\eta}_t) - (\widehat{X}_{t-1}, \widehat{\eta}_{t-1})\|_{\triangle, \delta}$ between

nested convex sets $P_t \subseteq P_{t-1}$:

$$\text{NCBC}_{\text{proj}}(P_t, \widehat{X}_{t-1}, \widehat{\eta}_{t-1}) := \underset{(X,\eta) \in P_t}{\arg\min} \|(X, \eta) - (\widehat{X}_{t-1}, \widehat{\eta}_{t-1})\|_{\triangle, \delta}. \qquad (7.14)$$

This achieves competitive ratio $\gamma_{\text{proj}}(m) = \pi(m-1)m^{m/2}$.

To keep the optimization problem (7.7) computationally tractable for consistent model chasing, our implementation does not use the full trajectory D as in the constraints of the consistent set (7.10). Instead, we include the 20 latest observations and 80 more observations sampled uniformly at random $(v(t), v(t+1), u(t), q^c(t)) \sim D$. This provides a computationally tractable approximation of the uncertainty set. In our experiments on linear system dynamics, we found that $\widehat{X}_t$ selected using this approximation was always in the consistent set defined by the full trajectory D, when allowing for small numerical inaccuracies introduced by the CVXPY solver.

Unless otherwise stated, we initialize $\widehat{\eta}_1 = 0$. We initialize $\widehat{X}_1$ by adding noise to the true $X^\star$ in two ways. First, we scale each line impedance x_{ij} by a random factor $\sigma_{ij} \overset{\text{iid}}{\sim} \text{Uniform}[0, 2]$. Second, we randomly permute the bus ordering, so $\widehat{X}_1$ corresponds to a permuted grid topology. Finally, we project $\widehat{X}_1$ into the uncertainty set $\mathcal{X}_\alpha$, with $\alpha = 1$.

Except for the experiments in Figure 7.8, we fix $\delta = 20$ which empirically strikes a balance between minimizing the modeling error $\|\widehat{X}_t - X^\star\|_{\triangle}$ and overfitting noise.

Table 7.1: Performance of our method simulated under linear system dynamics (top) and nonlinear system dynamics (bottom). See Section 7.5.

Info provided	# mistakes	avg. violation	max violation
Unknown	662.2 ± 435.1	0.43 ± 0.16	4.40 ± 2.59
Topo-14	917.0 ± 155.2	0.34 ± 0.12	4.93 ± 2.19
Lines-14	1085.8 ± 186.6	0.57 ± 0.29	2.55 ± 1.09
Known	88.0	0.07	0.12
Unknown	16.0 ± 15.8	0.68 ± 0.56	2.74 ± 2.39
Topo-14	0.5 ± 0.6	2.21 ± 2.56	2.90 ± 3.38
Lines-14	0.5 ± 0.6	1.01 ± 1.20	1.45 ± 1.73
Known	0.0	0.00	0.00

Experimental Results

Our experimental results demonstrate the ability of Algorithm 6 to stabilize the system without knowledge of the network topology, providing good voltage control performance even though it still has significant uncertainty about the topology at
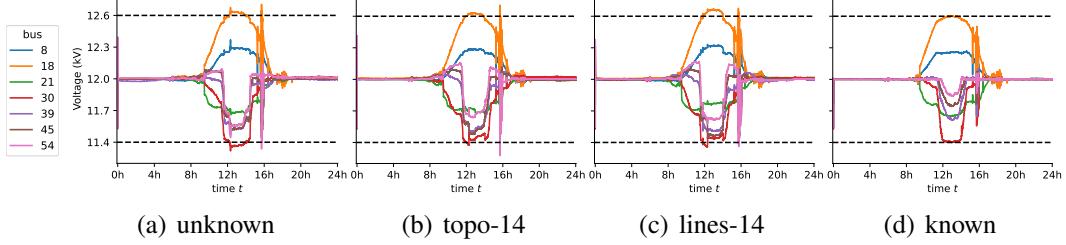
Figure 7.4: **(a)-(d)** Voltage profiles of 7 different buses simulated under linear system dynamics (7.2). Dotted black lines indicate voltage limits $[\underline{v}, \overline{v}]$. **(a)** Π+SEL initialized with random $\widehat{X} \in \mathcal{X}_\alpha$. **(b)** like (a) but the topology for buses 1-14 is known. **(c)** like (a) but the topology and line parameters for buses 1-14 are known. **(d)** like (a) but $\widehat{X} = X^\star$ is fixed and known so only $\widehat{\eta}$ is learned.
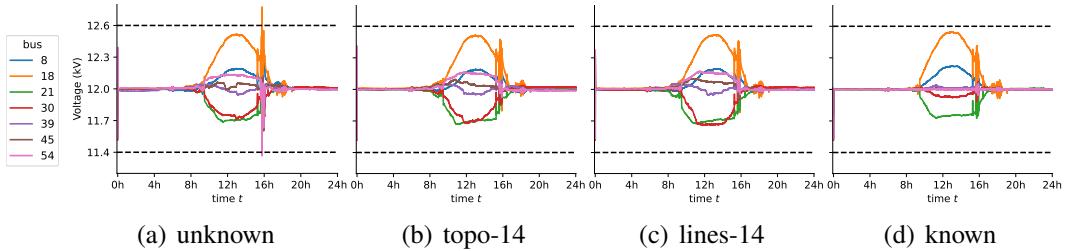
Figure 7.5: Parallels Figure 7.4. Voltage profiles of 7 different buses simulated under balanced nonlinear AC power flow (7.1).

the end of the experiments. We test our algorithm under both the linearized system dynamics (7.5) as well as the more realistic nonlinear balanced AC power flow setting (7.1) simulated using Pandapower [299]. The convex optimization problems for SEL and Π are solved with CVXPY [300] using the MOSEK solver [301]. Code for our simulations are available on GitHub.[1]

Linearized power flow with full control   Our first set of experiments, shown in Figure 7.4 and Table 7.1 (top), tests our algorithm's performance on the SCE-56 bus network under linearized system dynamics (7.5). Different amounts of network information are provided to the consistent model chasing algorithm SEL via the initial consistent set $\mathcal{X}_\alpha$, ranging from no information ("unknown," Figure 7.4(a)), information about the edges among the first 14 buses but not the line impedances ("topo-14," Figure 7.4(b)), information about the edges *and* line impedances among the first 14 buses ("lines-14," Figure 7.4(c)), and complete information about the network ("known," Figure 7.4(d)). Because the buses in the SCE 56-bus network
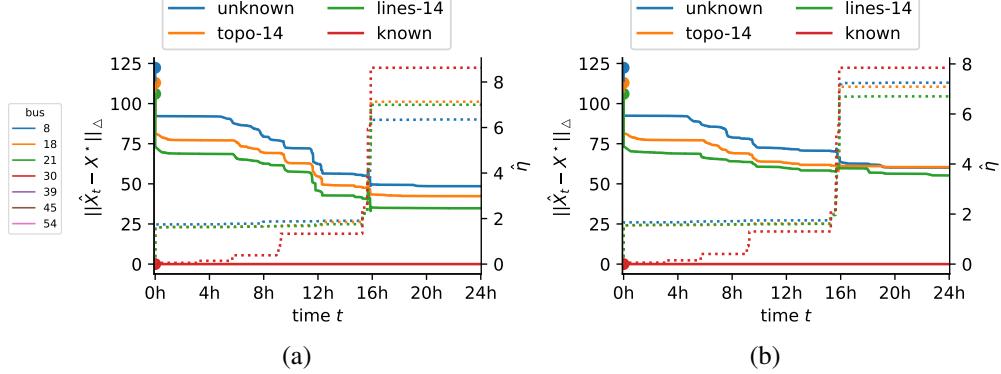
---

[1] https://github.com/chrisyeh96/voltctrl

Figure 7.6: Convergence of $\widehat{X}_t$ towards true $X^\star$ (solid lines, left axis) and estimated $\widehat{\eta}$ (dotted lines, right axis) for linear **(a)** and nonlinear **(b)** dynamics corresponding to the experiments in Figure 7.4 and Figure 7.5 respectively. Notice that even when $\|\widehat{X}_t - X^\star\|_\triangle$ does not reach 0, the controller still successfully achieve voltage safety.

are numbered in a topological ordering, the "topo-14" setting adds constraints of the form (7.9) for all of the first 14 buses, and the "lines-14" setting constrains all $X \in \mathcal{X}_\alpha$ such that $X_{ij} = X^\star_{ij}$ for all $i, j \in \{1, \ldots, 14\}$.

As shown in Figure 7.6(a), incorporating more prior knowledge about the network into the initial uncertainty set reduces the model estimation error $\|\widehat{X} - X^\star\|_\triangle$. Furthermore, the model estimation error decreases the most dramatically when the voltage violations are the largest. However, we note that lower model estimation error does not always result in fewer mistakes in our experiments.

Table 7.1 quantifies our algorithm's performance under varying amounts of initial network information. A "mistake" refers to any time step where any bus' voltage violated the limits $[\underline{v}, \overline{v}]$. "Avg. violation" refers to the average absolute squared-voltage violation

$$\underset{i \in [n],\, t \in [T]:\, v_i(t) \notin [\underline{v}_i, \overline{v}_i]}{\text{mean}} \max(v_i(t) - \overline{v}_i,\, \underline{v}_i - v_i(t)).$$

"Max violation" is like "avg. violation" but replaces the mean with a max. Results given show the mean and standard deviation over 4 random initializations of $\widehat{X}_1$.

Nonlinear power flow with full control   Our second set of experiments test our online controller on the standard balanced AC power flow model (7.1). As in the linearized power flow experiments, we compare Algorithm 6's performance across varying levels of prior information (Figure 7.5, Figure 7.6(b), and Table 7.1, bottom). Even though the controller is designed under the assumption of linearized
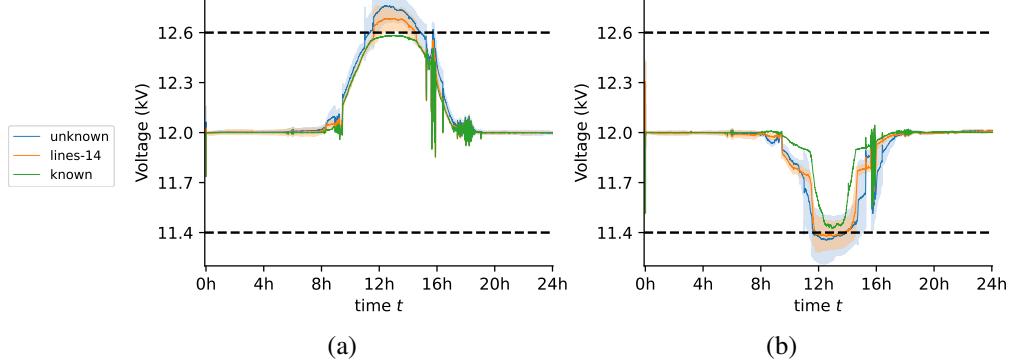
Figure 7.7: Balanced nonlinear AC power flow simulation of the voltage profiles under different algorithms with *partial control and observation*. The dark colors plot the mean voltages across 4 random initializations of $\widehat{X}_1$ and the light shading plots ±1 standard deviation. **(a)** bus 18 **(b)** bus 30.
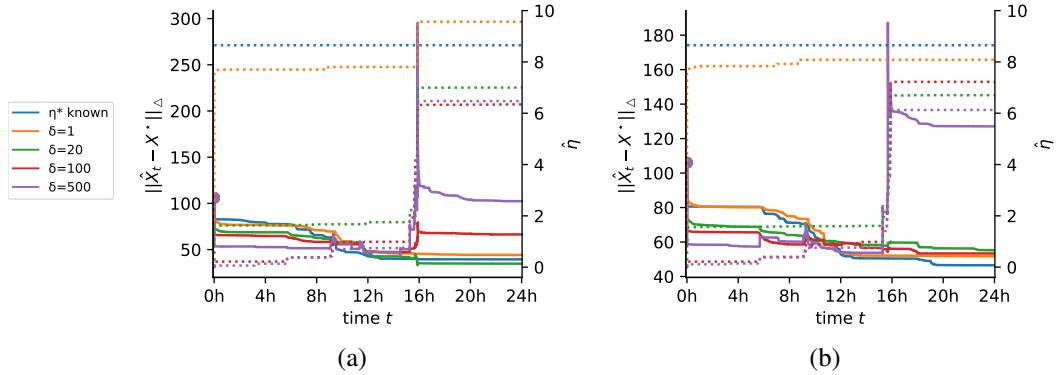

Figure 7.8: Effect of varying δ on consistent model chasing. As in Figure 7.6(a), convergence of $\widehat{X}_t$ towards $X^\star$ is plotted in solid lines (left axis), and estimated $\widehat{\eta}$ is plotted in dotted lines (right axis). In blue are results where we fix $\widehat{\eta} = \eta^\star = 8.65$ and δ has no effect. **(a)** linear dynamics **(b)** nonlinear dynamics.

voltage dynamics, our algorithm still performs well in the nonlinear simulation. The performance improves progressively, with less voltage violation and smaller overall deviation from the desired steady state voltage as it is provided more information.

Nonlinear power flow with partial observation and partial control   We also test our proposed online controller in the partial observation and partial control setting. In Figure 7.7, we withhold voltage observations and control authority from buses $i \in \{8, 18, 21, 30, 39, 45, 54\}$ by setting $q_i^c(t) = 0$ for all t. We simulate the voltage profiles across 4 random initializations of $\widehat{X}_1$ and plot the mean and ±1 standard deviation. Despite the more challenging setting, the performance of Algorithm 6 remains strong. We again observe in Figure 7.7 that adding prior topology and line

parameter information marginally improves the performance of Algorithm 6.

Varying δ   In Figure 7.8, we demonstrate the effect of varying δ on the performance of our algorithm. From a theoretical perspective, Theorem 27 shows that our algorithm achieves a finite mistake bound for every $\delta > 0$, and this bound is minimized by taking δ to be very large. What happens when using a large δ, though, is that the model chasing algorithm may overfit to noise until a time when the noise is too large, forcing the algorithm to increase the noise bound (*e.g.*, around the 16h mark in Figure 7.8). This leads to inconsistent performance in the short term, albeit with perhaps better worst-case performance. In contrast, a smaller δ allows more of the network uncertainty to be captured in a larger noise $\widehat{\eta}$ term at the cost of learning a less accurate $\widehat{X}$, but the decrease in modeling error $\|\widehat{X}_t - X^\star\|_\triangle$ becomes monotonic.

In practice, δ should be treated as a prior "confidence" about how close the initial guess of $\widehat{\eta}$ is to $\eta^\star$. δ should be larger when there is greater confidence that $\widehat{\eta}$ is close to the true $\eta^\star$.

Detecting topology changes   Finally, we consider the challenge of responding to a change in the distribution grid topology in real-time. If the topology changes from one radial grid to another due to switches, new observed data may render the consistent set empty. That is, when consistent model chasing (7.14) becomes infeasible, we are assured that the topology has changed. At this point, we may reset the algorithm by discarding the observed trajectory D_t and reinitializing consistent parameter estimates from the original consistent set P_1. Figure 7.9 demonstrates this on linear system dynamics, where we introduce a topology change at the 12h mark. We replace lines $33 \rightarrow 40$ and $46 \rightarrow 48$ with new lines $1 \rightarrow 40$ and $10 \rightarrow 48$, which maintains a radial distribution grid.

7.6   Conclusion

This chapter provides the first controller that establishes a finite-mistake guarantee for voltage control in a setting with uncertainty in both the grid topology and load and generation variations. We showed that our proposed algorithm is able to learn a model of the grid dynamics in an online fashion and provably (under linearized voltage dynamics) converge to a stable controller. Further, simulated experiments on a 56-bus distribution grid demonstrate the effectiveness of our algorithm even under more realistic nonlinear dynamics. We demonstrated how to incorporate prior knowledge
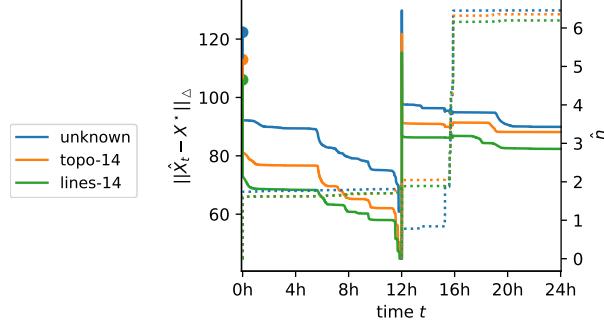
Figure 7.9: Demonstration of the detection of a topology change under linear system dynamics. Convergence of $\widehat{X}_t$ towards $X^\star$ is plotted in solid lines (left axis), where $X^\star$ changes at the 12h mark. The topology change triggers a reset of the consistent model chasing algorithm. Estimated $\widehat{\eta}$ is plotted in dotted lines (right axis).

about the network topology and line parameters to improve performance, while also extending our algorithm to the partial observability and partial controllability setting which may better reflect real-world scenarios.

As the current algorithm is centralized, future works may consider decentralized approaches to topology-robust voltage control in order to enable faster real-time control with ideas from [99]. Another direction is to extend the current algorithm to the time-varying topology setting with techniques from works such as [111]. Further studies may also explore loosening the radial topology assumption and test our algorithm on unbalanced 3-phase AC grids to accommodate a wider range of distribution grids. This would be a challenging, but important, extension. Finally, an interesting algorithmic extension is to consider computationally efficient convex body chasing algorithms with better competitive ratios. Existing methods based on Steiner point [191], [240] achieve nearly-optimal competitive ratio but are computationally inefficient in high dimension settings such as voltage control, so designing efficient approximate Steiner point algorithms could potentially lead to significant performance improvements.

BIBLIOGRAPHY

[1]   D. Silver, T. Hubert, J. Schrittwieser, *et al.*, "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play," *Science*, vol. 362, no. 6419, pp. 1140–1144, 2018.

[2]   J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1238–1274, 2013.

[3]   M. Soori, B. Arezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics, a review," *Cognitive Robotics*, vol. 3, pp. 54–70, 2023.

[4]   E. Kasneci, K. Seßler, S. Küchemann, *et al.*, "Chatgpt for good? on opportunities and challenges of large language models for education," *Learning and individual differences*, vol. 103, p. 102 274, 2023.

[5]   G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*. Springer Science & Business Media, 2013, vol. 36.

[6]   I. Khalil, J. Doyle, and K. Glover, *Robust and optimal control*. Prentice hall, 1996, vol. 2.

[7]   K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[8]   P. A. Ioannou and J. Sun, *Robust adaptive control*. Courier Corporation, 2012.

[9]   N. Hovakimyan and C. Cao, *1 adaptive control theory: Guaranteed robustness with fast adaptation*. SIAM, 2010.

[10] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*, Springer, 2007, pp. 207–226.

[11] S. Sastry and M. Bodson, *Adaptive control: stability, convergence and robustness*. Courier Corporation, 2011.

[12] E. Hazan, S. Kakade, and K. Singh, "The nonstochastic control problem," in *Algorithmic Learning Theory*, PMLR, 2020, pp. 408–421.

[13] D. Ho, H. Le, J. Doyle, and Y. Yue, "Online robust control of nonlinear systems with large uncertainty," in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR, vol. 130, San Diego, CA, USA: PMLR, Apr. 2021, pp. 3475–3483.

[14] Y. Li, J. A. Preiss, N. Li, Y. Lin, A. Wierman, and J. S. Shamma, "Online switching control with stability and regret guarantees," in *Learning for Dynamics and Control Conference*, PMLR, 2023, pp. 1138–1151.

[15] Y. Lin, J. A. Preiss, F. Xie, *et al.*, "Online policy optimization in unknown nonlinear systems," *arXiv preprint arXiv:2404.13009*, 2024.

[16] R. E. Kalman *et al.*, "Contributions to the theory of optimal control," *Bol. soc. mat. mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[17] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Reinforcement learning with fast stabilization in linear dynamical systems," in *International Conference on Artificial Intelligence and Statistics*, PMLR, 2022, pp. 5354–5390.

[18] G. Qu, Y. Shi, S. Lale, A. Anandkumar, and A. Wierman, "Stable online control of linear time-varying systems," in *Learning for Dynamics and Control*, PMLR, 2021, pp. 742–753.

[19] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, pp. 1–47, 2019.

[20] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2020.

[21] A. Cohen, A. Hasidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar, "Online linear quadratic control," in *International Conference on Machine Learning*, PMLR, 2018, pp. 1029–1038.

[22] K. Glover and J. C. Doyle, "A state space approach to h_∞ optimal control," *Three decades of mathematical system theory: a collection of surveys at the occasion of the 50th birthday of Jan C. Willems*, pp. 179–218, 2005.

[23] N. Agarwal, B. Bullins, E. Hazan, S. Kakade, and K. Singh, "Online control with adversarial disturbances," in *International Conference on Machine Learning*, PMLR, 2019, pp. 111–119.

[24] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained linear quadratic regulator," in *2019 American Control Conference (ACC)*, IEEE, 2019, pp. 5582–5588.

[25] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," in *Advances in Neural Information Processing Systems*, vol. 32, Curran Associates, Inc., 2019, pp. 10 154–10 164.

[26] X. Chen and E. Hazan, "Black-box control for linear dynamical systems," in *Conference on Learning Theory*, PMLR, 2021, pp. 1114–1143.

[27] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári, "Online least squares estimation with self-normalized processes: An application to bandit problems," *arXiv preprint arXiv:1102.2670*, 2011.

[28] T. Kargin, S. Lale, K. Azizzadenesheli, A. Anandkumar, and B. Hassibi, "Thompson sampling achieves $\widetilde{O}\sqrt{T}$ regret in linear quadratic control," in *Conference on Learning Theory*, PMLR, 2022, pp. 3235–3284.

[29] K. P. Wabersich, A. J. Taylor, J. J. Choi, *et al.*, "Data-driven safety filters: Hamilton-jacobi reachability, control barrier functions, and predictive methods for uncertain systems," *IEEE Control Systems Magazine*, vol. 43, no. 5, pp. 137–177, 2023.

[30] L. Brunke, M. Greeff, A. W. Hall, *et al.*, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 411–444, 2022.

[31] M. Petrik and R. H. Russel, "Beyond confidence regions: Tight bayesian ambiguity sets for robust mdps," *Advances in neural information processing systems*, vol. 32, 2019.

[32] T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.

[33] D. Ratasich, F. Khalid, F. Geissler, R. Grosu, M. Shafique, and E. Bartocci, "A roadmap toward the resilient internet of things for cyber-physical systems," *IEEE Access*, vol. 7, pp. 13 260–13 283, 2019.

[34] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial internet of things: A cyber-physical systems perspective," *Ieee access*, vol. 6, pp. 78 238–78 259, 2018.

[35] K. Cao, S. Hu, Y. Shi, A. W. Colombo, S. Karnouskos, and X. Li, "A survey on edge and edge-cloud computing assisted cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7806–7819, 2021.

[36] F.-J. Wu, Y.-F. Kao, and Y.-C. Tseng, "From wireless sensor networks towards cyber physical systems," *Pervasive and Mobile computing*, vol. 7, no. 4, pp. 397–413, 2011.

[37] J. M. Lemos and L. F. Pinto, "Distributed linear-quadratic control of serially chained systems: Application to a water delivery canal [applications of control]," *IEEE Control Systems Magazine*, vol. 32, no. 6, pp. 26–38, 2012.

[38] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[39] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.

[40] M. Rotkowitz, "On information structures, convexity, and linear optimality," in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 1642–1647.

[41] Y. Zheng, R. P. Mason, and A. Papachristodoulou, "Scalable design of structured controllers using chordal decomposition," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 752–767, 2017.

[42] Y. R. Sturz, A. Eichler, and R. S. Smith, "Distributed control design for heterogeneous interconnected systems," *IEEE Transactions on Automatic Control*, 2020.

[43] N. Matni and V. Chandrasekaran, "Regularization for design," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3991–4006, 2016.

[44] Y.-S. Wang, N. Matni, and J. C. Doyle, "Localized lqr optimal control," in *53rd IEEE Conference on Decision and Control*, IEEE, 2014, pp. 1661–1668.

[45] Y.-S. Wang, N. Matni, and J. C. Doyle, "A system-level approach to controller synthesis," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4079–4093, 2019.

[46] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour, "An input–output parametrization of stabilizing controllers: Amidst youla and system level synthesis," *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 1014–1019, 2019.

[47] Ş. Sabău, A. Sperilă, C. Oară, and A. Jadbabaie, "Network realization functions for optimal distributed control," *arXiv preprint arXiv:2112.09093*, 2021.

[48] E. Fogel and Y.-F. Huang, "On the value of information in system identification–bounded noise case," *Automatica*, vol. 18, no. 2, pp. 229–238, 1982.

[49] M. Lorenzen, M. Cannon, and F. Allgöwer, "Robust mpc with recursive model update," *Automatica*, vol. 103, pp. 461–471, 2019.

[50] M. Bujarbaruah, X. Zhang, M. Tanaskovic, and F. Borrelli, "Adaptive stochastic mpc under time-varying uncertainty," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2840–2845, 2020.

[51] K. Zhang, Q. Sun, and Y. Shi, "Trajectory tracking control of autonomous ground vehicles using adaptive learning mpc," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5554–5564, 2021.

[52] A. Parsi, A. Iannelli, M. Yin, M. Khosravi, and R. S. Smith, "Robust adaptive model predictive control with worst-case cost," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 4222–4227, 2020.

[53] A. Parsi, A. Iannelli, and R. S. Smith, "Active exploration in adaptive model predictive control," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 6186–6191.

[54] A. Sasfi, M. N. Zeilinger, and J. Köhler, "Robust adaptive mpc using control contraction metrics," *arXiv preprint arXiv:2209.11713*, 2022.

[55] C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Robust online voltage control with an unknown grid topology," in *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, 2022, pp. 240–250.

[56] D. J. Russo, B. Van Roy, A. Kazerouni, I. Osband, Z. Wen, *et al.*, "A tutorial on thompson sampling," *Foundations and Trends® in Machine Learning*, vol. 11, no. 1, pp. 1–96, 2018.

[57] M. Abeille and A. Lazaric, "Improved regret bounds for thompson sampling in linear quadratic control problems," in *International Conference on Machine Learning*, PMLR, 2018, pp. 1–9.

[58] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "Regret bounds for robust adaptive control of the linear quadratic regulator," *Advances in Neural Information Processing Systems*, vol. 31, pp. 4188–4197, 2018.

[59] M. Simchowitz and D. Foster, "Naive exploration is optimal for online lqr," in *International Conference on Machine Learning*, PMLR, 2020, pp. 8937–8948.

[60] Y. Abbasi-Yadkori and C. Szepesvári, "Regret bounds for the adaptive control of linear quadratic systems," in *Proceedings of the 24th Annual Conference on Learning Theory*, JMLR Workshop and Conference Proceedings, 2011, pp. 1–26.

[61] J. Umenberger and T. B. Schön, "Optimistic robust linear quadratic dual control," in *Learning for Dynamics and Control*, PMLR, 2020, pp. 550–560.

[62] I. Szita and A. Lőrincz, "The many faces of optimism: A unifying approach," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 1048–1055.

[63] "Global Survey of Regulatory Approaches for Power Quality and Reliability," EPRI, Palo Alto, CA, USA, Tech. Rep. 1008589, Mar. 2005, p. 146. [Online]. Available: https://www.epri.com/research/products/1008589.

[64] H. Haes Alhelou, M. E. Hamedani-Golshan, T. C. Njenda, and P. Siano, "A survey on power system blackout and cascading events: Research motivations and challenges," *Energies*, vol. 12, no. 4, 2019. DOI: 10.3390/en12040682.

[65] Y. Liao, Y. Weng, M. Wu, and R. Rajagopal, "Distribution grid topology reconstruction: An information theoretic approach," in *2015 North American Power Symposium (NAPS)*, Charlotte, NC, USA: IEEE, Oct. 2015, pp. 1–6, ISBN: 978-1-4673-7389-0. DOI: 10.1109/NAPS.2015.7335248. [Online]. Available: https://ieeexplore.ieee.org/document/7335248/.

[66] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, Sep. 2018. DOI: 10.1109/TCNS.2017.2673546.

[67] L. Lessard and S. Lall, "Optimal controller synthesis for the decentralized two-player problem with output feedback," in *2012 American Control Conference (ACC)*, IEEE, 2012, pp. 6314–6321.

[68] H. R. Feyzmahdavian, A. Alam, and A. Gattami, "Optimal distributed controller design with communication delays: Application to vehicle formations," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, IEEE, 2012, pp. 2232–2237.

[69] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid–the new and improved power grid: A survey," *IEEE communications surveys & tutorials*, vol. 14, no. 4, pp. 944–980, 2011.

[70] S. E. Li, Y. Zheng, K. Li, and J. Wang, "An overview of vehicular platoon control under the four-component framework," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2015, pp. 286–291.

[71] J. Han and R. E. Skelton, "An lmi optimization approach for structured linear controllers," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, IEEE, vol. 5, 2003, pp. 5143–5148.

[72] J. Skaf and S. P. Boyd, "Design of affine controllers via convex optimization," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2476–2487, 2010.

[73] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.

[74] E. L. Jensen, "Topics in optimal distributed control," Ph.D. dissertation, UC Santa Barbara, 2020.

[75] M. W. Fisher, G. Hug, and F. Dörfler, "System level synthesis beyond finite impulse response using approximation by simple poles," *arXiv preprint arXiv:2203.16765*, 2022.

[76] Y. Zheng, L. Furieri, M. Kamgarpour, and N. Li, "System-level, input–output and new parameterizations of stabilizing controllers, and their numerical computation," *Automatica*, vol. 140, p. 110 211, 2022.

[77] C. A. Alonso and N. Matni, "Distributed and localized closed loop model predictive control via system level synthesis," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 5598–5605.

[78] Y.-S. Wang, N. Matni, and J. C. Doyle, "Separable and localized system-level synthesis for large-scale systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4234–4249, 2018.

[79] S.-H. Tseng and J. Anderson, "Deployment architectures for cyber-physical control systems," in *2020 American Control Conference (ACC)*, IEEE, 2020, pp. 5287–5294.

[80] S.-H. Tseng and J. Anderson, "Deployment architectures for cyber-physical control systems," *arXiv preprint arXiv:1911.01510*, pp. 5287–5294, 2019.

[81] Y.-S. Wang, S. You, and N. Matni, "Localized distributed Kalman filters for large-scale systems," *IFAC-PapersOnLine*, vol. 48, no. 22, pp. 52–57, 2015.

[82] C. A. Alonso, J. S. Li, N. Matni, and J. Anderson, "Distributed and localized model predictive control–part ii: Theoretical guarantees," *IEEE Transactions on Control of Network Systems*, vol. 10, no. 3, pp. 1113–1123, 2023.

[83] J. M. Maciejowski, *Predictive control: with constraints*. Pearson education, 2002.

[84] L. Zhang, J. Wang, and C. Li, "Distributed model predictive control for polytopic uncertain systems subject to actuator saturation," *Journal of Process Control*, vol. 23, no. 8, pp. 1075–1089, 2013.

[85] Z. Li and J. Sun, "Disturbance compensating model predictive control with application to ship heading control," *IEEE transactions on control systems technology*, vol. 20, no. 1, pp. 257–265, 2011.

[86] Y. Chen and J. Anderson, "System level synthesis with state and input constraints," *CoRR*, vol. abs/1903.07174, 2019. arXiv: 1903.07174. [Online]. Available: http://arxiv.org/abs/1903.07174.

[87] S. Galeani, S. Tarbouriech, M. Turner, and L. Zaccarian, "A tutorial on modern anti-windup design," *European Journal of Control*, vol. 15, no. 3-4, pp. 418–440, 2009.

[88] I. Kolmanovsky, E. Garone, and S. Di Cairano, "Reference and command governors: A tutorial on their theory and automotive applications," in *2014 American Control Conference*, IEEE, 2014, pp. 226–241.

[89] T. Hu, Z. Lin, and B. M. Chen, "Analysis and design for discrete-time linear systems subject to actuator saturation," *Systems & Control Letters*, vol. 45, no. 2, pp. 97–112, 2002.

[90] M. V. Kothare, P. J. Campo, M. Morari, and C. N. Nett, "A unified framework for the study of anti-windup designs," *Automatica*, vol. 30, no. 12, pp. 1869–1883, 1994.

[91] T. Hu, A. R. Teel, and L. Zaccarian, "Anti-windup synthesis for linear control systems with input saturation: Achieving regional, nonlinear performance," *Automatica*, vol. 44, no. 2, pp. 512–519, 2008.

[92] D. Ho and J. Doyle, "System level approach to general nonlinear systems:the universal principle of realizing closed loop maps," *ACC*, 2020.

[93] A. Zheng, M. V. Kothare, and M. Morari, "Anti-windup design for internal model control," *International Journal of Control*, vol. 60, no. 5, pp. 1015–1024, 1994.

[94] A. Sarker, P. Fisher, J. E. Gaudio, and A. M. Annaswamy, "Accurate parameter estimation for safety-critical systems with unmodeled dynamics," *Artificial Intelligence*, p. 103 857, 2023.

[95] A. Wagenmaker and K. Jamieson, "Active learning for identification of linear dynamical systems," in *Conference on Learning Theory*, PMLR, 2020, pp. 3487–3582.

[96] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Conference On Learning Theory*, PMLR, 2018, pp. 439–473.

[97] T.-J. Chang and S. Shahrampour, "Regret analysis of distributed online lqr control for unknown LTI systems," *IEEE Transactions on Automatic Control*, vol. 69, no. 1, pp. 667–673, 2024.

[98] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-based model predictive control: Toward safe learning in control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 269–296, 2020.

[99] J. Yu, D. Ho, and A. Wierman, "Online adversarial stabilization of unknown networked systems," in *Prof. ACM Meas. Anal. Comput. Syst.*, vol. 7, 2023, pp. 1–43.

[100] M. Lauricella and L. Fagiano, "Set membership identification of linear systems with guaranteed simulation accuracy," *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5189–5204, 2020.

[101] M. M. Livstone and M. A. Dahleh, "Asymptotic properties of set membership identification algorithms," *Systems & control letters*, vol. 27, no. 3, pp. 145–155, 1996.

[102] D. P. Bertsekas, "Control of uncertain systems with a set-membership description of the uncertainty.," Ph.D. dissertation, Massachusetts Institute of Technology, 1971.

[103] J. Qi, G. He, S. Mei, and F. Liu, "Power system set membership state estimation," in *2012 IEEE Power and Energy Society General Meeting*, IEEE, 2012, pp. 1–7.

[104] J. R. Benevides, M. A. Paiva, P. V. Simplício, R. S. Inoue, and M. H. Terra, "Disturbance observer-based robust control of a quadrotor subject to parametric uncertainties and wind disturbance," *IEEE Access*, vol. 10, pp. 7554–7565, 2022.

[105] K. Narendra and A. Annaswamy, "Robust adaptive control in the presence of bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 31, no. 4, pp. 306–315, 1986.

[106] X. Zhang, W. Shi, X. Li, B. Yan, A. Malkawi, and N. Li, "Decentralized temperature control via HVAC systems in energy efficient buildings: An approximate solution procedure," in *Proceedings of 2016 IEEE Global Conference on Signal and Information Processing*, 2016, pp. 936–940.

[107] X. Lu and M. Cannon, "Robust adaptive model predictive control with persistent excitation conditions," *Automatica*, vol. 152, p. 110 959, 2023.

[108] Y. Li, S. Das, and N. Li, "Online optimal control with affine constraints," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, May 2021, pp. 8527–8537. DOI: 10.1609/aaai.v35i10.17035. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/17035.

[109] Y. Li, S. Das, J. Shamma, and N. Li, "Safe adaptive learning-based control for constrained linear quadratic regulators with regret guarantees," *arXiv preprint arXiv:2111.00411*, 2021.

[110] X. Liu, Z. Yang, and L. Ying, "Online nonstochastic control with adversarial and static constraints," *arXiv preprint arXiv:2302.02426*, 2023.

[111] J. Yu, V. Gupta, and A. Wierman, "Online adversarial stabilization of unknown linear time-varying systems," in *Prof. 62nd IEEE Conf. Decis. Control (CDC)*, 2023, pp. 8320–8327.

[112] E.-W. Bai, H. Cho, and R. Tempo, "Convergence properties of the membership set," *Automatica*, vol. 34, no. 10, pp. 1245–1249, 1998.

[113] H. Akçay, "The size of the membership-set in a probabilistic framework," *Automatica*, vol. 40, no. 2, pp. 253–260, 2004.

[114] W. Kitamura, Y. Fujisaki, and E.-W. Bai, "The size of the membership set in the presence of disturbance and parameter uncertainty," in *Proceedings of the 44th IEEE Conference on Decision and Control*, IEEE, 2005, pp. 5698–5703.

[115] E.-W. Bai, R. Tempo, and H. Cho, "Membership set estimators: Size, optimal inputs, complexity and relations with least squares," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 42, no. 5, pp. 266–277, 1995.

[116] J. Eising, S. Liu, S. Martínez, and J. Cortés, "Using data informativity for online stabilization of unknown switched linear systems," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, IEEE, 2022, pp. 8–13.

[117] J. Köhler, E. Andina, R. Soloperto, M. A. Müller, and F. Allgöwer, "Linear robust adaptive model predictive control: Computational complexity and conservatism," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, IEEE, 2019, pp. 1383–1388.

[118] X. Lu, M. Cannon, and D. Koksal-Rivet, "Robust adaptive model predictive control: Performance and parameter estimation," *International Journal of Robust and Nonlinear Control*, 2019.

[119] C. Wu, K. L. Teo, and S. Wu, "Min–max optimal control of linear systems with uncertainty and terminal state constraints," *Automatica*, vol. 49, no. 6, pp. 1809–1815, 2013.

[120] S. L. Tu, *Sample complexity bounds for the linear quadratic regulator*. University of California, Berkeley, 2019.

[121] Y. Li, Y. Tang, R. Zhang, and N. Li, "Distributed reinforcement learning for decentralized linear quadratic control: A derivative-free policy optimization approach," *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6429–6444, 2021.

[122] K. S. Narendra and A. M. Annaswamy, "Persistent excitation in adaptive systems," *International Journal of Control*, vol. 45, no. 1, pp. 127–160, 1987.

[123] B. T. Lopez, J.-J. E. Slotine, and J. P. How, "Robust adaptive control barrier functions: An adaptive and data-driven approach to safety," *IEEE Control Systems Letters*, vol. 5, no. 3, pp. 1031–1036, 2020.

[124] J. B. Rawlings and D. Q. Mayne, *Model predictive control: Theory and design*. Nob Hill Pub., 2009.

[125] Y. Li, T. Zhang, S. Das, J. Shamma, and N. Li, "Non-asymptotic system identification for linear systems with nonlinear policies," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 1672–1679, 2023, 22nd IFAC World Congress, ISSN: 2405-8963.

[126] J. P. Hespanha, *Linear systems theory*. Princeton university press, 2018.

[127] H. Xu and Y. Li, "On the convergence rates of set membership estimation of linear systems with disturbances bounded by general convex sets," *arXiv preprint arXiv:2406.00574*, 2024.

[128] R. Zhang, Y. Li, and N. Li, "On the regret analysis of online lqr control with predictions," in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 697–703.

[129] Y. Li, X. Chen, and N. Li, "Online optimal control with linear dynamics and predictions: Algorithms and regret analysis," in *Advances in Neural Information Processing Systems*, 2019, pp. 14 887–14 899.

[130] A. Mete, R. Singh, and P. Kumar, "Augmented rbmle-ucb approach for adaptive control of linear quadratic systems," *Advances in Neural Information Processing Systems*, vol. 35, pp. 9302–9314, 2022.

[131] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design*. Nob Hill Publishing Madison, WI, 2017, vol. 2.

[132] D. Q. Mayne, M. M. Seron, and S. Rakovic, "Robust model predictive control of constrained linear systems with bounded disturbances," *Automatica*, vol. 41, no. 2, pp. 219–224, 2005.

[133] Y. Sattar, S. Oymak, and N. Ozay, "Finite sample identification of bilinear dynamical systems," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, IEEE, 2022, pp. 6705–6711.

[134] D. Foster, T. Sarkar, and A. Rakhlin, "Learning nonlinear dynamical systems from a single trajectory," in *Learning for Dynamics and Control*, PMLR, 2020, pp. 851–861.

[135] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," in *International Conference on Machine Learning*, PMLR, 2019, pp. 5610–5618.

[136] M. Milanese, J. Norton, H. Piet-Lahanier, and É. Walter, *Bounding approaches to system identification*. Springer Science & Business Media, 2013.

[137] M. Abeille and A. Lazaric, "Thompson sampling for linear-quadratic control problems," in *Artificial intelligence and statistics*, PMLR, 2017, pp. 1246–1254.

[138] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2737–2752, 2018.

[139] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory–Implementation–Applications*. Springer Science & Business Media, 2012.

[140] L. Ljung, *System identification*. Springer, 1998.

[141] M. C. Campi and E. Weyer, "Finite sample properties of system identification methods," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1329–1334, 2002.

[142] M. Vidyasagar and R. L. Karandikar, "A learning theory approach to system identification and stochastic adaptive control," *Probabilistic and randomized methods for design under uncertainty*, pp. 265–302, 2006.

[143] D. Hsu, S. M. Kakade, and T. Zhang, "Random design analysis of ridge regression," in *Conference on learning theory*, JMLR Workshop and Conference Proceedings, 2012, pp. 9–1.

[144] S. Oymak and N. Ozay, "Non-asymptotic identification of lti systems from a single trajectory," in *2019 American control conference (ACC)*, IEEE, 2019, pp. 5655–5661.

[145] Y. Zheng and N. Li, "Non-asymptotic identification of linear dynamical systems using multiple trajectories," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1693–1698, 2020.

[146] A. Rantzer, "Concentration bounds for single parameter adaptive control," in *2018 Annual American Control Conference (ACC)*, IEEE, 2018, pp. 1862–1866.

[147] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, "Finite-time adaptive stabilization of linear systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3498–3505, 2018.

[148] A. Tsiamis, I. Ziemann, N. Matni, and G. J. Pappas, "Statistical learning theory for control: A finite sample perspective," *arXiv preprint arXiv:2209.05423*, 2022.

[149] Z. Zhao and Q. Li, "Adaptive sampling methods for learning dynamical systems," in *Mathematical and Scientific Machine Learning*, PMLR, 2022, pp. 335–350.

[150] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.

[151] V. Adetola and M. Guay, "Robust adaptive mpc for constrained uncertain nonlinear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 2, pp. 155–167, 2011.

[152] M. Tanaskovic, L. Fagiano, R. Smith, P. Goulart, and M. Morari, "Adaptive model predictive control for constrained linear systems," in *2013 European Control Conference (ECC)*, IEEE, 2013, pp. 382–387.

[153] W. Kitamura and Y. Fujisaki, "Convergence properties of the membership set in the presence of disturbance and parameter uncertainty," *Transactions of the Society of Instrument and Control Engineers*, vol. 39, no. 4, pp. 382–387, 2003.

[154] H. J. van Waarde, J. Eising, M. K. Camlibel, and H. L. Trentelman, "The informativity approach to data-driven analysis and control," *arXiv preprint arXiv:2302.10488*, 2023.

[155] J. Eising and J. Cortes, "When sampling works in data-driven control: Informativity for stabilization in continuous time," *arXiv preprint arXiv:2301.10873*, 2023.

[156] Y. Liu, Y. Zhao, and F. Wu, "Ellipsoidal state-bounding-based set-membership estimation for linear system with unknown-but-bounded disturbances," *IET Control Theory & Applications*, vol. 10, no. 4, pp. 431–442, 2016.

[157] Y. Gao, Y. Tang, H. Qi, and H. Yang, "Closure: Fast quantification of pose uncertainty sets," *arXiv preprint arXiv:2403.09990*, 2024.

[158] Y. Tang, J.-B. Lasserre, and H. Yang, "Uncertainty quantification of set-membership estimation in control and perception: Revisiting the minimum enclosing ellipsoid," *arXiv preprint arXiv:2311.15962*, 2023.

[159]   C. Rogers, "Covering a sphere with spheres," *Mathematika*, vol. 10, no. 2, pp. 157–164, 1963.

[160]   J.-L. Verger-Gaugry, "Covering a ball with smaller equal balls in $\mathbb{R}^n$," *Discrete and Computational Geometry*, vol. 33, pp. 143–155, 2005.

[161]   S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[162]   P. Gradu, E. Hazan, and E. Minasyan, "Adaptive regret for control of time-varying dynamics," *arXiv preprint arXiv:2007.04393*, 2020.

[163]   B. Nortmann and T. Mylvaganam, "Data-driven control of linear time-varying systems," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 3939–3944.

[164]   Y. Luo, V. Gupta, and M. Kolar, "Dynamic regret minimization for control of non-stationary linear dynamical systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 6, no. 1, pp. 1–72, 2022.

[165]   Y. Lin, J. Preiss, E. Anand, Y. Li, Y. Yue, and A. Wierman, "Online adaptive controller selection in time-varying systems: No-regret via contractive perturbations," *arXiv preprint arXiv:2210.12320*, 2022.

[166]   E. Minasyan, P. Gradu, M. Simchowitz, and E. Hazan, "Online control of unknown time-varying dynamical systems," *Advances in Neural Information Processing Systems*, vol. 34, pp. 15 934–15 945, 2021.

[167]   R. Tedrake, "Underactuated robotics: Learning, planning, and control for efficient and agile machines course notes for mit 6.832," *Working draft edition*, vol. 3, p. 4, 2009.

[168]   P. Falcone, F. Borrelli, H. E. Tseng, J. Asgari, and D. Hrovat, "Linear time-varying model predictive control and its application to active steering systems: Stability analysis and experimental validation," *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 18, no. 8, pp. 862–875, 2008.

[169]   K. S. Tsakalis and P. A. Ioannou, *Linear time-varying systems: control and adaptation*. Prentice-Hall, Inc., 1993.

[170]   R. Tóth, *Modeling and identification of linear parameter-varying systems*. Springer, 2010, vol. 403.

[171]   J. Mohammadpour and C. W. Scherer, *Control of linear parameter varying systems with applications*. Springer Science & Business Media, 2012.

[172]   W. Zhang, Q.-L. Han, Y. Tang, and Y. Liu, "Sampled-data control for a class of linear time-varying systems," *Automatica*, vol. 103, pp. 126–134, 2019.

[173] R. Mojgani and M. Balajewicz, "Stabilization of linear time-varying reduced-order models: A feedback controller approach," *International Journal for Numerical Methods in Engineering*, vol. 121, no. 24, pp. 5490–5510, 2020.

[174] G. Goel and B. Hassibi, "Regret-optimal estimation and control," *IEEE Transactions on Automatic Control*, 2023.

[175] K. Tsakalis and P. Ioannou, "Adaptive control of linear time-varying plants," *Automatica*, vol. 23, no. 4, pp. 459–468, 1987.

[176] J.-J. Slotine and J. Coetsee, "Adaptive sliding controller synthesis for nonlinear systems," *International Journal of Control*, vol. 43, no. 6, pp. 1631–1651, 1986.

[177] R. Marino and P. Tomei, "Adaptive control of linear time-varying systems," *Automatica*, vol. 39, no. 4, pp. 651–659, 2003.

[178] M. Verhaegen and X. Yu, "A class of subspace model identification algorithms to identify periodically and arbitrarily time-varying systems," *Automatica*, vol. 31, no. 2, pp. 201–216, 1995.

[179] B. Bamieh and L. Giarre, "Identification of linear parameter varying models," *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 12, no. 9, pp. 841–853, 2002.

[180] T. Sarkar, A. Rakhlin, and M. Dahleh, "Nonparametric system identification of stochastic switched linear systems," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, IEEE, 2019, pp. 3623–3628.

[181] S. Talebi, S. Alemzadeh, N. Rahimi, and M. Mesbahi, "On regularizability and its application to online control of unstable lti systems," *IEEE Transactions on Automatic Control*, 2021.

[182] M. Rotulo, C. De Persis, and P. Tesi, "Online learning of data-driven controllers for unknown switched linear systems," *Automatica*, vol. 145, p. 110 519, 2022.

[183] S. Baros, C.-Y. Chang, G. E. Colon-Reyes, and A. Bernstein, "Online data-enabled predictive control," *Automatica*, vol. 138, p. 109 926, 2022.

[184] B. Pang, T. Bian, and Z.-P. Jiang, "Data-driven finite-horizon optimal control for linear time-varying discrete-time systems," in *2018 IEEE Conference on Decision and Control (CDC)*, IEEE, 2018, pp. 861–866.

[185] S.-J. Liu, M. Krstic, and T. Başar, "Batch-to-batch finite-horizon lq control for unknown discrete-time linear systems via stochastic extremum seeking," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4116–4123, 2016.

[186] Y. Ouyang, M. Gagrani, and R. Jain, "Learning-based control of unknown linear systems with thompson sampling," *arXiv preprint arXiv:1709.04047*, 2017.

[187] Y. Han, R. Solozabal, J. Dong, X. Zhou, M. Takac, and B. Gu, "Learning to control under time-varying environment," *arXiv preprint arXiv:2206.02507*, 2022.

[188] M. Sellke, "Chasing convex bodies optimally," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 2020, pp. 1509–1518.

[189] C. Argue, A. Gupta, Z. Tang, and G. Guruganesh, "Chasing convex bodies with linear competitive ratio," *Journal of the ACM (JACM)*, vol. 68, no. 5, pp. 1–10, 2021.

[190] N. Bansa, M. Böhm, M. Eliáš, G. Koumoutsos, and S. W. Umboh, "Nested convex bodies are chaseable," in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 2018, pp. 1253–1260.

[191] S. Bubeck, B. Klartag, Y. T. Lee, Y. Li, and M. Sellke, "Chasing nested convex bodies nearly optimally," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, Salt Lake City, Utah: SIAM, 2020, pp. 1496–1508, ISBN: 9781611975994. DOI: 10.1137/1.9781611975994.91.

[192] B. Ramasubramanian, B. Xiao, L. Bushnell, and R. Poovendran, "Safety-critical online control with adversarial disturbances," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 3731–3738.

[193] A. Cohen, T. Koren, and Y. Mansour, "Learning linear-quadratic regulators efficiently with only sqrt(t) regret," in *International Conference on Machine Learning*, PMLR, 2019, pp. 1300–1309.

[194] W. Hahn *et al.*, *Stability of motion*. Springer, 1967, vol. 138.

[195] F. Amato, G. Celentano, and F. Garofalo, "New sufficient conditions for the stability of slowly varying linear systems," *IEEE Transactions on Automatic Control*, vol. 38, no. 9, pp. 1409–1411, 1993.

[196] J. Xiong and J. Lam, "Stabilization of discrete-time markovian jump linear systems via time-delayed controllers," *Automatica*, vol. 42, no. 5, pp. 747–753, 2006.

[197] J. Jiang and Y. Zhang, "A revisit to block and recursive least squares for parameter estimation," *Computers & Electrical Engineering*, vol. 30, no. 5, pp. 403–416, 2004.

[198] U. Shaked, "Guaranteed stability margins for the discrete-time linear quadratic optimal regulator," *IEEE Transactions on Automatic Control*, vol. 31, no. 2, pp. 162–165, 1986.

[199] A. S. Householder, *The theory of matrices in numerical analysis*. Courier Corporation, 2013.

[200] P. Gahinet and A. Laub, "Computable bounds for the sensitivity of the algebraic riccati equation," *SIAM journal on control and optimization*, vol. 28, no. 6, pp. 1461–1480, 1990.

[201] D. Morgan, S.-J. Chung, and F. Y. Hadaegh, "Model predictive control of swarms of spacecraft using sequential convex programming," *Journal of Guidance, Control, and Dynamics*, vol. 37, no. 6, pp. 1725–1740, 2014.

[202] J. Bu, A. Mesbahi, M. Fazel, and M. Mesbahi, "Lqr through the lens of first order methods: Discrete-time case," *arXiv preprint arXiv:1907.08921*, 2019.

[203] S. Fattahi, N. Matni, and S. Sojoudi, "Efficient learning of distributed linear-quadratic control policies," *SIAM Journal on Control and Optimization*, vol. 58, no. 5, pp. 2927–2951, 2020.

[204] L. Furieri, Y. Zheng, and M. Kamgarpour, "Learning the globally optimal distributed lq regulator," in *Learning for Dynamics and Control*, PMLR, 2020, pp. 287–297.

[205] L. Ye, H. Zhu, and V. Gupta, "On the sample complexity of decentralized linear quadratic regulator with partially nested information structure," *arXiv preprint arXiv:2110.07112*, 2021.

[206] M. Fardad and M. R. Jovanović, "On the design of optimal structured and sparse feedback gains via sequential convex programming," in *2014 American Control Conference*, IEEE, 2014, pp. 2426–2431.

[207] Y. Zheng, L. Furieri, A. Papachristodoulou, N. Li, and M. Kamgarpour, "On the equivalence of youla, system-level, and input–output parameterizations," *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 413–420, 2020.

[208] Y. Hu, A. Wierman, and G. Qu, "On the sample complexity of stabilizing lti systems on a single trajectory," *arXiv preprint arXiv:2202.07187*, 2022.

[209] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[210] C. A. Alonso, F. Yang, and N. Matni, "Data-driven distributed and localized model predictive control," *arXiv preprint arXiv:2112.12229*, 2021.

[211] J. Sieber, S. Bennani, and M. N. Zeilinger, "A system level approach to tube-based model predictive control," *IEEE Control Systems Letters*, vol. 6, pp. 776–781, 2021.

[212] P. Ioannou and B. Fidan, *Adaptive control tutorial*. SIAM, 2006.

[213] J. Perdomo, J. Umenberger, and M. Simchowitz, "Stabilizing dynamical systems via policy gradient methods," *Advances in Neural Information Processing Systems*, vol. 34, 2021.

[214] F. Zhao, X. Fu, and K. You, "Learning stabilizing controllers of linear systems via discount policy gradient," *arXiv preprint arXiv:2112.09294*, 2021.

[215] L. Treven, S. Curi, M. Mutnỳ, and A. Krause, "Learning stabilizing controllers for unstable linear quadratic regulators from a single trajectory," in *Learning for Dynamics and Control*, PMLR, 2021, pp. 664–676.

[216] A. Lamperski, "Computing stabilizing linear controllers via policy iteration," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 1902–1907.

[217] M. Kashyap and L. Lessard, "Explicit agent-level optimal cooperative controllers for dynamically decoupled systems with output feedback," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, IEEE, 2019, pp. 8254–8259.

[218] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440–446, 1985.

[219] A. Lamperski and L. Lessard, "Optimal decentralized state-feedback control with sparsity and delays," *Automatica*, vol. 58, pp. 143–151, 2015.

[220] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, vol. 20, no. 4, pp. 633–679, 2020.

[221] A. Didier, J. Sieber, and M. N. Zeilinger, "A system level approach to regret optimal control," *IEEE Control Systems Letters*, 2022.

[222] M. K. S. Faradonbeh and A. Modi, "Joint learning-based stabilization of multiple unknown linear systems," *arXiv preprint arXiv:2201.01387*, 2022.

[223] G. Jing, H. Bai, J. George, A. Chakrabortty, and P. K. Sharma, "Learning distributed stabilizing controllers for multi-agent systems," *IEEE Control Systems Letters*, 2021.

[224] S. Alemzadeh and M. Mesbahi, "Distributed q-learning for dynamically decoupled systems," in *2019 American Control Conference (ACC)*, IEEE, 2019, pp. 772–777.

[225] S. Talebi, S. Alemzadeh, and M. Mesbahi, "Distributed model-free policy iteration for networks of homogeneous systems," in *2021 60th IEEE Conference on Decision and Control (CDC)*, IEEE, 2021, pp. 6970–6975.

[226] S. Alemzadeh, S. Talebi, and M. Mesbahi, "D3pi: Data-driven distributed policy iteration for homogeneous interconnected systems," *arXiv preprint arXiv:2103.11572*, 2021.

[227] D. Ho and J. C. Doyle, "Scalable robust adaptive control from the system level perspective," in *2019 American Control Conference (ACC)*, 2019, pp. 3683–3688. DOI: 10.23919/ACC.2019.8814896.

[228] G. Goel and A. Wierman, "An online algorithm for smoothed regression and lqr control," in *The 22nd International Conference on Artificial Intelligence and Statistics*, PMLR, 2019, pp. 2504–2513.

[229] G. Shi, Y. Lin, S.-J. Chung, Y. Yue, and A. Wierman, "Online optimization with memory and competitive control," *Advances in Neural Information Processing Systems*, vol. 33, pp. 20 636–20 647, 2020.

[230] A. Antoniadis, N. Barcelo, M. Nugent, K. Pruhs, K. Schewior, and M. Scquizzato, "Chasing convex bodies and functions," in *LATIN 2016: Theoretical Informatics*, Springer, 2016, pp. 68–81.

[231] C. Argue, "Chasing convex bodies and functions," Ph.D. dissertation, Carnegie Mellon University, 2022.

[232] S. Mukherjee and T. L. Vu, "Reinforcement learning of structured stabilizing control for linear systems with unknown state matrix," *IEEE Transactions on Automatic Control*, 2022.

[233] Y. Zhang and T. Zhou, "Controllability analysis for a networked dynamic system with autonomous subsystems," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3408–3415, 2016.

[234] A. Gholami and X. A. Sun, "A fast certificate for power system small-signal stability," in *2020 59th IEEE Conference on Decision and Control (CDC)*, IEEE, 2020, pp. 3383–3388.

[235] Y.-S. Wang and N. Matni, "Localized lqg optimal control for large-scale systems," in *2016 American Control Conference (ACC)*, IEEE, 2016, pp. 1954–1961.

[236] J. Yu, Y.-S. Wang, and J. Anderson, "Localized and distributed $\mathcal{H}_2$ state feedback control," in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 2732–2738.

[237] E. D. Sontag, "Input to state stability: Basic concepts and results," in *Nonlinear and optimal control theory*, Springer, 2008, pp. 163–220.

[238] Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.

[239] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.

[240] C. Argue, S. Bubeck, M. B. Cohen, A. Gupta, and Y. T. Lee, "A nearly-linear bound for chasing nested convex bodies," in *Proc. 30th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2019, pp. 117–122.

[241] S. Tu and B. Recht, "The gap between model-based and model-free methods on the linear quadratic regulator: An asymptotic viewpoint," in *Conference on Learning Theory*, PMLR, 2019, pp. 3036–3083.

[242] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, "Optimism-based adaptive regulation of linear-quadratic systems," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1802–1808, 2020.

[243] Y.-C. Ho *et al.*, "Team decision theory and information structures in optimal control problems–part i," *IEEE Transactions on Automatic control*, vol. 17, no. 1, pp. 15–22, 1972.

[244] P. Shah and P. A. Parrilo, "H2-optimal decentralized control over posets: A state-space solution for state-feedback," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3084–3096, 2013.

[245] L. Ye, M. Chi, and V. Gupta, "Regret bounds for learning decentralized linear quadratic regulator with partially nested information structure," *arXiv preprint arXiv:2210.08886*, 2022.

[246] Y. Shi, J. Huang, and B. Yu, "Robust tracking control of networked control systems: Application to a networked dc motor," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 12, pp. 5864–5874, 2012.

[247] D. Ma and J. Zhao, "Stabilization of networked switched linear systems: An asynchronous switching delay system approach," *Systems & Control Letters*, vol. 77, pp. 46–54, 2015.

[248] G. Qu, A. Wierman, and N. Li, "Scalable reinforcement learning of localized policies for multi-agent networked systems," in *Learning for Dynamics and Control*, PMLR, 2020, pp. 256–266.

[249] Y. Lin, G. Qu, L. Huang, and A. Wierman, "Distributed reinforcement learning in multi-agent networked systems," *arXiv*, 2020.

[250] G. Qu, Y. Lin, A. Wierman, and N. Li, "Scalable multi-agent reinforcement learning for networked systems with average reward," *arXiv preprint arXiv:2006.06626*, 2020.

[251] M. Ibrahimi, A. Javanmard, and B. Roy, "Efficient reinforcement learning for high dimensional linear quadratic systems," *Advances in Neural Information Processing Systems*, vol. 25, pp. 2636–2644, 2012.

[252] M. Yazdanian and A. Mehrizi-Sani, "Distributed control techniques in microgrids," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2901–2909, 2014.

[253] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 253–279, 2019.

[254] J. Anderson and N. Matni, "Structured state space realizations for sls distributed controllers," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2017, pp. 982–987.

[255]   P.-Å. Wedin, "Perturbation theory for pseudo-inverses," *BIT Numerical Mathematics*, vol. 13, no. 2, pp. 217–232, 1973.

[256]   Y. G. Rebours, D. S. Kirschen, M. Trotignon, and S. Rossignol, "A survey of frequency and voltage control ancillary services–part i: Technical features," *IEEE Trans. Power Syst.*, vol. 22, no. 1, 2007. DOI: 10.1109/TPWRS.2006.888963.

[257]   H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, 2015. DOI: 10.1109/TPWRS.2015.2504419.

[258]   D. K. Molzahn, F. Dörfler, H. Sandberg, *et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, 2017. DOI: 10.1109/TSG.2017.2720471.

[259]   T. Senjyu, Y. Miyazato, A. Yona, N. Urasaki, and T. Funabashi, "Optimal distribution voltage control and coordination with distributed generation," *IEEE Trans. Power Del.*, vol. 23, no. 2, 2008. DOI: 10.1109/TPWRD.2007.908816.

[260]   C. Gao and M. A. Redfern, "A review of voltage control techniques of networks with distributed generations using on-load tap changer transformers," in *45th International Universities Power Engineering Conference UPEC2010*, Cardiff, UK: IEEE, 2010.

[261]   Y. Guo, Q. Wu, H. Gao, X. Chen, J. Østergaard, and H. Xin, "MPC-Based Coordinated Voltage Regulation for Distribution Networks With Distributed Generation and Energy Storage System," *IEEE Trans. Sustainable Energy*, vol. 10, no. 4, Oct. 2019, ISSN: 1949-3037. DOI: 10.1109/TSTE.2018.2869932.

[262]   S. Maharjan, A. M. Khambadkone, and J. C. H. Peng, "Robust constrained model predictive voltage control in active distribution networks," *IEEE Trans. Sustainable Energy*, vol. 12, no. 1, Jun. 2021, ISSN: 19493037. DOI: 10.1109/TSTE.2020.3001115.

[263]   S. Park, D. Deka, and M. Chcrtkov, "Exact topology and parameter estimation in distribution grids with minimal observability," in *2018 Power Syst. Comput. Conf.*, IEEE, 2018. DOI: 10.23919/PSCC.2018.8442881.

[264]   H. Li, Y. Weng, Y. Liao, B. Keel, and K. E. Brown, "Robust Hidden Topology Identification in Distribution Systems," *arXiv preprint arXiv:1902.01365*, Feb. 2019. DOI: 10.48550/arXiv.1902.01365.

[265]   P. Dehghanian and M. Kezunovic, "Probabilistic impact of transmission line switching on power system operating states," in *2016 IEEE PES Transmiss. Distrib. Conf. Exhib.*, IEEE, 2016, pp. 1–5.

[266] Z. Li, S. Jazebi, and F. De Leon, "Determination of the optimal switching frequency for distribution system reconfiguration," *IEEE Trans. Power Del.*, vol. 32, no. 4, pp. 2060–2069, 2016.

[267] S. Esmaeili, A. Anvari-Moghaddam, S. Jadid, and J. M. Guerrero, "Optimal simultaneous day-ahead scheduling and hourly reconfiguration of distribution systems considering responsive loads," *International Journal of Electrical Power & Energy Systems*, vol. 104, 2019.

[268] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Trans. Autom. Control*, vol. 60, no. 4, Apr. 2015. DOI: 10.1109/TAC.2014.2363931.

[269] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *Proc. 52nd Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA: IEEE, Sep. 2014. DOI: 10.1109/ALLERTON.2014.7028508.

[270] Z. Tang, D. J. Hill, and T. Liu, "Fast distributed reactive power control for voltage regulation in distribution networks," *IEEE Trans. Power Syst.*, vol. 34, no. 1, 2019. DOI: 10.1109/TPWRD.2018.2868158.

[271] G. Qu and N. Li, "Optimal distributed feedback voltage control under limited reactive power," *IEEE Trans. Power Syst.*, vol. 35, no. 1, Jan. 2020. DOI: 10.1109/TPWRS.2019.2931685.

[272] J. Duan, D. Shi, R. Diao, *et al.*, "Deep-reinforcement-learning-based autonomous voltage control for power grid operations," *IEEE Trans. Power Syst.*, vol. 35, no. 1, Jan. 2020. DOI: 10.1109/TPWRS.2019.2941134.

[273] S. Wang, J. Duan, D. Shi, *et al.*, "A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning," *IEEE Trans. Power Syst.*, vol. 35, no. 6, 2020. DOI: 10.1109/TPWRS.2020.2990179.

[274] H. Xu, A. D. Domínguez-García, and P. W. Sauer, "Optimal tap setting of voltage regulation transformers using batch reinforcement learning," *IEEE Trans. Power Syst.*, vol. 35, no. 3, May 2020. DOI: 10.1109/TPWRS.2019.2948132.

[275] Y. Gao, W. Wang, and N. Yu, "Consensus multi-agent reinforcement learning for volt-var control in power distribution networks," *IEEE Trans. Smart Grid*, vol. 12, no. 4, 2021. DOI: 10.1109/TSG.2021.3058996.

[276] X. Sun and J. Qiu, "Two-stage volt/var control in active distribution networks with multi-agent deep reinforcement learning method," *IEEE Trans. Smart Grid*, vol. 12, no. 4, 2021. DOI: 10.1109/TSG.2021.3052998.

[277]   Y. Shi, G. Qu, S. Low, A. Anandkumar, and A. Wierman, "Stability constrained reinforcement learning for real-time voltage control," in *2022 American Control Conference (ACC)*, 2022, pp. 2715–2721. DOI: `10.23919/ACC53348.2022.9867476`.

[278]   J. Feng, Y. Shi, G. Qu, S. H. Low, A. Anandkumar, and A. Wierman, "Stability constrained reinforcement learning for real-time voltage control in distribution systems," *arXiv preprint arXiv:2209.07669*, 2022.

[279]   W. Cui, J. Li, and B. Zhang, "Decentralized safe reinforcement learning for inverter-based voltage control," *Electr. Power Syst. Res.*, vol. 211, 2022.

[280]   V. Kekatos, G. B. Giannakis, and R. Baldick, "Grid topology identification using electricity prices," in *2014 IEEE Power Eng. Soc. Gen. Meet.*, National Harbor, MD, USA: IEEE, 2014. DOI: `10.1109/PESGM.2014.6939474`.

[281]   E. Fabbiani, P. Nahata, G. De Nicolao, and G. Ferrari-Trecate, "Identification of AC distribution networks with recursive least squares and optimal design of experiment," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 4, pp. 1750–1757, 2021.

[282]   J.-S. Brouillon, E. Fabbiani, P. Nahata, K. Moffat, F. Dörfler, and G. Ferrari-Trecate, "Bayesian error-in-variables models for the identification of distribution grids," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1289–1299, Mar. 2023.

[283]   G. Cavraro and V. Kekatos, "Graph algorithms for topology identification using power grid probing," *IEEE Control Syst. Lett.*, vol. 2, no. 4, pp. 689–694, 2018.

[284]   O. Ardakanian, V. W. Wong, R. Dobbe, *et al.*, "On identification of distribution grids," *IEEE Trans. on Control Netw. Syst.*, vol. 6, no. 3, pp. 950–960, Sep. 2019.

[285]   H. Xu, A. D. Domínguez-García, V. V. Veeravalli, and P. W. Sauer, "Data-driven voltage regulation in radial power distribution systems," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2133–2143, May 2020.

[286]   S. Nowak, Y. C. Chen, and L. Wang, "Measurement-based optimal DER dispatch with a recursively estimated sensitivity model," *IEEE Trans. Power Systems*, vol. 35, no. 6, pp. 4792–4802, Nov. 2020.

[287]   D. Deka, V. Kekatos, and G. Cavraro, "Learning distribution grid topologies: A tutorial," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 999–1013, Jan. 2024.

[288]   Y. Chen, Y. Shi, and B. Zhang, "Data-driven optimal voltage regulation using input convex neural networks," *Electr. Power Syst. Res.*, vol. 189, Dec. 2020, ISSN: 03787796. DOI: `10.1016/j.epsr.2020.106741`.

[289]   J.-S. Brouillon, K. Moffat, F. Dörfler, and G. Ferrari-Trecate, "Robust online joint state/input/parameter estimation of linear systems," in *Prof. IEEE 61st Conf. Decis. Control (CDC)*, 2022, pp. 2153–2158.

[290] J. Yu, Y. Weng, and R. Rajagopal, "PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4335–4347, Jul. 2018.

[291] D. Deka, S. Backhaus, and M. Chertkov, "Estimating distribution grid topologies: A graphical learning based approach," in *2016 Power Syst. Comput. Conf.*, Genoa, Italy: IEEE, Jun. 2016, ISBN: 978-88-941051-2-4. DOI: 10.1109/PSCC.2016.7541005.

[292] Y. Sharon, A. M. Annaswamy, A. L. Motto, and A. Chakraborty, "Topology identification in distribution network with limited measurements," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, 2012. DOI: 10.1109/ISGT.2012.6175638.

[293] Y. Li*, J. Yu*, L. Conger, T. Kargin, and A. Wierman, "Learning the uncertainty sets for control dynamics via set membership: A non-asymptotic analysis," in *Forty-first International Conference on Machine Learning (ICML)*, 2024.

[294] M. Farivar, R. Neal, C. Clarke, and S. Low, "Optimal inverter VAR control in distribution systems with high PV penetration," in *2012 IEEE Power Eng. Soc. Gen. Meet.*, San Diego, CA, USA: IEEE, 2012. DOI: 10.1109/PESGM.2012.6345736.

[295] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, 1989.

[296] M. E. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 735–743, 1989. DOI: 10.1109/61.19266.

[297] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *Proc. IEEE Conf. Decision Control*, Firenze, Italy: IEEE, Dec. 2013, ISBN: 978-1-4673-5717-3. DOI: 10.1109/CDC.2013.6760555.

[298] A. Bernstein and E. Dall'Anese, "Real-time feedback-based optimization of distribution grids: A unified approach," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 3, pp. 1197–1209, 2019.

[299] L. Thurner, A. Scheidler, F. Schäfer, *et al.*, "pandapower – An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, Nov. 2018, ISSN: 0885-8950. DOI: 10.1109/TPWRS.2018.2829021.

[300] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, 2016.

[301] M. ApS, *Mosek optimizer api for python 10.0.46*, 2023. [Online]. Available: https://docs.mosek.com/10.0/pythonapi/index.html.